

Draft 2024

Dynamic
Environmental and Resource
Economics

Larry S. Karp & Christian P. Traeger

UC Berkeley & University of Oslo

Contents

1	The Basics of Dynamic Optimization (Catching Fish)	10
1.1	A Heuristic Derivation of the Euler Equation	10
1.2	The Dynamic Programming Equation	13
1.3	Using the DPE to Obtain the Euler Equation	17
1.4	Approximating the Control Rule	19
1.5	Generalizations	24
1.6	An Analogy	25
1.7	Exercises	26
2	An Application to Climate Policy	28
2.1	The method of Lagrange	29
2.1.1	Additional functional assumptions	34
2.2	The dynamic programming approach	37
2.3	Comments on functional assumptions	43
2.4	Problem set	47
3	The Linear Quadratic Problem and Pollution Control	49
3.1	Taxes versus quotas in a static setting	50
3.1.1	A different perspective	55
3.1.2	Cost variations	57
3.2	The LQ control problem	59
3.2.1	The infinite horizon limit	66
3.3	Two variations	69
3.3.1	Multiplicative noise	69
3.3.2	Risk sensitivity	70
3.4	Taxes versus quotas with stock pollutants	71
3.5	Dynamic investment decisions	80

<i>CONTENTS</i>	3
3.6 Further discussion	80
3.7 Problem set	83
4 Reactions to Risk in a Simple Resource Exploitation Model	85
4.1 A Simple Model	85
4.2 Defining a reduction or increase in risk	87
4.3 Reaction to Risk	89
4.4 A characterization of prudence	91
4.5 Reaction to Increasing Risk*	92
4.6 Insurance against Risk	93
4.7 Common functional forms*	94
4.8 Anticipated Learning	97
4.8.1 Complete learning	97
4.8.2 Remark on Discounting*	101
4.8.3 Partial Learning*	102
4.8.4 Learning under risk neutrality	109
5 Multiperiod Anticipated Learning	115
5.1 Discrete distribution	116
5.2 Learning about Δ	116
5.2.1 Δ is a random variable with unknown distribution	117
5.2.2 The "star information structure"	118
5.2.3 Comparing the two ways of modeling learning.	119
5.3 Describing the optimization problem	120
5.3.1 The DP problem	120
5.3.2 The SP problem	123
5.3.3 Digression: active and passive learning again	124
5.4 Continuous distribution	124
5.4.1 Example 1: Poisson and gamma distributions	125

5.4.2	Example 2: Linear quadratic functions and normal distributions	127
5.4.3	The role of conjugates	132
5.4.4	Criticism of these examples	133
6	Numerical Methods	134
6.1	An abstract algorithm: function iteration	134
6.2	Approximating a function	137
6.3	Approximation and value function iteration	141
6.4	Choosing a basis functions and interpolation nodes	142
6.5	Higher dimensional state space	146
6.6	Other considerations	151
6.6.1	Uncertainty	151
6.6.2	Approximation interval	152
6.6.3	More on infinite horizon problems	154
6.6.4	Various	155
6.7	Other functional equations	158
6.8	Finite Markov chains	159
6.9	Comments on Markov chains	163
6.10	Application to Integrated Assessment	165
6.10.1	Welfare	166
6.10.2	The Climate Economy	167
6.10.3	The Code	169
6.10.4	The Questions	171
6.11	Related literature	172
6.12	Problems	173

7	The Maximum Principle and Pollution Control	174
7.1	Intertemporal Optimization: Motivation	174
7.2	The Maximum Principle: Derivation	177
7.3	The Maximum Principle: Formal statement	181
7.4	Economic Interpretation	184
7.5	Current Value Hamiltonian	186
7.6	Infinite Time Horizon	187
7.7	Stock Pollutants: The Euler Equation	189
7.8	Stock Pollutants: Phase Space Analysis	192
7.9	Stock Pollutants: Local Dynamics in the Neighborhood of the Steady State	197
7.10	Dynamics in the Neighborhood of a Steady State: General Remarks	201
7.11	Stock Pollutants: Comparative Statics and Dynamics	205
8	Nonrenewable Resources: Socially Optimal Extraction	209
8.1	Problem formulation	209
8.2	Euler equation	211
8.3	How to stop?	212
8.4	Power utility – the eternal stretch	216
8.5	Constant absolute risk aversion – running out	219
8.6	Constant absolute risk aversion – a variation	221
8.7	More on CARA utility & terminal conditions	223
8.8	CRRA utility	225
9	Nonrenewable Resources: Market-Based Extraction	228
9.1	Extraction by a Competitive Firm	228
9.2	Competitive Market Equilibrium & Comparison to Social Op- timum	233
9.3	Monopolistic resource extraction	240

9.4	Monopolistic versus socially optimal extraction	243
9.5	Stock-dependent extraction costs and exogenous technological progress	246
9.6	Taxes: Profits versus Revenue	250
10	Non-Convex Control Problems	253
10.1	'Convex' Control Problems	253
10.2	The Shallow Lake Problem	255
10.3	Optimal Control of the Shallow Lake	259
10.4	A Close Up of the Steady States	262
10.5	The Optimal Consumption Paths	266
11	Dynamic Programming in Continuous Time	274
11.1	The Continuous Time Limit of the Dynamic Programming Equation	274
11.2	Derivation of HJB as Sufficient Condition	276
11.3	Relation to the Maximum Principle	277
11.4	Autonomous Problems with Infinite Time Horizon, Example and General Approach	279
12	Event Uncertainty	283
12.1	Preliminaries	283
12.2	Endogenous risk	288
12.3	A jump process	297
12.3.1	The DPE for a jump process	298
12.3.2	The jump process and catastrophic risk	300
12.4	Recent applications	302

13 Stochastic Control	303
13.1 Probability Theory and Stochastic Processes - A Brief Review	303
13.2 Brownian Motion	307
13.3 Stochastic Integration	310
13.4 Ito calculus	314
13.5 Stochastic Differential Equations	317
13.6 Stochastic Dynamic Programming	321
13.7 The Linear Quadratic Example under Uncertainty	324
13.8 Resource Extraction under Uncertainty	327
13.9 Appendix	332
13.9.1 More on Brownian Motion	332
13.9.2 Stochastic Integration	333
13.10 Problems	334
14 Discounting	336
14.1 The social discount rate	336
14.2 The SDR and environmental policy	338
14.3 The positive and the normative perspective	339
14.4 Discounting under Uncertainty	341
14.5 Stochastic Ramsey equation	341
14.5.1 General risk attitude	342
14.5.2 General uncertainty attitude	344
14.5.3 Catastrophic risk	345
14.5.4 The Weitzman-Gollier puzzle	346
14.6 Discounting: Extensions	348
14.6.1 Environmental versus produced consumption	348
14.6.2 Hyperbolic discounting	350
14.6.3 Overlapping generations	352

15 Time Inconsistency	355
15.1 An optimal tax example	356
15.2 Two definitions	358
15.3 The durable goods monopoly	359
15.3.1 An extension to convex costs	366
15.3.2 An extension to depreciation	366
15.4 A nonrenewable resource	368
15.5 Disadvantageous power	371
15.6 Problems	374
16 Disentangling RRA from IES	376
16.1 An Intertemporal Model of Risk Attitude	376
16.2 Disentangling Risk Aversion and Intertemporal Substitutability	381
16.3 Intertemporal Risk Aversion	389
16.4 Preference for the Timing of Risk Resolution	391
17 Ambiguity	394
17.1 Introduction to ambiguity	394
17.2 The smooth ambiguity model	396
17.3 The social discount rate under ambiguity	398
17.4 Epilogue	402
A Dynamic Programming - Formal Background	408
A.1 Two Statements of the Dynamic Optimization Problem	408
A.2 Establishing Equivalence	410
A.3 Banach's Contraction Mapping	412
A.4 Applicability of the Contraction Mapping	415
A.5 Existence of the Dynamic Programming Solution	417
A.6 Further Properties of the Solution	419

B	The Basics of Ordinary Differential Equations	420
B.1	Elements of ODEs	421
B.2	Phase portrait analysis	425
B.3	Linear approximations	430
B.4	Back to the predator-prey model	438
B.5	More on saddle points	438
B.6	Examples of dynamic systems in economics	442
	B.6.1 Population dynamics on Easter Island	442
	B.6.2 Multiple rational expectations equilibria	443
	B.6.3 History and expectations with two stock variables	447
B.7	Problems	449

1 The Basics of Dynamic Optimization (Catching Fish)

The Euler equation is the basic necessary condition for optimization in dynamic problems. Here we discuss the Euler equation corresponding to a discrete time, deterministic control problem where both the state variable and the control variable are continuous, e.g. they are members of the real line. Later chapters consider continuous time and stochastic control problems and problems where the control or state variable can take a finite or countably infinite number of values. This chapter also introduces the dynamic programming equation (DPE) as an intermediate step in deriving the Euler equation. Later chapters consider the DPE in a more general setting, and discuss its use in solving dynamic problems. We show that by evaluating the Euler equation in a steady state, and using the condition for local stability of that steady state, we can approximate the optimal control rule in the neighborhood of the steady state and conduct comparative statics exercises. Final sections discuss generalizations to the simple problem that we use to introduce the material, and also discuss an analogy to the dynamic programming problem. An appendix contains a technical treatment of this material.

1.1 A Heuristic Derivation of the Euler Equation

We use a simple version of a fishing model to provide a heuristic derivation of the Euler equation. The stock of fish might be measured by the number of fish or the weight of the population, known as the biomass. Denote the stock of fish at time t as x_t and the harvest at t as h_t . After harvest, the remaining biomass is $y_t = x_t - h_t$, also called the *escapement*. The growth function $F(y_t)$ determines the stock in the next period:

$$x_{t+1} = F(y_t). \tag{1.1}$$

In the problem here, x is the state variable, and equation 1.1 is referred to as the equation of motion, or the state equation. The control variable is h , and y (escapement) is simply a definition.

Profits in the current period, $\pi(x, h)$, depend on the state and the control variables. This formulation allows harvest costs to be stock dependent;

for example, it may be more expensive to catch fish when the stock is low. The price might depend on the harvest or be exogenous and constant. The former case corresponds to the situation where the decision maker faces a downward sloping demand function, either because she is a monopoly or a social planner; in the latter, case, the decision maker is a price taker. The discount factor is $0 \leq \beta < 1$ and the objective of the planner is to maximize the present discounted flow of harvests from the current period, time t , to infinity:

$$\max_{\{h_\tau\}_{\tau=t}^\infty} \sum_{\tau=t}^\infty \beta^{\tau-t} \pi(x_\tau, h_\tau) \quad (1.2)$$

subject to $x_{\tau+1} = F(x_\tau - h_\tau)$, $\tau \geq t$, and x_t given.

For the rest of this chapter we take the initial time to be $t = 0$ in order to simplify notation. The predetermined value of the state variable at the initial time is referred to as the *initial condition*.

This optimization problem is *autonomous*, meaning that it has no dependence on calendar time except via constant discounting. In particular, time enters neither the profit function π nor the growth function F explicitly, but only indirectly because the arguments of these functions change over time. In addition, the horizon for the problem is infinite. For example, if we were to solve problem 1.2 at time $s \neq 0$ given $x_s = x_0$ (i.e. with the same initial condition), we would have exactly the same optimization problem. The solution would also be the same, with the obvious adjustment for time index. We frequently work with autonomous problems, but the methods discussed in this chapter (excluding the steady state analysis) are also applicable for non-autonomous problems.

The first questions that the researcher creating a dynamic model answers are “What are the control variables and what are the state variables?” The next questions are “What is the single period payoff (here, π) and what is the equation of motion (here, $x_{t+1} = F(x_t - h_t)$)?” These are the ingredients of the optimization problem.

We want to find a necessary condition for an optimal solution. The solution to the problem can be written in two ways. We can write it as an infinite sequence of control variables, $\{h_\tau^*\}_{\tau=0}^\infty$, where the $*$ denotes optimality, or we can write it as a function of the state, $h^* = g(x)$, for some function g . The first representation is usually called the *open loop* form and the second

is usually called the *feedback* form. The function g is referred to as the control rule. Note that time is not an argument of this function, because, as we observed above, the solution to an autonomous control problem does not depend on calendar time. For a non-autonomous problem, the control rule typically does depend on calendar time.

Here we work with the open loop representation. If $\{h_\tau^*\}_{\tau=0}^\infty$ is the optimal solution, then it must be the case that any *perturbation* from that solution yields a zero first order welfare change. Because we have not yet established optimality of the trajectory $\{h_\tau^*\}_{\tau=0}^\infty$ we refer to it as a *candidate* for optimality; we also refer to the trajectory of the state variable, when this sequence of control variables is applied, as a candidate trajectory of the state variable.

For an interior optimum of a familiar static problem, optimality requires that the derivative of the maximand with respect to the control is zero. Recall that the first order Taylor approximation of a function, with respect to a control variable, equals the value of that function at a point, plus the derivative of the function with respect to the control variable multiplied by the change in the control variable. Thus, the first order condition in the static problem implies that at an optimum point the first order term in the Taylor approximation of the function is zero. An analogous condition holds in the dynamic problem. The complication in the infinite horizon setting is that we are looking for an infinite sequence rather than a finite number of control variables. Fortunately, we can obtain the necessary condition for optimality by considering a particular change, or perturbation, in the sequence of control variables, in the neighborhood of the candidate.

Suppose that at an arbitrary time t (not necessarily the initial time) we make a slight change in the control h_t relative to the candidate trajectory and then make an offsetting change in the control in the next period, h_{t+1} , so as to leave the state variable x_{t+2} equal to its level under the candidate trajectory. Using the fact that $x_{t+2} = F(F(x_t - h_t) - h_{t+1}) = F(y_{t+1})$, the requirement that the perturbation not alter the value of x_{t+2} implies

$$0 = -\frac{dF(y_{t+1})}{dy} \left(\frac{dF(y_t)}{dy} dh_t + dh_{t+1} \right) \Rightarrow \frac{dh_{t+1}}{dh_t} = -\frac{dF(y_t)}{dy}.$$

If we increase current harvest by ε then we have to decrease next period harvest by $\varepsilon \frac{dF(y_t)}{dy}$ in order to return the state variable to its candidate trajectory. The perturbation alters the level of x_{t+1} , but leaves unchanged the

level of the state at all other times. The perturbation changes the value of only h_t, x_{t+1}, h_{t+1} . The present value at time 0 of the ϵ -perturbation at time t therefore depends only on the payoff during period t and $t + 1$, and is

$$\begin{aligned} & \beta^t \left[\pi_h(x_t, h_t) + \beta \left(\pi_x(x_{t+1}, h_{t+1}) \frac{dx_{t+1}}{dh_t} + \pi_h(x_{t+1}, h_{t+1}) \frac{dh_{t+1}}{dh_t} \right) \right] \epsilon \\ & = \beta^t \left[\pi_h(x_t, h_t) - \beta \left(\pi_x(x_{t+1}, h_{t+1}) + \pi_h(x_{t+1}, h_{t+1}) \right) \frac{dF(y_t)}{dy} \right] \epsilon. \end{aligned}$$

Setting this expression equal to zero yields the Euler equation

$$\pi_h(x_t, h_t) = \beta [\pi_x(x_{t+1}, h_{t+1}) + \pi_h(x_{t+1}, h_{t+1})] \frac{dF(y_t)}{dy}. \quad (1.3)$$

The left side is the marginal profit of an additional unit of harvest in period t . If the perturbation is positive, this term equals the benefit of the perturbation. The right side is the present value at time t of the sum of two terms: the marginal value of an additional unit of stock, $\pi_x(x_{t+1}, h_{t+1})$, times the change in stock due to the extra harvest at t , $-\frac{dF(y_t)}{dy}$, plus the marginal value of an additional unit of harvest at $t + 1$, $\pi_h(x_{t+1}, h_{t+1})$, times the changed harvest at $t + 1$, $-\frac{dF(y_t)}{dy}$. If the perturbation dh_t is positive then the compensating change dh_{t+1} must be negative, so the term on the right hand side is the cost of the perturbation.

1.2 The Dynamic Programming Equation

Here we introduce the dynamic programming equation, retaining the simple version of the fishing problem. Again we take the initial time to be $t = 0$, and the initial condition is $x_0 = x$. We denote the solution to the optimization problem as $V(x)$:

$$V(x) = \max_{\{h_\tau\}_{\tau=0}^\infty} \sum_{\tau=0}^\infty \beta^\tau \pi(x_\tau, h_\tau)$$

subject to $x_{\tau+1} = F(x_\tau - h_\tau)$, $\tau \geq 0$ and $x_0 = x$, given.

The function V depends only on the initial condition of the state variable, x , not on calendar time, because this problem is autonomous. For a non-autonomous problem, the value function includes the argument t .

The *horizon* of a discrete time control problem is the number of periods before the end of the program, which equals the number of remaining periods in which the planner makes a decision. We obtain the dynamic programming equation for the infinite horizon problem above by taking the limit of a sequence of finite horizon problems, as the horizon goes to infinity. A finite horizon problem is not autonomous. For example suppose that the problem begins at the calendar time $t = 2015$, the length of each period is one year, and the problem lasts for 50 years. As calendar time advances, the number of remaining periods diminishes. Therefore, the problem facing a planner in year 2015, at the beginning of the problem, is different from the problem facing the planner at calendar time 2030, even if the value of the state variable happens to be the same at the two points in time. In 2015 the planner has 50 years left, and in 2030 she has 35 years left. The notation must therefore keep track of this difference. There are two ways that we might adapt the notation to achieve this accounting. Probably the most natural-seeming method is to introduce an argument equal to calendar time, and also keep track of the time when the problem begins – 2015 in the example above. The alternative is to keep track of *time-to-go*.

Apart from the constant discounting, the problem depends on time only because of the finite horizon. This fact makes it convenient to use the second alternative, where we keep track of time-to-go rather than calendar time. The use of time-to-go emphasizes the manner in which we actually solve the finite horizon problem, beginning with the final period when the time-to-go is 1 and then working backwards until the first period in calendar time, when (in the example above) the time-to-go is 50 and the calendar time is 2015. This alternative also simplifies the procedure for finding the dynamic programming equation for the infinite horizon, autonomous problem.

At an arbitrary calendar time, we denote the length of the remaining horizon as T , which we take to be finite. We use a superscript T to indicate time-to-go. Thus, $V^T(x)$ is the value of a program with initial condition x when there are T periods to go before the end of the program, and $V(x) = \lim_{T \rightarrow \infty} V^T(x)$. We assume that the limit exists; the appendix provides the conditions under which this assumption is valid.

The fact that the horizon is finite means that we can consider the decision in the last period, where $T = 1$. In our fishing problem, given arbitrary stock

x , the problem is

$$V^1(x) = \max_{h \leq x} \pi(x, h). \quad (1.4)$$

The superscript on the value function denotes the fact that a single period remains. The lack of time subscripts on the arguments of the functions is intentional; it emphasizes that actual calendar time plays no role in the problem. The important data are the value of the current stock and the fact that the current decision is the last one. The notation conveys this information.

However, the absence of time subscripts means that we need some other means of distinguishing between the values of variables in contiguous periods. In the following, we denote the values of the state and control variables at an arbitrary time t as x and h and the values of these variables in the next period as x' and h' . In some contexts we view t as the current time, so that x is the value of the initial condition. However, the notation is general: t is an arbitrary time. With this notation, when there are two periods remaining and the current stock is x , the optimization problem is

$$\begin{aligned} V^2(x) &= \max_{h \leq x \text{ and } h' \leq x'} \pi(x, h) + \beta \pi(x', h') \text{ subject to } x' = F(x - h) \\ &= \max_{h \leq x} \left[\max_{h' \leq x'} \{ \pi(x, h) + \beta \pi(x', h') \} \right] \text{ subject to } x' = F(x - h) \\ &= \max_{h \leq x} \left\{ \pi(x, h) + \max_{h' \leq x'} \beta \pi(x', h') \right\} \text{ subject to } x' = F(x - h) \\ &= \max_{h \leq x} \left\{ \pi(x, h) + \beta V^1(x') \right\} \text{ subject to } x' = F(x - h). \end{aligned}$$

The first equality defines the value function when there are two periods remaining: the maximized value of the program. The second equality breaks the joint maximization over h and h' into two separate stages of maximization. The third equality moves the $\max_{h' \leq x'}$ operator past the function $\pi(x, h)$, taking advantage of the fact that this function does not depend on h' . The fourth equality uses the definition in equation 1.4, replacing the maximization problem inside the brackets with its optimal value, *the function* $V^1(x')$. We need the superscripts on the functions V to keep track of the number of periods remaining.

If there are $T \geq 2$ periods remaining in the fishing problem, we can proceed

inductively to write the dynamic programming problem

$$V^T(x) = \max_{h \leq x} \{ \pi(x, h) + \beta V^{T-1}(x') \} \quad \text{subject to } x' = F(x - h). \quad (1.5)$$

The left side is the value of the program when there are T periods remaining and the stock is x . The equation states that this value equals the maximized value of the sum of the current payoff and the present value of future payoffs, $V^{T-1}(x')$; the last term is often referred to as the *continuation payoff*.

The procedure above enables us to replace the dynamic problem that requires T harvest decisions, with a sequence of T “static” problems, each requiring a single decision. If T is large, the second alternative is more efficient. Of course, there is a cost to considering many small problems rather than a single large problem. The cost is that for each value of $T \geq 2$, each of the small problems involves the *function* $V^{T-1}(x)$. In order to solve any of these problems, we need to solve the prior problem – the one associated with the next time period – in order to construct $V^{T-1}(x)$. We work backwards from the final period in order to construct the solution. Note that it is not enough to know a particular value of $V^{T-1}(x)$; we need the actual function, because in considering the effect of a change in the current control we need to know how that change alters subsequent payoffs.

The discussion here provides the sketch of an algorithm for numerically solving the original dynamic problem, an issue that we return to in Chapter xx. Our objective here is primarily to motivate the dynamic programming equation for the autonomous problem, where $T = \infty$. If $\lim_{T \rightarrow \infty} V^T(x) = V(x)$ exists, then we can formally take limits of both sides of equation 1.5 to write the DPE for the original autonomous problem:

$$V(x) = \max_{h \leq x} \{ \pi(x, h) + \beta V(x') \} \quad \text{subject to } x' = F(x - h). \quad (1.6)$$

An alternative derivation of the DPE does not involve the detour to the finite horizon problem. We simply define $V(x)$ as the value of the program, and recognize that the value of the program today, given state x , equals the maximized value of the current payoff, $\pi(x, h)$, plus the discounted value of the payoff in the future, given that we behave optimally in the future, $\beta V(x')$.

Given the state variable and the control variable, and given the single period payoff and the equation of motion (here π and $x' = F(x - h)$, respectively),

and given the discount factor β , the student should be able to write down the autonomous control problem and the corresponding DPE (see exercise in section 1.7).

1.3 Using the DPE to Obtain the Euler Equation

We continue to use the notation where, for an arbitrary calendar time t (not necessarily the initial time) a prime over a variable indicates the value of that variable in the next period, $t + 1$; the absence of a time subscript on a state or control variable indicates that we evaluate that variable in period t . Continuing with the simple fishery model, here we use the DPE to provide an alternative derivation of the Euler equation. This systematic approach is useful for problems that are too complicated for the heuristic approach taken in section 1.1 to be practical.

The first order condition for the optimization problem on the right side of equation 1.6, assuming an interior optimum, is

$$\pi_h(x, h) = -\beta V_x(x') \frac{dF(x-h)}{dh}.$$

Using the fact that

$$\frac{dF(x-h)}{dh} = -\frac{dF(x-h)}{dx} \quad (1.7)$$

we rewrite the first order condition as

$$\pi_h(x, h) = \beta V_x(x') \frac{dF(x-h)}{dx}. \quad (1.8)$$

The derivative $V_x(x')$ is often referred to as the *shadow value* of the state variable; in other contexts it is referred to as the *costate variable*. Recall that in constrained optimization problems the Lagrange multiplier associated with a constraint is also referred to as the shadow value of that constraint. The derivative $V_x(x')$ equals the increase in the value of the program due to an increase in the level of the state variable.

Applying the envelope theorem to equation 1.6 implies

$$V_x(x) = \pi_x(x, h^*) + \beta V_x(x') \frac{dF(x-h^*)}{dx}. \quad (1.9)$$

Equation 1.9 relates the shadow value at different periods, and is sometime referred to as the equation of motion for the shadow value, or the *costate equation* (because it determines the change in the costate variable). The current shadow value equals the marginal change in the current payoff following a marginal change in the current state variable, plus the marginal change in the state variable in the next period, times the shadow value in the next period, discounted back to the current period. The *'s in equation 1.9 emphasize that the functions are evaluated at the optimal harvest.

The first order condition and the costate equation hold at every period; in particular they hold at periods t and $t + 1$, i.e. in contiguous periods. At an arbitrary time t we have $x_t = x$, and period $t + 1$ is the next period. We can advance both the first order condition and the costate equation 1 period, i.e. evaluate them at time $t + 1$, and write the equations as

$$\pi_h(t+1) = \beta V_x(t+2) \frac{dF(t+1)}{dx} \quad (1.10)$$

$$V_x(t+1) = \pi_x(t+1) + \beta V_x(t+2) \frac{dF(t+1)}{dx} \quad (1.11)$$

Here we abuse notation by writing, for example, $\pi_x(t+1)$ instead of $\pi_x(x_{t+1}, h_{t+1}^*)$; that is we replace the arguments of the various functions by the time index at which the arguments are evaluated. This notation merely simplifies the appearance of the equations. In the interest of clarity we rewrite equation 1.8 using this notation:

$$\pi_h(t) = \beta V_x(t+1) \frac{dF(t)}{dx},$$

which we rewrite as

$$\frac{\pi_h(t)}{\beta \frac{dF(t)}{dx}} = V_x(t+1). \quad (1.12)$$

Substituting equations 1.10 and 1.12 into the costate equation 1.11 we obtain

$$\begin{aligned} \frac{\pi_h(t)}{\beta \frac{dF(t)}{dx}} &= \pi_x(t+1) + \pi_h(t+1) \iff \\ \pi_h(t) &= \beta [\pi_x(t+1) + \pi_h(t+1)] \frac{dF(t)}{dx}. \end{aligned}$$

The last equation is equivalent to the Euler equation we previously obtained, because $\frac{dF(t)}{dx} = \frac{dF(t)}{dy}$.

1.4 Approximating the Control Rule

Up to this point we have emphasized the open loop representation of the solution to the control problem. Here we consider the feedback representation – the control rule that gives the level of the control variable as a function of the state variable. We remain with the autonomous control problem.

A *steady state* is a level of the state variable at which the optimal decision causes the state variable to remain constant. The steady state may not be unique. It is straightforward, at least in principle, to determine the value or the values of the steady state. We can also conduct comparative statics of the steady state, i.e. obtain qualitative information about how the steady state value changes with exogenous parameter changes. However, except for very simple special cases, we cannot obtain analytic solutions to the optimal control rule, and we cannot learn how parameter changes alter the optimal decision outside the steady state. We can, however, use the Euler equation to approximate the optimal control rule in the neighborhood of the steady state, and we can do comparative statics on this approximation. That is, we can determine how an exogenous change in a parameter alters the approximately optimal decision at a level of the state variable near the steady state.

In almost any situation, it would be more interesting to obtain an approximation of the optimal control rule in the neighborhood of the current level of the state variable, the initial condition. If the initial condition is far from the steady state, the approximation of the control rule in the neighborhood of the steady state may not provide much information about the optimal decision in the current period. Thus, the practical value of the analysis in this section depends on the specifics of the problem. In particular, it depends on the distance between the initial condition and the steady state, and on how closely the linear approximation tracks the optimal control rule. We can learn the first piece of information by calculating the steady state and comparing it with the initial condition, but we cannot learn the second piece without solving the dynamic problem – which would of course render the approximation useless.

We emphasize the assumption that the steady state is *locally stable* (also referred to as *asymptotically stable*). Local (or asymptotic) stability in this context means that the optimally controlled state variable approaches the steady state if it begins close to the steady state. This assumption means that if the decision maker uses the optimal decision rule, a function $g(x)$,

and if the state begins in the neighborhood of the steady state, then the trajectory asymptotically (as $t \rightarrow \infty$) approaches the steady state. We can insure that this assumption is satisfied by putting enough structure on the primitives of the model, the single period payoff function and the equation of motion and the discount factor. Alternatively, we may want to assume local stability: if the control rule is not locally stable, then there is not much interest in that particular steady state, because the controlled state would not remain in its neighborhood unless it happens to begin at exactly the steady state. Therefore, there would be little point in learning about the control rule in that neighborhood.

We depart from the fishing problem and consider instead a capital accumulation problem with costly (convex) adjustment. The state variable is the stock of capital, K , and the control variable is the level of investment, I . The equation of motion is

$$K_{t+1} = \delta K_t + I_t, \quad (1.13)$$

with $0 < \delta \leq 1$ and the single period payoff is

$$\pi(K_t) - c(I_t)$$

where the restricted profit function, π , is concave and the investment cost, c , is strictly convex. This convexity reflects the idea that doubling investment in a single period more than doubles costs, i.e. it is more expensive to adjust rapidly than slowly. The discount factor is $0 < \beta < 1$. Now we have the primitives of the problem.

This problem has two features that distinguish it from the fishery problem. First, the equation of motion is linear in both the state and the control. This fact simplifies the local approximation, and is our main reason for departing from the fishery problem; a linear equation of motion does not make much sense for a renewable resource problem. Second, in our formulation of the fishery problem, the stock in the next period depends only on the difference between the state variable and the control. We used that fact in writing equation 1.7, and this equation played a role in writing the Euler equation. In the capital accumulation problem, the state variable depends on the current state and control variables, not just their difference, unless $\delta = 0$.

With these primitives, the Euler equation (see Exercise 3) is

$$c'(I_t) = \beta (\pi'(K_{t+1}) + \delta c'(I_{t+1})). \quad (1.14)$$

By definition, in a steady state, capital and investment are unchanging. Thus, in a steady state, equations 1.13 and 1.14 must hold as algebraic (instead of difference) equations. The steady state conditions are therefore

$$K = \delta K + I, \quad (1.15)$$

$$c'(I) = \beta(\pi'(K) + \delta c'(I)). \quad (1.16)$$

Solving the first equation ($I = (1 - \delta)K$) to eliminate I from the second equation gives the steady state condition

$$c'((1 - \delta)K) = \beta(\pi'(K) + \delta c'((1 - \delta)K)), \text{ or} \quad (1.17)$$

$$q(K) \equiv c'((1 - \delta)K)(1 - \beta\delta) - \beta\pi'(K) = 0. \quad (1.18)$$

Note that

$$q'(K) = (1 - \beta\delta)(1 - \delta)c''((1 - \delta)K) - \beta\pi''(K) > 0.$$

This monotonicity means that there is at most one root to equation 1.18, i.e. at most one steady state. Therefore, if a steady state exists, it is unique. Not all problems of this sort have a unique steady state.

Some comparative statics questions are straightforward. For example, to determine how δ affects the steady state level of K , totally differentiate the steady state condition 1.17 to obtain

$$((1 - \beta\delta)(1 - \delta)c''(I_\infty) - \beta\pi''(K_\infty))dK_\infty = ((1 - \beta\delta)K_\infty c''(I_\infty) + \beta c'(I_\infty))d\delta. \quad (1.19)$$

Here we use the subscript ∞ to emphasize that we evaluate variables in the steady state. When there is no ambiguity, we often drop this subscript. From equation 1.19 conclude that

$$\frac{dK_\infty}{d\delta} > 0;$$

i.e., if capital lasts longer (larger δ) the steady state stock of capital is higher.

A more interesting question is to determine the characteristics of the decision rule in the neighborhood of the steady state. The feedback form of the solution to the Euler equation is a control rule of the form

$$I_t = g(K_t),$$

where $g(\cdot)$ is the control rule. For general functional forms we cannot find the function g in closed form, but we can always approximate it in the neighborhood of the steady state. Write the first order Taylor approximation of the control rule, in the neighborhood of the steady state, as

$$I(K) \approx g(K_\infty) + g'(K_\infty)(K - K_\infty).$$

The trick is to find $g'(K_\infty)$, since everything else on the right side of the equation is known. For example, $g(K_\infty) = I_\infty$, and we know this value from solving the pair of algebraic equations 1.15 and 1.16. Denote $m_t = g'(K_t)$; the object that we want to evaluate is $m_\infty = m(K_\infty)$, the steady state value of m_t .

In order to find the value of $g'(K_\infty)$, replace I_t with $g(K_t)$ in the Euler equation 1.14 and differentiate with respect to K_t . The result is a difference equation involving m_t and m_{t+1} . Next, evaluate this difference equation in the steady state, where $m_t = m_{t+1}$. Denote this common value as m rather than m_∞ in order not to encumber the notation. The result is an algebraic (rather than a difference) equation in m . This equation is

$$z(m) \equiv (- (1 - \beta\delta^2 c'' + \beta\pi'') m + \beta\delta c'' m^2 + \beta\delta\pi'') = 0. \quad (1.20)$$

This equation does not show the arguments of the functions; the notation c'' , for example, means $c''(I_\infty)$.

Equation 1.20 is a quadratic in m . The coefficient of m^2 is positive, the coefficient of m is negative and the intercept ($\beta\delta\pi''$) is negative. From a sketch, one sees that one root of the function $z(m)$ is negative and the other positive. We are interested in a steady state that is locally stable; this means that if the state is in the neighborhood of the steady state, it approaches the steady state. The controlled equation of motion is

$$K_{t+1} = \delta K_t + g(K_t) \equiv L(K_t).$$

In general, a steady state K_∞ , associated with a difference equation of the form

$$K_{t+1} = L(K_t)$$

is stable if and only if

$$|L'(K_\infty)| < 1. \quad (1.21)$$

Figure 1 provides a geometric illustration of this fact. The figure shows the 45 degree line, at which $K_{t+1} = K_t$, as must be the case at a steady state. The figure also shows the graphs of two possible functions $L(K)$. For the lower of these, $L'(K) > 1$ at a steady state; for the higher graph, $0 < L'(K) < 1$ at a steady state. We can use this construction to map out the trajectory given a particular initial condition, as the arrows show. This mapping proceeds as follows. A steady state is the K_t coordinate of the intersection between the function $L(K_t)$ and the the 45 degree line. We pick an arbitrary value of K_t near but not exactly at a steady state. At this initial value of K_t , the value of the state variable in the next period is $L(K)$. We reflect this point off the 45 degree line; that is we move horizontally toward the 45 degree line until hitting it and then move vertically towards the horizontal axis until hitting the axis, to obtain the next value of the state variable. Proceeding in this way, we obtain a sequence of points on the K_t axis. These points are successive values of the state variable.

The sequence of points corresponding to the lower graph diverges from the steady state, so that steady state is unstable. The sequence of points corresponding to the higher graph converge to the steady state, so that steady state is stable. In the lower graph the inequality 1.21 is not satisfied, and in the upper graph the inequality is satisfied. Both of these graphs are increasing functions; an exercise considers the case where the graphs are decreasing.

For our problem

$$L'(K_\infty) = \delta + m,$$

so the stability condition 1.21 can be written as

$$\begin{aligned} -1 < \delta + m < 1 \\ \text{or} \quad -1 - \delta < m < 1 - \delta \end{aligned}$$

An exercise asks the student to confirm that the negative root of $z(m)$ defined in equation 1.20 satisfies the stability condition, and the positive root violates it.

Therefore, we conclude that the negative root is the correct value of m . It is (fairly) straightforward to do comparative statics of the negative root. Thus, we can learn something about the solution to the optimal control problem in the neighborhood of a steady state by studying the algebraic solution to the Euler equation, and using the stability of the steady state.

1.5 Generalizations

In both of the examples in this Chapter, the state variable and the control variable are scalars. In general, the state and the control “variables” can be vectors. The state variable contains all of the information needed to make an optimal decision. The control variable contains the variables that the decision maker chooses. Models used to obtain qualitative insights often involve scalar state and control variables. In some cases, particularly where we use numerical solutions, the state variable and control variable may be vectors. In addition to this practical reason for considering the multi-dimensional case, it is also worth thinking about it in order to help fix ideas about the meaning of the state and the control variables.

To this end, consider the fishing problem where in each period a decision maker chooses the amount of money to invest in new boats and the number of fish to harvest. Suppose that the price of fish changes exogenously and deterministically. The payoff in a period equals revenue (price times harvest) minus the cost of harvest minus the investment. The cost of harvest depends on the current stock of boats (i.e. “fishing capital”) and might also depend on the biomass of fish, because it is cheaper to catch fish when they are abundant. The growth of the fish stock also depends on the biomass of fish. In this problem there are three state variables, the price of fish, the biomass, and the fishing capital. At the beginning of a period, the level of each of these variables is given, i.e. is pre-determined. The price of fish changes exogenously, and the level of biomass and the stock of fishing capital change endogenously. Changes in the latter two stocks depend on current decisions, harvest and investment. Thus, state variables can be either exogenous or endogenous, but at the beginning of a period their levels are predetermined. The control variables in this setting are harvest and investment.

By adding more state and/or control variables, we obtain a more realistic problem. For example, rather than describing the stock of fish using a single variable, it might be important to distinguish age classes, e.g. young and mature fish. In that case, we need two state variables to describe these two stocks, rather than a single variable. If there are n age classes, we need n state variables. In addition, it might be important to keep track of the stock of different types of fish. For example, if the fish being harvested is preyed upon by a second type of fish, the stock level of this second species likely affects the change in the stock of the harvested species. In that case, the

stock of the predator should also be part of the state variable. There may also be other exogenous state variables. For example, the price of fish may depend on income or the prices of other commodities that evolve exogenously, but which affect the payoff to fishing.

As this example suggests, it is easy to make a plausible case for including many variables in the state and the control vectors. However, parsimony is the hallmark of a good model. The optimization problem becomes difficult to solve, even numerically, as the dimension of the state variable increases. Qualitative insights, such as those based on approximations around the steady state, are also harder to obtain in a multi-dimensional setting. We often lose clarity in creating a complex model. Our goal is to create a model that captures enough of the real world to be relevant, and that is simple enough to be tractable. In moving from a description of an economic situation, to a formal dynamic model, the researcher should begin by asking: What is the state variable and what is the control variable for this model?

1.6 An Analogy

Students who have not previously studied explicitly dynamic models, are nevertheless likely to have encountered methods that are analogous to dynamic programming. This section develops that analogy by means of an example. The point of the discussion is help the reader see that dynamic programming is a natural way to approach problems.

Consider an agent who has one unit of time that he divides between leisure, L , and working for wage w . The labor supply decision provides income $y = w(1 - L)$ which he spends on a consumption vector c , obtaining utility $u(c, L)$. The vector of prices is p . We can think of the labor supply and the consumption decisions occurring at the same time, in which case the problem is

$$\max_{L,c} u(c, L) \text{ subject to } pc \leq y = w(1 - L).$$

Equivalently, we can break the problem up into two stages, first choosing labor supply and then choosing consumption. The labor supply decision determines the agent's income, and that variable feeds into the problem of choosing the consumption bundle. In order to determine how much labor to supply, the agent needs to know the benefit of income. In this two-stage problem, we therefore consider the consumption stage first, even though

consumption occurs after supplying labor; that is, we work backwards, just as in the fishing model. The indirect utility function, with arguments income, prices, and leisure, y , p , and L , is

$$v(y, p, L) = \max_c u(c, L) \quad \text{subject to } pc \leq y.$$

The indirect utility function equals the maximized value of utility for a given level of income and leisure. This function is familiar from intermediate theory, although often the argument leisure does not appear. The utility function is primitive (data to the problem), and the indirect utility function is endogenous, i.e. it is induced by the utility function and the operation of maximization. The indirect utility function is analogous to the function $V^1(x)$ that we used in the finite horizon fishing model. The model here corresponds to the case where $T = 2$. The labor supply decision is the solution to

$$\max_L v(w(1 - L), p, L).$$

In this simple problem, there is no obvious advantage to considering the two decisions jointly or to considering them sequentially. (In a slightly more complicated setting, e.g. where prices are a random variable at the time of choosing labor, and known at the time of choosing consumption, there would be a material reason for breaking the problem into two parts.) However, the point of this discussion is to emphasize that the problem is the same, regardless of whether we consider it as one problem over two types of decisions, L and c , or two problems each over one type of decision. In a dynamic setting, this idea of breaking a single problem into smaller parts has much greater power. The relation between the value function V and the single period profit function π and growth function F in the dynamic problem is analogous to the relation between the indirect and direct utility functions in the static problem.

1.7 Exercises

Exercise 1 *Profits from farming depend on soil quality. The evolution of soil quality depends on the application of fertilizer, which is costly. The farmer wants to maximize the present discounted value of profits net of the cost of fertilizer. Define a state and a control variable, a payoff function and an equation of motion. Write down the optimization problem and then write the corresponding DPE.*

Exercise 2 Consider replacing the equation of motion in the fishing problem, $x' = F(x - h)$, with $x' = F(x) - h$. (a) What is the physical interpretation of this change? (b) Derive the Euler equation for the new model using the heuristic approach. Interpret this Euler equation. (c) Write down the DPE for this problem and derive the Euler equation using the DPE.

Exercise 3 (a) Write the maximization problem for the capital accumulation problem, using the definitions of the primitives in Section 1.4. (b) Derive the Euler equation for this problem in both the heuristic method and the DPE. (c) Interpret the Euler equation.

Exercise 4 Confirm that the negative root of equation 1.20 satisfies the stability condition, and the positive root violates it. In order to do this, show that $z(-1 - \delta) > 0 > z(1 - \delta)$. Use a graph to see why these inequalities establish the claim.

Exercise 5 Using the tools developed in Section 1.4 find the comparative statics of the local approximation of the control rule, with respect to δ .

Exercise 6 Both of the graphs in Figure 1 are increasing functions. Use the procedure described in the text to confirm that inequality 1.21 is also necessary and sufficient for stability when the graph of the function $L(K)$ is negatively sloped.

2 An Application to Climate Policy

This section explains two methods of obtaining the necessary conditions for an optimal control problem involving climate policy. The first method, following Golosov et al. (2011), treats the problem as a stochastic programming problem solved using the method of Lagrange. In addition to illustrating how to use this method, the discussion also shows how judicious choice of functional forms can hugely simplify a problem. The second method uses dynamic programming. For the particular functional forms, dynamic programming (DP) offers no advantages over the method of Lagrange. By showing the two approaches side-by-side, we: (i) illustrate that they are merely two routes to the same end; (ii) introduce the reader to a solution method that is often used, but which we do not consider hereafter; and (iii) develop intuition about the meaning of the partial derivatives of a value function. For functional forms more general than those considered here, the DP approach offers computational advantages, as will be apparent later in this series of lectures.

The economy has a final goods sector (1) and an energy sector (2). Agents can consume the final good or save it as capital. Production of this good requires capital, K_1 , labor, N_1 , and energy, E_1 . The stock of carbon, S , creates damages, reducing output of the final good. Production of the final good equals $Y_t = F_{1,t}(K_{1,t}, N_{1,t}, E_{1,t}, S_t)$.

If the economy begins a period with a resource stock R_t and produces E_t units of energy (fossil fuels) in period t , the change in the fossil fuel stock is $R_t - R_{t+1} = E_t$. Energy production create carbon emissions, which alter the carbon stock, S . Sector 2 produces energy, using capital, K_2 , labor, N_2 , and energy, E_2 . The production function for energy is $E_t = F_{2,t}(K_{2,t}, N_{2,t}, E_{2,t}, R_t, R_{t+1})$; the arguments R_t and R_{t+1} allow production costs to be stock-dependent. For example, if the resource stock is oil, production costs increase as stocks fall.

The objective is to maximize the expectation of the present discounted stream of utility,

$$\mathbf{E}_t \sum_{\tau=0}^{\infty} \beta^\tau U(C_{t+\tau}),$$

with $0 < \beta < 1$ the utility discount factor. The constraints are

$$C_t + K_{t+1} = F_{1,t}(K_{1,t}, N_{1,t}, E_{1,t}, S_t) + (1 - \delta) K_t \quad (2.1)$$

$$E_t = F_{2,t}(K_{2,t}, N_{2,t}, E_{2,t}, R_t, R_{t+1}) \quad (2.2)$$

$$S_{t+\tau} = \sum_{s=0}^{T+t+\tau} (1 - d_s) E_{t+\tau-s} \quad (2.3)$$

$$R_{t+1} = R_t - E_t \quad (2.4)$$

$$K_{1,t} + K_{2,t} = K_t \quad N_{1,t} + N_{2,t} = N_t \quad E_{1,t} + E_{2,t} = E_t. \quad (2.5)$$

Equation 2.3 states that the carbon stock S , at time $t + \tau$, is a linear function of emissions from time $-T$ to the current time. This model provides one of many ways to represent the carbon cycle.

The subscripts t on the functions $F_{1,t}$ and $F_{2,t}$ allow for the possibility that there is exogenous change, e.g. due to technological progress. Therefore, this problem is not stationary. These functions may also contain stochastic variables, e.g. random variables that affect technical progress or climate-related damages. We suppress those arguments. We do not need specify the precise stochastic structure, except to assume that the realization of any random variable entering the functions $F_{1,t}$ and $F_{2,t}$ are known at time t . This assumption enables us to move certain functions outside of the expectations operator, \mathbf{E}_t .

2.1 The method of Lagrange

We use the first constraint in equation 2.5 to replace the choice variable $K_{1,t}$, appearing in the function $F_{1,t}$ by $K_t - K_{2,t}$; we use the third constraint in equation 2.5 to eliminate the choice variable $E_{2,t}$, appearing in the function $F_{2,t}$. (This “asymmetric” substitution, eliminating $K_{1,t}$ and $E_{2,t}$, rather than, for example $K_{1,t}$ and $E_{1,t}$ is arbitrary. It leads quickly to the conditions shown in Golosov et al (2011).) The third constraint in equation 2.5 is an implicit function of $E_{t+\tau}$. We cannot eliminate this constraint by substitution. Constraints appearing as implicit functions do not present any problem, but they do affect the appearance of the dynamic programming equation (Section 2.2). Here, we merely note that some constraints can and others cannot be eliminated by substitution.

The variables $\beta^\tau \lambda_{1,t+\tau}$, $\beta^\tau \lambda_{2,t+\tau}$, $\beta^\tau \zeta_{t+\tau}$, and $\beta^\tau \mu_{t+\tau}$ are the present value Lagrange multipliers for the constraints in equations 2.1 – 2.4. The Lagrangian for the problem at time t is

$$\begin{aligned}
\mathbf{L} = & \max \mathbf{E}_t \sum_{\tau=0}^{\infty} \beta^\tau \{ U(C_{t+\tau}) \\
& + \lambda_{1,t+\tau} (F_{1,t+\tau} (K_{t+\tau} - K_{2,t+\tau}, N_{1,t+\tau}, E_{1,t+\tau}, S_{t+\tau}) + (1 - \delta) K_{t+\tau} - C_{t+\tau} - K_{t+\tau+1}) \\
& + \lambda_{2,t+\tau} (F_{2,t+\tau} (K_{2,t+\tau}, N_{2,t+\tau}, E_{t+\tau} - E_{1,t+\tau}) - E_{t+\tau}) \\
& + \zeta_{t+\tau} \left(\sum_{s=0}^{T+t+\tau} (1 - d_s) E_{t+\tau-s} - S_{t+\tau} \right) \\
& + \mu_{t+\tau} (R_{t+\tau} - E_{t+\tau} - R_{t+\tau+1}) \}.
\end{aligned} \tag{2.6}$$

Each of the last four lines of the Lagrangian, \mathbf{L} , corresponds to one of the four constraints in equations 2.1 – 2.4. The expectation operator and β^τ apply to each of these constraints. The interpretation of the Lagrange multipliers is worth noting. $\lambda_{1,t} = \frac{\partial \mathbf{L}}{\partial F_{1,t}}$ is the shadow value at t of an additional unit of the final good. $\lambda_{2,t} = \frac{\partial \mathbf{L}}{\partial F_{2,t}}$ is the shadow value at t of an additional unit of energy supply. For example, the planner would be willing to pay $\lambda_{2,t}$ for one more unit of energy, if that additional unit did not require the reallocation of factors of production or increased emissions or the reduction in the stock of fossil fuels. Similarly, ζ_t is the shadow value at t of the carbon cycle constraint. We show below that it equals the utility loss due to an additional unit of the carbon stock S_t , holding fixed all future stocks; ζ_t is not the shadow value of carbon, because a change in that stock would in fact change subsequent stocks. μ_t is the shadow value of the resource, the amount that the planner would pay for an additional unit of the stock of fossil fuels.

There are two plausible interpretations of this optimization problem. If we think of the planner as choosing the entire sequence of endogenous variables, from the current period to infinity, we obtain the open loop equilibrium. Under uncertainty, this solution is not optimal: the planner can do better by conditioning future decisions on information that becomes available in the future, rather than, for example, selecting the value of C_{t+5} at time t . A planner who implements only the first period decision of the open loop solution, and then resolves the problem in the next period, to obtain the second period decision, uses “*open loop with revision*”. In this scenario, the planner uses new information as it becomes available, and obtains a higher

payoff than if she were to implement the open loop solution obtained in the first period.

However, the open loop with revision outcome is not optimal, because in selecting the first period decision (the one that she actually implements) she behaves as if she will also implement subsequent decisions (the ones that she will actually revise). Thus, although better than the open loop solution, the open loop with revision solution still fails to incorporate all useful information. In particular, in this scenario the planner ignores the fact that she will, in fact, condition future decisions on information that becomes available in the future.

The second interpretation of the optimization problem, and the one that we adopt, is that the planner chooses state contingent actions, i.e. actions based on all information available at the time of implementation. More careful, but substantially more cumbersome notation, would explicitly show that future decisions are state contingent; we use the simpler, but less precise notation. The necessary conditions to this optimization problem do not lead directly to those state contingent policies, but only provide information about their qualities. However, for some functional forms, these necessary conditions make it possible to obtain the state contingent policies.

To simplify the notation, we sometimes suppress the arguments of a function, writing, for example, $\frac{\partial F_{1,t+1}(t+1)}{\partial K}$, or $\frac{\partial F_{1,t+1}}{\partial K}$ or (where the additional notation might promote clarity), $\frac{\partial F_{1,t+1}}{\partial K_{1,t}}$ as shorthand for

$$\frac{\partial F_{1,t+1}(K_{t+1} - K_{2,t+1}, N_{1,t+1}, E_{1,t+1}, S_{t+1})}{\partial K_{t+1}}.$$

The reader knows that $F_{1,t+1}$ is a function of $K_{1,t+1}$, which we set equal to $K_{t+1} - K_{2,t+1}$ in order to eliminate a constraint. A change in K_t therefore creates a change in $K_{1,t}$.

The first order conditions with respect to C_t and K_{t+1} are¹

$$\begin{aligned} \mathbf{E}_t [U'(C_t) - \lambda_{1,t}] &= 0 & (2.7) \\ \mathbf{E}_t \left[-\lambda_{1,t} + \beta \lambda_{1,t+1} \left(\frac{\partial F_{1,t+1}(t+1)}{\partial K} + 1 - \delta \right) \right] &= 0 \implies \end{aligned}$$

¹In some places, e.g. equation 2.7, we leave non-random variables under the expectations operator, merely to give the equations a more unified appearance.

$$\begin{aligned} \lambda_{1,t} &= \mathbf{E}_t \left[\beta \lambda_{1,t+1} \left(\frac{\partial F_{1,t+1}(t+1)}{\partial K} + 1 - \delta \right) \right] \implies \\ U'(C_t) &= \beta \mathbf{E}_t \left[U'(C_{t+1}) \left(\frac{\partial F_{1,t+1}(t+1)}{\partial K} + 1 - \delta \right) \right]. \end{aligned} \quad (2.8)$$

We can take $U'(C_t)$ out of the expectations operator because C_t is a choice variable, not a random variable, at time t . However, C_{t+1} will be conditioned on information not available at time t , and therefore it is a random variable at time t . Equation 2.8 is the Euler equation for consumption, and has the standard interpretation: perturbing an optimal program by saving one more unit today and changing next period consumption in order to leave the subsequent level of capital unchanged, has a zero first order effect on welfare; the marginal cost of the perturbation equals the marginal benefit. Saving rather than consuming the marginal unit of output today reduces current consumption by one unit, so the cost of the perturbation is $U'(C_t)$. The present value of an additional unit of consumption next period equals $\beta U'(C_{t+1})$, and the additional unit of capital makes it possible to consume $\frac{\partial F_{1,t+1}(t+1)}{\partial K} + 1 - \delta$ additional units next period, while leaving the period $t+2$ capital stock at its original (before the perturbation) level. The right side of equation 2.8 is the expected present value benefit of the perturbation.

The first order conditions for $E_{1,t}$, E_t , and S_t are, respectively,

$$\mathbf{E}_t \left[\lambda_{1,t} \frac{\partial F_{1,t}(t)}{\partial E} - \lambda_{2,t} \frac{\partial F_{2,t}(t)}{\partial E} \right] = 0 \quad (2.9)$$

$$\mathbf{E}_t \left[-\lambda_{2,t} \left(1 - \frac{\partial F_{2,t}(t)}{\partial E} \right) + \sum_{\tau=0}^{\infty} \beta^\tau \zeta_{t+\tau} (1 - d_\tau) - \mu_t \right] = 0 \quad (2.10)$$

$$\mathbf{E}_t \left[\lambda_{1,t} \frac{\partial F_{1,t}}{\partial S} - \zeta_t \right] = 0 \quad (2.11)$$

Equation 2.11 shows the relation between two shadow values. The shadow value of the carbon constraint, ζ_t , equals the shadow value of the final good, $\lambda_{1,t}$, times the amount by which a change in S alters the supply of the final good. Equation 2.10 implies

$$\mathbf{E}_t \left[\lambda_{2,t} \frac{\partial F_{2,t}(t)}{\partial E} \right] = \mathbf{E}_t \left[\lambda_{2,t} - \sum_{\tau=0}^{\infty} \beta^\tau \zeta_{t+\tau} (1 - d_\tau) + \mu_t \right].$$

Substituting this relation into equation 2.9 and then using equation 2.11 gives

$$\begin{aligned} \lambda_{1,t} \frac{\partial F_{1,t}(t)}{\partial E} &= \mathbf{E}_t [\lambda_{2,t} - \sum_{\tau=0}^{\infty} \beta^{\tau} \zeta_{t+\tau} (1 - d_{\tau}) + \mu_t] \\ &= \lambda_{2,t} + \mu_t + \mathbf{E}_t \left[- \sum_{\tau=0}^{\infty} \beta^{\tau} \lambda_{1,t+\tau} \frac{\partial F_{1,t+\tau}}{\partial S} (1 - d_{\tau}) \right]. \end{aligned} \quad (2.12)$$

The constraint multipliers $\lambda_{2,t}$ and μ_t are known (not random) at time t , but the stream of future damages depends on future realizations of random variables, and is therefore random at time t . Using equation 2.7, we express the expectation of the present discounted value of the utility cost of an additional unit of emissions, denoted λ_t^s , as

$$\begin{aligned} \lambda_t^s &= -\mathbf{E}_t [\sum_{\tau=0}^{\infty} \beta^{\tau} \zeta_{t+\tau} (1 - d_{\tau})] = \\ &= -\mathbf{E}_t \sum_{\tau=0}^{\infty} \beta^{\tau} \lambda_{1,t+\tau} \frac{\partial F_{1,t+\tau}}{\partial S} (1 - d_{\tau}) = \\ &= -\mathbf{E}_t \sum_{\tau=0}^{\infty} \beta^{\tau} U' (C_{t+\tau}) \frac{\partial F_{1,t+\tau}}{\partial S} (1 - d_{\tau}); \end{aligned} \quad (2.13)$$

λ_t^s equals the marginal social cost of carbon, in units of utility.² The distinction between ζ_t and λ_t^s is worth emphasizing; in particular, ζ_t does not equal the shadow value (the marginal social cost) of carbon. We noted above that ζ_t is the shadow value of the carbon constraint, the change in utility due to a marginal reduction in the current carbon stock, holding all else fixed (equation 2.11). The function λ_t^s shows the utility cost of an additional unit of carbon emissions in t . These emissions increase $S_{t+\tau}$ for all $\tau \geq 0$, thus decreasing output and lowering current and future utility. Note also that λ_t^s is the expectation of a random variable, and is therefore not random. Using the definition of λ_t^s and equation 2.7, we rewrite equation 2.12 as

$$U' (C_t) \frac{\partial F_{1,t}(t)}{\partial E} = \lambda_{2,t} + \lambda_t^s + \mu_t. \quad (2.14)$$

Equation 2.14 is an optimality condition for resource extraction, showing the balance between marginal benefits and costs of an additional unit of emissions. An extra unit of extraction leads to higher production of the final

²Recall our convention for notation: $\frac{\partial F_{1,t+\tau}}{\partial S}$ is shorthand for

$$\frac{\partial F_{1,t+\tau}(K_{t+\tau} - K_{2,t+\tau}, N_{1,t+\tau}, E_{1,t+\tau}, S_{t+\tau})}{\partial S_{t+\tau}}.$$

good and higher current utility. The left side of equation 2.14 equals the marginal benefit of emissions. The right side contains the three components of marginal costs. Recall that $\lambda_{2,t}$ equals the shadow value of the energy constraint, the amount the planner would pay for an additional unit of energy, if that could be obtained without reallocating factors or changing emissions; λ_t^s equals the social marginal cost of emissions; and μ_t equals the shadow value of the stock of fossil fuels. Optimal extraction requires balancing all of these considerations.

To convert the marginal social cost of carbon from units of utility to units of current consumption, we divide by the current (time t) marginal utility:

$$\Lambda_t^s = \frac{\lambda_t^s}{U'(C_t)} = -\mathbf{E}_t \sum_{\tau=0}^{\infty} \beta^\tau \frac{U'(C_{t+\tau})}{U'(C_t)} \frac{\partial F_{1,t+\tau}}{\partial S} (1 - d_\tau). \quad (2.15)$$

If a regulator charges the social marginal cost of carbon, Λ_t^s , per unit of emissions, a competitive economy results in the first best trajectory for consumption, savings, and emissions. The regulator can levy the tax on either fuel producers or users. This type of result is familiar: in the presence of an externality or some other market failure, the socially optimal program can be decentralized, i.e. supported as a competitive equilibrium, by appropriate taxes. The proof of this assertion proceeds by showing that the equilibrium conditions for a competitive economy with emissions tax Λ_t^s , are identical to the optimality conditions of the social planner who chooses all variables (consumption, savings, and emissions).

The only externality in this problem arises from stock-related environmental damages. Therefore, an emissions tax that varies with the state of the economy suffices to correct the externality. The correction of this externality moves us to a first-best world, so it is not necessary to use other policies, such as investment taxes or subsidies. If, for some reason, it is not feasible to implement the first best emissions tax, a second best policy, such as an investment tax or subsidy, could improve welfare.

2.1.1 Additional functional assumptions

We now make two assumptions about functions, leading to a simpler expression for Λ_t^s :

$$U(C) = \ln C \text{ and } Y_t = F_{1,t}(K_1, N_1, E_1, S) = \exp(-\gamma_t(S - \bar{S})) \tilde{F}_{1,t}(K_1, N_1, E_1),$$

where γ_t is a possibly random but exogenous parameter. With these functional forms $\frac{\partial F_{1,t+\tau}}{\partial S} = \gamma_t Y_t$. Define the savings rate as s_t , so $C_t = (1 - s_t) Y_t$, or $\frac{Y_t}{C_t} = \frac{1}{1-s_t}$. With these definitions and assumptions, using equation 2.15

$$\begin{aligned}\Lambda_t^s &= C_t \mathbf{E}_t \sum_{\tau=0}^{\infty} \beta^\tau \gamma_{t+\tau} \frac{Y_{t+\tau}}{C_{t+\tau}} (1 - d_\tau) \\ &= (1 - s_t) Y_t \mathbf{E}_t \sum_{\tau=0}^{\infty} \beta^\tau \gamma_{t+\tau} \frac{1}{1 - s_{t+\tau}} (1 - d_\tau).\end{aligned}$$

We can take $C_t = (1 - s_t) Y_t$ outside the expectations operator because C_t is chosen at time t , and thus is non-random. The *future* savings rates are the only endogenous variables under the expectation operator.

Dividing Λ_t^s by current output, Y_t gives the social marginal cost of carbon in units of consumption (Λ_t^s), per unit of output:

$$\hat{\Lambda}_t^s = \frac{\Lambda_t^s}{Y_t} = - (1 - s_t) \mathbf{E}_t \sum_{\tau=0}^{\infty} \beta^\tau \gamma_{t+\tau} \frac{1}{1 - s_{t+\tau}} (1 - d_\tau).$$

If it happens to be the case that the savings rate is constant, i.e. $s_t = s_{t+\tau}$ for all τ , then the consumption equivalent of the social marginal cost of carbon, per unit of output (hereafter, simply “social cost of carbon per unit of output”), simplifies to

$$\hat{\Lambda}_t^s = -\mathbf{E}_t \sum_{\tau=0}^{\infty} \beta^\tau \gamma_{t+\tau} (1 - d_\tau).$$

If the distribution of the damage parameter, $\gamma_{t+\tau}$, is stationary, i.e. independent of calendar time, then $\mathbf{E}_t \gamma_{t+\tau} = \bar{\gamma}$, a constant, and

$$\hat{\Lambda}_t^s = \hat{\Lambda}^s = \bar{\gamma} \sum_{\tau=0}^{\infty} \beta^\tau (1 - d_\tau), \quad (2.16)$$

a constant.

Without additional assumptions, the savings rate s_t is not constant. Suppose, in addition to the previous assumptions, that production of the final

good is Cobb Douglas and in addition, extraction is costless up to the capacity constraint.³ With these assumptions,

$$Y_t = \exp(-\gamma_t (S_t - \bar{S})) A_t K_t^\alpha N_t^{1-\alpha-\nu} E_{1,t}^\nu.$$

Here, the final goods sector uses all labor and capital, so we do not use a 1 subscript on K, N . In addition, assume that capital fully depreciates in a period: $\delta = 1$. The Euler equation for consumption (2.8) simplifies (using $\frac{K_{t+1}}{s_t Y_t} = 1$) to

$$\begin{aligned} \frac{1}{C_t} &= \beta \mathbf{E}_t \left[\frac{\alpha}{C_{t+1}} (\exp(-\gamma_{t+1} (S_{t+1} - \bar{S}))) A_{t+1} K_{t+1}^{\alpha-1} N_{t+1}^{1-\alpha-\nu} E_{1,t+1}^\nu \right] \\ &\implies \\ \frac{1}{(1-s_t) Y_t} &= \beta \mathbf{E}_t \left[\frac{\alpha}{(1-s_{t+1}) Y_{t+1}} \frac{K_{t+1}}{s_t Y_t} (\exp(-\gamma_{t+1} (S_{t+1} - \bar{S}))) A_{t+1} K_{t+1}^{\alpha-1} N_{t+1}^{1-\alpha-\nu} E_{1,t+1}^\nu \right] \\ &\implies \\ \frac{1}{(1-s_t) Y_t} &= \beta \mathbf{E}_t \left[\frac{\alpha}{(1-s_{t+1}) Y_{t+1}} \frac{Y_{t+1}}{s_t Y_t} \right] = \frac{\alpha \beta}{s_t Y_t} \mathbf{E}_t \left[\frac{1}{(1-s_{t+1})} \right] \\ &\implies \\ \frac{s_t}{(1-s_t)} &= \alpha \beta \mathbf{E}_t \left(\frac{1}{(1-s_{t+1})} \right). \end{aligned}$$

The unique solution to the optimization problem solves the Euler equation and a transversality condition (see Appendix). This unique solution is the constant savings rule, $s_t = s = \beta \alpha$. The savings rate is the product of the discount factor and capital's share in the value of output

Given the assumptions above, the optimal savings rate is constant, and equals $\alpha \beta$. If the distribution of the damage parameter does not depend on calendar time, then $\mathbf{E}_t \gamma_{t+\tau} = \mathbf{E} \gamma$, a constant. With this assumption and using the fact that constant savings rate is constant, the optimal emissions tax is a constant fraction of output:

$$\Lambda_t^s = \Lambda^s = \alpha \beta Y_t \left(\sum_{\tau=0}^{\infty} \beta^\tau (1-d_\tau) \right) \mathbf{E} \gamma. \quad (2.17)$$

³Golosov et al. also study a model with an infinite stock of the resource (e.g. coal) where each unit extracted requires a constant amount of labor.

2.2 The dynamic programming approach

This section reproduces the necessary conditions already obtained using the method of Lagrange. The purpose of this exercise is to clarify the relation between two methods of approaching a dynamic problem, and to develop intuition concerning the meaning of the derivatives of the value function.

To use DP, we begin by identifying the state and control variables to the problem. Denote the vector $E^t = (E_t, E_{t-1}, \dots, E_0, E_{-1}, \dots, E_{-T})$, the $t + 1 + T$ dimensional vector consisting of all emissions from time $-T$ to the current time, t . The passage of each period increases the dimension of the vector by one. At time t , the state variable is K_t, R_t, E^{t-1} . Here we suppress N_t , labor supply, because we saw that it does not play an interesting role in this model. This variable may be important in calibrating the model, but that is not our objective here. As noted above, the time indices on the production functions capture exogenous changes in, for example, technology, and in addition the production functions may depend on random variables, which we suppress. The value function at time t is $J_t(K_t, R_t, E^{t-1})$. This problem is non-stationary, both because of exogenous change in the production functions, and the time-dependent changing dimension of the state variable. We assume that the value function and its derivatives exist.

It is important to remember the meaning of the value function and the state variables. The value function equals the maximized value of the program. That value, and typically also the decisions at a point in time, depend on the state variable (here, a vector). In “standard” models of the climate cycle, of the type discussed in Section 2.3, the dimension of the state variable is constant; in those models, a constant number of pieces of information about the past contain all the information needed to make current decisions. The situation is slightly different for the climate model in equation 2.3. Here, the value of the program and the optimal current decisions depend on the entire history of emissions, up to time $-T$. In that model, we need one more piece of information in every period. Every element of E^{t-1} (in general) affects the payoff and the optimal decision at time t ; the decision at time t determines the last element of E^t , part of the state variable at $t + 1$.

Example: The following example illustrates how the optimal policy depends on the history of emissions as opposed to only the current stock of at-

mospheric carbon. Suppose that $d_0 = d_1 = 1 - 0.9$, $d_\tau = 1 - 0.8^\tau$ for $\tau \geq 2$, so that emissions do not decay at a constant rate. Consider two different histories of emissions, that yield the same current stock of carbon. One trajectory is $E^{t-1} = (1, 1, 1, 1, \dots, 1)$ and the second is $E^{t-1} = (0.7, 1.1375, 1, 1, \dots, 1)$. A calculation confirms that the stock of carbon at t resulting from the first emission path and emissions E_t in period t

$$0.9 + 0.9E_t + \sum_{s=2}^{t-1+T} (0.8)^s + 0.9E_t = 4.1 - 5.0 \times 0.8^{T+t} + 0.9E_t$$

is the same as the period t carbon stock resulting from the second emission history

$$0.9 + 0.9E_t(.7) + 0.8(1.1375) + \sum_{s=3}^{t-1+T} (0.8)^s + E_t = 4.1 - 5.0 \times 0.8^{T+t} + 0.9E_t$$

Thus, regarding the time t stock, the planner is indifferent between the two histories. However, the stock in the next period equals

$$S_{t+1} = 0.9(E_{t+1} + E_t) + \sum_{s=3}^{t+T} (0.8)^s$$

under the first history, and equals

$$S_{t+1} = 0.9(E_{t+1} + E_t) + (0.8)^3(0.7) + (0.8)^4(1.1375) + \sum_{s=5}^{t+T} (0.8)^s$$

under the second history. These two stocks differ by

$$(0.8)^3(0.7) + (0.8)^4(1.1375) - [(0.8)^3 + (0.8)^4] = -0.09728,$$

so the planner at $t + 1$ prefers the second history, the one that leads to a lower current stock, for any current emissions level. In addition, she would (typically) choose different emissions levels under the two histories, even if the planner at t had chosen the same emissions level. The planner at t understands this fact, and therefore would (typically) choose different emissions levels under the two histories.

The dynamic programming equation is

$$J_t(K_t, R_t, E^{t-1}) = \max_{C_t, K_{2,t}, E_{1,t}, E_t} \{U(C_t) + \beta \mathbf{E}_t J_{t+1}(K_{t+1}, R_{t+1}, E^t) + \lambda_{2,t}(F_{2,t}(K_{2,t}, E_t - E_{1,t}) - E_t)\} \quad (2.18)$$

with

$$\begin{aligned} K_{t+1} &= F_{1,t}(K_t - K_{2,t}, N_{1,t}, E_{1,t}, S_t) - (1 - \delta)K_t - C_t \\ R_{t+1} &= R_t - E_t. \end{aligned}$$

Because the constraint associated with energy, in the second line of equation 2.18, is an implicit function, we cannot use it to eliminate a choice variable. Instead, we attach the constraint, using a Lagrange multiplier, to the DPE problem. Had we not explicitly attached this constraint, we would have had to note it's presence beneath the max operator. One way or the other, the problem must include the constraint. We denote the Lagrange multiplier for the energy sector as $\lambda_{2,t}$, as in Section 2.1, in order to facilitate comparison between the two solution methods.

We first obtain the Euler equation for consumption. The only difference in the procedure here, relative to the fishery problem considered in Chapter \label{dpe-discrete}, is that here the problem is stochastic. We have to be careful with the expectations operator. The first order condition with respect to C_t is

$$U'(C_t) = \beta \mathbf{E}_t \frac{\partial J_{t+1}}{\partial K}. \quad (2.19)$$

As above, we conserve notation by using the shorthand

$$\frac{\partial J_{t+1}}{\partial K} = \frac{\partial J_{t+1}(K_{t+1}, R_{t+1}, E^t)}{\partial K_{t+1}}.$$

When there is a risk of confusion, we write $\frac{\partial J_{t+1}}{\partial K_{t+1}}$ instead of $\frac{\partial J_{t+1}}{\partial K}$. The envelope theorem, applied to the DPE 2.18, implies

$$\frac{\partial J_t}{\partial K_t} = \beta \mathbf{E}_t \frac{\partial J_{t+1}}{\partial K_{t+1}} \left(\frac{\partial F_{1,t}}{\partial K_t} + 1 - \delta \right) = \left(\frac{\partial F_{1,t}}{\partial K_t} + 1 - \delta \right) \beta \mathbf{E}_t \frac{\partial J_{t+1}}{\partial K_{t+1}}.$$

Advancing this equation one period, and recognizing that the planner has new information at time $t + 1$ and conditions her decision on that information, gives

$$\frac{\partial J_{t+1}}{\partial K_{t+1}} = \left(\frac{\partial F_{1,t+1}}{\partial K_{t+1}} + 1 - \delta \right) \beta \mathbf{E}_{t+1} \frac{\partial J_{t+2}}{\partial K_{t+2}} = \left(\frac{\partial F_{1,t+1}}{\partial K_{t+1}} + 1 - \delta \right) U'(C_{t+1}),$$

where the second equality uses the first order condition 2.19, evaluated at time $t+1$. The right side of this equation contains the expectations operator \mathbf{E}_{t+1} , because the decision at time $t+1$ is conditioned on information available at that time. Substituting this result into equation 2.19, gives the Euler equation for consumption

$$U'(C_t) = \beta \mathbf{E}_t U'(C_{t+1}) \left(\frac{\partial F_{1,t+1}}{\partial K_{t+1}} + 1 - \delta \right),$$

shown previously as equation 2.8. The time $t+1$ decisions are conditioned on information available at time $t+1$. The optimal time t decisions depend on the information, available at time t , about the time $t+1$ decisions.

Next, we show that the remaining necessary conditions reproduce those obtained using the method of Lagrange. In particular, we want to show that the social cost of carbon reproduces equation 2.13 and that the optimality condition for extraction reproduces equation 2.14. The first order conditions with respect to $K_{2,t}$, $E_{1,t}$ and E_t are, respectively,

$$\begin{aligned} -\frac{\partial F_{1,t}}{\partial K_t} \beta \mathbf{E}_t \left(\frac{\partial J_{t+1}}{\partial K_{t+1}} \right) + \lambda_{2,t} \frac{\partial F_{2,t}}{\partial K_{2,t}} &= 0, \\ \frac{\partial F_{1,t}}{\partial E_{1,t}} \beta \mathbf{E}_t \frac{\partial J_{t+1}}{\partial K_{t+1}} &= \lambda_{2,t} \frac{\partial F_{2,t}}{\partial E_{2,t}}, \end{aligned}$$

and

$$\beta \mathbf{E}_t \left(\frac{\partial J_{t+1}}{\partial K_{t+1}} \frac{\partial F_{1,t}}{\partial S_t} (1 - d_0) + \frac{\partial J_{t+1}}{\partial E_t} - \frac{\partial J_{t+1}}{\partial R_{t+1}} \right) = \lambda_{2,t} \left(1 - \frac{\partial F_{2,t}}{\partial E_{2,t}} \right). \quad (2.20)$$

An increase in E_t increases S_t by $1 - d_0$ units, leading to a decrease in output (and a corresponding decrease in next period capital stock) of $\frac{\partial F_{1,t}}{\partial S_t} (1 - d_0)$. The increase in E_t also increases the final element in the vector E^t , a state variable in period $t+1$. The additional extraction of fossil fuels also decreases by one unit the $t+1$ stock of fossil fuels, another state variable in period $t+1$. The three terms on the left side of equation 2.20 measure the present value expected effect of these changes. The right side of the equation equals the increase in energy available for the final goods sector, due to the increased extraction of fossil fuels, times the shadow value of this increase.

With a view to emphasizing the correspondence between the two solution methods, and also in order to develop intuition about the meaning of the

partial derivatives of the value function, we *define* $\lambda_{1,t} = \beta \mathbf{E}_t \left(\frac{\partial J_{t+1}}{\partial K_{t+1}} \right)$. With this definition, for both methods, $U'(C_t) = \lambda_{1,t}$; compare equations 2.7 and 2.19. With the method of Lagrange, $\lambda_{1,t} = \frac{\partial \mathbf{L}}{\partial F_{1,t}}$, the shadow value of the final good: the marginal increase in the Lagrangian \mathbf{L} , with respect to an increase in the final good, $F_{1,t}$. With the DP method, $\lambda_{1,t} = \beta \mathbf{E}_t \left(\frac{\partial J_{t+1}}{\partial K_{t+1}} \right)$, the expected present value of an additional unit of capital in the next period, which (because of optimality) equals the marginal value of an additional unit of the final good in the current period. We also define $\beta \mathbf{E}_t \frac{\partial J_{t+1}}{\partial R_{t+1}} = \mu_t$, the shadow value of the resource. In the DP setting $\beta \mathbf{E}_t \frac{\partial J_{t+1}}{\partial R_{t+1}}$ equals the present value expected marginal increase in the continuation payoff, due to an extra unit of the stock. With the method of Lagrange, μ_t equals the marginal increase in the Lagrangian, due to an extra unit of the stock.

With the second definition, and rewriting equation 2.20, we obtain

$$\lambda_{2,t} \frac{\partial F_{2,t}}{\partial E_{2,t}} = \lambda_{2,t} - \underbrace{\beta \mathbf{E}_t \left(\frac{\partial J_{t+1}}{\partial K_{t+1}} \frac{\partial F_{1,t}}{\partial S_t} (1 - d_0) + \frac{\partial J_{t+1}}{\partial E_t} \right)}_{\text{social cost of carbon}} + \mu_t. \quad (2.21)$$

The underlined expression equals the social cost of carbon. The first term equals

$$\beta \mathbf{E}_t \frac{\partial J_{t+1}}{\partial K_{t+1}} \frac{\partial F_{1,t}}{\partial S_t} (1 - d_0) = \beta \frac{\partial F_{1,t}}{\partial S_t} (1 - d_0) \mathbf{E}_t \frac{\partial J_{t+1}}{\partial K_{t+1}} = U'(C_t) \frac{\partial F_{1,t}}{\partial S_t} (1 - d_0),$$

where the last equality uses equation 2.19. This expression equals the utility cost of climate-related damage arising from current emissions. The second term, $\beta \mathbf{E}_t \frac{\partial J_{t+1}}{\partial E_t}$, equals the change in expected present value of the subsequent payoff (the continuation value $J(t+1)$) due to an additional unit of lagged emissions.

Comparing equation 2.21 with equation 2.14 shows that the two are identical if and only if the social cost of carbon, λ_{st} , defined in equation 2.13, actually equals the negative of the underlined expression in equation 2.21 – as we asserted in the previous paragraph. Using equation 2.19, we know that the first term in this expression equals $U'(C_t) \frac{\partial F_{1,t}}{\partial S_t} (1 - d_0)$, so we only need to obtain a formula for $\beta \mathbf{E}_t \frac{\partial J_{t+1}}{\partial E_t}$.

To obtain this formula, we differentiate the DPE 2.18 with respect to E_{t-1} , using the envelope theorem. Recall that E_{t-1} is the first element of the state variable E^{t-1} . A marginal increase in E_{t-1} has two effects. First, it

increases S_t by $1 - d_1$ units, lowering period t output by $\frac{\partial F_{1,t}}{\partial S_t} (1 - d_1)$ and second, it increases the second element of E^t (part of the state variable in period $t + 1$). Using these facts and the envelope theorem,

$$\frac{\partial J_t}{\partial E_{t-1}} = \beta \mathbf{E}_t \left(\frac{\partial J_{t+1}}{\partial K_{t+1}} \frac{\partial F_{1,t}}{\partial S_t} (1 - d_1) + \frac{\partial J_{t+1}}{\partial E_{t-1}} \right).$$

Advancing this equation one time period (because we want an expression for $\frac{\partial J_{t+1}}{\partial E_t}$) gives

$$\frac{\partial J_{t+1}}{\partial E_t} = \beta \mathbf{E}_{t+1} \left(\frac{\partial J_{t+2}}{\partial K_{t+2}} \frac{\partial F_{1,t+1}}{\partial S_{t+1}} (1 - d_1) + \frac{\partial J_{t+2}}{\partial E_t} \right). \quad (2.22)$$

Our next goal is to eliminate $\frac{\partial J_{t+2}}{\partial E_t}$ from equation 2.22 using repeated substitution; we want to write the right side of this equation as the discounted sum of the single period effects of higher lagged emissions. Repeating the procedure that led to equation 2.22, we write $\frac{\partial J_{t+2}}{\partial E_t}$ as

$$\frac{\partial J_{t+2}}{\partial E_t} = \beta \mathbf{E}_{t+2} \left(\frac{\partial J_{t+3}}{\partial K_{t+3}} \frac{\partial F_{1,t+2}}{\partial S_{t+2}} (1 - d_2) + \frac{\partial J_{t+3}}{\partial E_t} \right).$$

Substitution gives

$$\frac{\partial J_{t+1}}{\partial E_t} = \beta \mathbf{E}_{t+1} \left(\frac{\partial J_{t+2}}{\partial K_{t+2}} \frac{\partial F_{1,t+1}}{\partial S_{t+1}} (1 - d_1) + \beta \mathbf{E}_{t+2} \left(\frac{\partial J_{t+3}}{\partial K_{t+3}} \frac{\partial F_{1,t+2}}{\partial S_{t+2}} (1 - d_2) + \frac{\partial J_{t+3}}{\partial E_t} \right) \right).$$

Recall that the law of iterated expectations states that the unconditional time t expectation of an event at (for example) time $t + 2$ equals the time t expectation of the conditional time $t + 1$ expectation of this event. We now use repeated substitution the law of iterated expectations to obtain

$$\frac{\partial J_{t+1}}{\partial E_t} = \mathbf{E}_{t+1} \left(\sum_{s=1}^{\infty} \beta^{s-1} \left(\beta \frac{\partial J_{t+1+s}}{\partial K_{t+1+s}} \right) \frac{\partial F_{1,t+s}}{\partial S_{t+s}} (1 - d_s) \right).$$

We simplify this expression (for $s \geq 1$) as follows:

$$\begin{aligned} \mathbf{E}_{t+1} \beta \frac{\partial J_{t+1+s}}{\partial K_{t+1+s}} \frac{\partial F_{1,t+s}}{\partial S_{t+s}} (1 - d_s) &= \mathbf{E}_{t+1} \mathbf{E}_{t+s} \beta \frac{\partial J_{t+1+s}}{\partial K_{t+1+s}} \frac{\partial F_{1,t+s}}{\partial S_{t+s}} (1 - d_s) \\ &= \mathbf{E}_{t+1} \frac{\partial F_{1,t+s}}{\partial S_{t+s}} (1 - d_s) \mathbf{E}_{t+s} \beta \frac{\partial J_{t+1+s}}{\partial K_{t+1+s}} = \mathbf{E}_{t+1} U'(C_{t+s}) \frac{\partial F_{1,t+s}}{\partial S_{t+s}} (1 - d_s). \end{aligned}$$

The first equality uses the law of iterated expectations, the second uses the fact that $\frac{\partial F_{1,t+s}}{\partial S_{t+s}}(1 - d_s)$ is known at time $t + s$, and the third uses equation 2.19. With this expression, we have

$$\frac{\partial J_{t+1}}{\partial E_t} = \mathbf{E}_{t+1} \left(\sum_{s=1}^{\infty} \beta^{s-1} U'(C_{t+s}) \frac{\partial F_{1,t+s}}{\partial S_{t+s}} (1 - d_s) \right).$$

Consequently,

$$\mathbf{E}_t \frac{\partial J_{t+1}}{\partial E_t} = \mathbf{E}_t \left(\sum_{s=1}^{\infty} \beta^{s-1} U'(C_{t+s}) \frac{\partial F_{1,t+s}}{\partial S_{t+s}} (1 - d_s) \right). \quad (2.23)$$

Using this expression, the underlined term in equation 2.21 simplifies to

$$\mathbf{E}_t \left(\sum_{s=0}^{\infty} \beta^s U'(C_{t+s}) \frac{\partial F_{1,t+s}}{\partial S_{t+s}} (1 - d_s) \right)$$

which equals the negative of the social cost of carbon, $-\lambda_{st}$, given in equation 2.13.

In summary, the method of Lagrange and the DP approach yield the same Euler equation for consumption, the same optimality condition for extraction, and the same expression for the social cost of carbon.

2.3 Comments on functional assumptions

Here we discuss the climate model in equation 2.3, and then the choice of functional forms that lead to closed form expressions for the savings rule and the optimal carbon tax.

Equation 2.3 expresses the current climate variable, S_t , as a linear function of emissions from periods $-T$ through t . This formulation is quite general, because it does not restrict the decay parameters d_i . Golosov et al choose a parameterization of d_i involving three parameters. Gerlagh and Liski (2012) choose d_i in order to match the historical record of emissions releases and carbon accumulation. This model requires that carbon stocks are independent of emissions prior to $-T$. For sufficiently large T , that limitation is unimportant when d_τ approaches 1 as τ becomes large.

However, this model requires one additional element of the state variable, with each additional time period. Due to the closed form expressions for savings and the optimal tax, arising from the functional assumptions described in Section 2.1.1, the requirement of additional information does not lead to an increase in complexity. However, more general functional forms do not lead to closed form decision rules, and require numerical solutions. The difficulty of solving such a model increases rapidly in the dimension of the state variable, a feature known as the “*curse of dimensionality*”. This feature makes parsimonious models desirable.

A more common approach to modeling the climate uses “boxes” to allow past emissions to affect the variable of interest. Each box requires a state variable. For example, in DICE, the average atmospheric temperature in the next period depends on current temperature (one state variable) and on the stock of atmospheric carbon (a second state variable) and on average oceanic temperature (a third state variable). Current emissions change the stock of atmospheric carbon; the next period oceanic temperature depends on the current oceanic and atmospheric temperatures. Climate-related damages depend on the atmospheric temperature. In this model, current emissions affect only future temperatures, and their effect dies out slowly.

Golosov et al.’s formulation achieves the same feature, by appropriate choice of the d_τ parameters. If the three (or finite n) box climate model is linear in state variables, the current temperature can be written as an infinite weighted sum of past emissions, in which the weights are functions of the parameters of the dynamic system. (DICE’s climate model is not linear in the state.) Thus, the linear box model and Golosov et al.’s general formulation are not nested. The former imposes restrictions (arising from the dynamics in the box model) on the coefficients of lagged emissions, whereas Golosov et al.’s general formulation does not. The latter ignores emissions prior to $-T$, whereas the former takes into account all previous emissions. For the purpose of calibrating a model, these differences are likely unimportant. However, the linear box model is parsimonious, requiring only a few state variables to describe the climate. In contrast, the Golosov et al. description of the climate requires a state variable containing $t + T$ variables.

To illustrate the relation between the box model and Golosov et al.’s formulation, consider a one-box linear model, with

$$S_{t+1} = \eta S_t + \rho M_t \text{ and } M_{t+1} = \delta M_t + E_t, \quad (2.24)$$

where S_t is the average atmospheric temperature, as a deviation from pre-industrial levels, M_t is the concentration of atmospheric greenhouse gasses, as a deviation from pre-industrial level, and E_t is emissions, as before. Climate damages depend on S . Emissions decay at a constant rate, $1 - \delta$, and current emissions affect the temperature with a lag of two periods. A one unit increase in emissions at t increases the carbon stock at $t+1$ by one unit, and increases temperature at $t+2$ by ρ units. We assume that $0 < \eta < 1$ and $0 < \delta < 1$, so that the effect of emissions, on future temperature, eventually dies out.

We use the lag operator, L , where $L^s x_t = x_{t-s}$ so rewrite system 2.24 as

$$(1 - \eta L) S_t = \rho L M_t \text{ and } (1 - \delta L) = L E_t$$

or

$$S_t = \frac{\rho L}{(1 - \eta L)} M_t \text{ and } M_t = \frac{L}{(1 - \delta L)} E_t.$$

Substituting the second equation into the first gives

$$S_t = \frac{\rho L}{(1 - \eta L)} \frac{L}{(1 - \delta L)} E_t = \frac{\rho L^2}{(1 - \eta L)(1 - \delta L)} E_t$$

Using the fact that (for $0 < \eta < 1$)

$$\frac{1}{1 - \eta L} = \sum_{s=0}^{\infty} (\eta L)^s,$$

we write the last expression as

$$S_t = \left(\sum_{s=0}^{\infty} (\eta L)^s \right) \left(\sum_{s=0}^{\infty} (\delta L)^s \right) \rho L^2 E_t.$$

This expression shows that S_t depends on emissions during periods $t-2$ and before, as noted above. The product

$$\left(\sum_{s=0}^{\infty} (\eta L)^s \right) \left(\sum_{s=0}^{\infty} (\delta L)^s \right)$$

is a polynomial in L , with coefficients depending on η and δ . Assuming that $\eta \neq \delta$, the coefficient of L^τ equals

$$\sum_{s=0}^{\tau} \eta^s \delta^{\tau-s} = \frac{1}{\eta - \delta} (\eta^{\tau+1} - \delta^{\tau+1}).$$

This one-box model is a special case of the Golosov et al. formulation, with

$$d_0 = d_1 = 1 \text{ and } d_s = 1 - \rho \frac{1}{\eta - \delta} (\eta^{s-1} - \delta^{s-1}) \text{ for } s \geq 2$$

and $T = \infty$.

For example, if the time step equals a decade, and the half life of a unit of atmospheric carbon is 15 decades, then $\delta^{15} = 0.5$, or $\delta \approx 0.955$. Suppose also that at a constant carbon stock, M_∞ , it would take 10 decades for a deviation of the temperature, from its steady state level $S_\infty = \frac{\rho}{1-\eta} M_\infty$, to fall by 50%. This assumption implies $\eta \approx 0.933$. The time profile of $1 - d_s$, the effect of a unit of emissions at $t - s$ on the temperature at t , is more informative than the time profile of d_s . The magnitude of ρ depends on the choice of units, and affects the magnitude but not the profile of $1 - d_s$. We therefore ignore ρ (equivalently, set $\rho = 1$), and show the graph of $1 - d_s$, for $s \geq 2$ in Figure ??, for the values of η and δ given above. For this example, the effect of a unit of emissions at time t on the temperature at time $t + s$ increases over time, (for $s > 2$) over the next 18.6 periods (186 years). The effect of a unit emissions at t , on temperature at $t + s$, is 6.8 times as large in 186 years as in twenty years. Moreover, it takes more than 80 decades for the magnitude of the effect, on temperature, of emissions at t to fall below its effect two decades after release; it takes over 96 decades for the effect to fall to 50% of its two-decade level.

This simple example illustrates several points. The effect of emissions on temperature change is likely non-monotonic with respect to time. Emissions may have negligible effect for decades, after which its effect may increase and remain pronounced for many decades. If we try to approximate the type of time profile shown in figure ??, using the history of emissions, we may need to include many lags – 80 to 100 in the example above. Unless the problem has a closed form solution, it's analysis (using numerical methods) would likely be impractical when we represent the climate using the history of emissions. However, the same model can be represented exactly (without truncating the history) using only two state variables, lagged temperature and carbon stock.

The particular functional assumptions described in Section 2.1.1 lead to a simple solution, making the lack of parsimony of the climate model irrelevant. The important functional assumptions are: utility is logarithmic in

consumption; damages are exponential in S , which is linear in lagged emissions; production of the final good is Cobb Douglas; all capital decays in a period; resource extraction is costless when stocks are positive, and infinite thereafter. Collectively, these assumptions imply that the optimal savings rate is constant, and that fact implies that marginal social damages are a constant proportion of output. With constant marginal damages, a familiar result states that the optimal emissions tax is a constant. Here, marginal damages are a constant fraction of output, leading to a straightforward modification of that result: the optimal emissions tax is constant fraction of output.

The optimal savings rate is not only constant, but also independent of all climate-related concerns: the savings rate depends only on the discount factor, β , and on capital's share of value added, α . A model without any climate variables or resource constraint leads to the same savings rate, under logarithmic utility and Cobb Douglas production. Arguably, the reason for using a general equilibrium model, in which capital is endogenous, to study climate policy, is that climate related damages are sufficiently important that they might affect savings. That is, it might be optimal to change the savings portfolio, leaving (perhaps) more environmental capital (a lower stock of S) and less man-made capital to agents in the future. The functional assumptions in this model rule out this possibility. The closed form solution arises because the functional assumptions eliminate (arguably) the reason for studying the emissions and the savings decisions in a single model.

This observation suggest a modelling alternative: take the savings rule as an exogenous, but not necessarily linear function of output. Then solve a more general model in which damages are not necessarily linear in output, so that the optimal tax is not necessarily a constant fraction of output. With such a model, we can study the relation between the social cost of carbon and the stock of capital and climate-related variables.

2.4 Problem set

Use the linear “one-box” model in equation 2.24 to obtain the Euler equation for consumption and the necessary condition for optimal emissions, using dynamic programming. You need only mimic the procedure used in Section 2.2, making modifications to account for the different climate model. (The

purpose of this exercise is to give students the opportunity to practice the manipulations needed to derive optimality conditions.)

3 The Linear Quadratic Problem and Pollution Control

A famous problem in environmental economics ranks social welfare under taxes and quotas when the policymaker and firms have asymmetric information about abatement costs. Early versions of this problem study the situation where the firm creates a flow pollutant, i.e. one that dissipates quickly. In this case, both the regulator and the firms' problems are static. Under particular functional assumptions, taxes lead to higher welfare than quotas if and only if the slope of the marginal abatement cost curve is greater than the slope of the marginal damage curve. We begin this chapter by discussing the problem of taxes versus quotas in the static setting.

A famous problem in control theory involves the case where the payoff is linear-quadratic in the state and control, and the equation of motion is linear in the state and the control. This problem has a closed form solution. If, in addition to the linear-quadratic structure, the stochastic element enters the equation of motion additively, the problem becomes even simpler. Section 3.2 presents the basics of the linear-quadratic (LQ) control problem. This problem is intrinsically important because it can be applied to many settings, and it is pedagogically useful because it provides one of the simplest examples of a dynamic problem that can be solved in closed form. Most dynamic problems do not admit a closed form solution, but the intuition obtained from solving the LQ problem is useful for understanding algorithms employed in numerical solutions of more complicated problems.

These two problems fit together because the LQ control problem provides a means of generalizing the taxes versus quotas question from a static to a dynamic setting. In the most obvious generalization, and the one that has attracted most attention, the dynamics arise because firms create a stock rather than a flow pollution. A flow pollution is one that dissipates in a fairly short period of time; examples might include particulates and SO_2 . A stock pollutant, in contrast, persists for a long period of time; CO_2 is the quintessential stock pollutant. The LQ control problem thus provides a model for comparing taxes and quotas for the control of greenhouse gasses.

In order to construct a tractable model, stock pollutants are often modeled as if they decay at a constant rate. As with any model, this one is only an approximation to a complex reality. The constant decay rate implies that

the equation of motion for the pollution stock is linear in the stock and in emissions. The model can be calibrated using the half-life of a the pollutant, the amount of time it takes a given quantity to decay into half its initial level. [A table with the approximate half-life of different pollutants would be useful here.] For example, CO_2 has an estimated half-life greater than 80 years.

Although the half-life of a pollutant provides an obvious means of distinguishing between stock and flow pollutants, this measure can be misleading. Pollutants such as particulates disappear rapidly from the atmosphere once emissions cease; due to this short half-life, particulates are typically considered flow pollutants. However, the effects of this kind of pollution may accumulate in people or animals. The effect of an extended period of past exposure may persist long after emissions stop. In this case, if we are interested in the health effects of pollution, a dynamic model is appropriate. Here, however, the stock of the pollution should be understood not as the atmospheric stock, but as a measure of accumulated exposure.

The stock aspect of pollution is the most obvious source of dynamics, but dynamics can also arise because of economic rather than chemical/physical considerations. For example, suppose that firms have convex investment costs, as discussed in Chapter 1. Here, their current investment depends on their belief about the future stream of payoffs. When firms have convex investment costs for abatement capital, their investment depends on their beliefs about future pollution policies. In this case, firms have a dynamic problem even when they create a flow pollutant. The regulator's problem is also dynamic in this case. We briefly consider this kind of problem.

3.1 Taxes versus quotas in a static setting

We begin with the static problem, due to Weitzman (1974). In this setting, firms have private information about a parameter that affects their marginal abatement costs. The regulator knows the distribution but not the realization of this random variable, and must choose between regulating pollution by means of a tax or a quota. We assume that the optimal tax and the optimal quota are both binding for all realizations of the cost parameter: firms facing the quota emit at the quota level, and firms facing the tax emit at a positive level.

Under a tax, the firm emits to the point where marginal abatement costs

equal the tax. Since marginal abatement costs are random (from the perspective of the regulator) emissions are also random. By means of a tax the regulator can choose the expected level of emissions. The firm then “arbitrages emissions over states of nature”, in the sense that the marginal abatement costs are the same in every state of nature. In contrast, under a binding quota on emissions, the firm emits at the same level in every state of nature.

Jensen’s Inequality plays an important role in the analysis of this problem, and many others in economics. Jensen’s inequality states that if $f(\theta)$ is a convex function of θ , a random variable with mean $\bar{\theta}$, then $E f(\theta) \geq f(\bar{\theta})$, with the inequality strict if f is strictly convex. The inequality is reversed if f is a concave function.

In evaluating social welfare under the tax, we always consider tax-exclusive abatement costs. In this setting we treat the tax payments as a pure transfer from firms to the general taxpayer; their cost to firms is offset by their benefit to the public, so their net direct effect on welfare is zero. The tax affects welfare only indirectly, because it affects the level of pollution.

Consider an arbitrary quota (not necessarily the optimal one) and a tax that leads to the same level of emissions in expectation. The ability under the tax to arbitrage emissions over states of nature, i.e. the flexibility of choosing actual emissions, and the lack of this flexibility under the quota, means that the expected (tax exclusive) abatement costs are always lower under the tax than under the quota. However, when damages are convex in emissions, Jensen’s Inequality implies that expected damages under the tax (where emissions are a random variable) is greater than the damages under the quota. Thus, the sources of the relative advantage of the tax over the quota is straightforward: the tax leads to lower expected (tax-exclusive) abatement costs and higher expected environmental damages compared to the quota. With this intuition, we now turn to the specific model which makes it possible to see how parameter values determine the ranking of policies.

Abatement equals the difference between the actual level of emissions and the Business as Usual (BAU) level. Abatement costs are quadratic in abatement, so the benefit of emissions is a quadratic function of emissions. We assume that the intercept of the marginal benefit function equals a constant a plus a mean-zero random variable θ with a constant and known variance σ^2 . The slope of marginal benefits is a known constant b . The firm, but not the

regulator, knows the value of θ . The firm's benefit of emissions is a concave function. An increase in emissions corresponds to a decrease in abatement, and therefore a decrease in abatement costs. In the linear-quadratic setting, the benefit of emissions, e , is

$$\tilde{f} + (a + \theta)e - \frac{b}{2}e^2. \quad (3.1)$$

Under Business as Usual (BAU), the firm chooses emissions, conditional on the realization of θ , to maximize its benefits. Thus, the BAU level of benefits and the BAU level of emissions are the random variables

$$\text{Firm's Benefits}^{\text{BAU}} = \frac{1}{2} \frac{2\tilde{f}b + a^2 + 2a\theta + \theta^2}{b} \quad (3.2)$$

$$e^{\text{BAU}} = \frac{a + \theta}{b}. \quad (3.3)$$

This model is appropriate if there is a single representative firm with cost stock $\theta \sim (0, \sigma^2)$, if there are many firms with this distribution of cost shocks, or many firms each with a particular cost parameter. In the latter case, we model the distribution across firms of the cost parameter using the distribution $\theta \sim (0, \sigma^2)$. In every case except for the representative firm model, different firms have different realized costs. In cases where firms have different realizations of θ , we assume that they are able to trade their emissions quotas, and that the resulting competitive equilibrium is efficient. Therefore, in equilibrium, all firms have the same marginal abatement cost under the quota, but the level of marginal abatement costs differs for different realization of the cost shock, i.e. for different states of nature.

When the regulator sets a tax p per unit of emissions, the firm maximizes the benefit of emissions minus the cost of tax. The firm's problem is

$$\max_e \tilde{f} + (a + \theta)e - \frac{b}{2}e^2 - pe.$$

The first order condition to this problem implies that the level of emissions is

$$e^* = \frac{a - p}{b} + \frac{\theta}{b} \equiv z + \frac{\theta}{b}. \quad (3.4)$$

In order to simplify comparisons with quotas, we can think of the regulator as choosing z , the expected level of emissions under a tax, rather than the

tax, p . Substituting e^* into the firm's benefit function, expression 3.1, and taking expectations, gives the expected benefit of emissions under the tax policy z :

$$\tilde{f} + az + \frac{\sigma^2}{2b} - \frac{b}{2}z^2. \quad (3.5)$$

As we noted above, this expression for expected benefits excludes the firm's tax payments. From the standpoint of the regulator, these tax payments are a pure transfer: they count as a cost to the firm, but a benefit to taxpayers at large, and therefore do not enter the calculation of social benefits.

The damage is quadratic in emissions

$$D(e) = \frac{G}{2}(e - \bar{e})^2$$

where \bar{e} is a known parameter. Under taxes the level of expected damage is

$$E_\theta \frac{G}{2}(e - \bar{e})^2 = E_\theta \frac{G}{2} \left(z + \frac{\theta}{b} - \bar{e} \right)^2 = \frac{G}{2}(z - \bar{e})^2 + \frac{G}{2} \frac{\sigma^2}{b^2}$$

The regulator's problem under taxes is

$$\begin{aligned} \max_z \left[\tilde{f} + az + \frac{\sigma^2}{2b} - \frac{b}{2}z^2 - \left(\frac{G}{2}(z - \bar{e})^2 + \frac{G}{2} \frac{\sigma^2}{b^2} \right) \right] = \\ \left(\tilde{f} + \frac{\sigma^2}{2b^2}(b - G) \right) + \max_z \left[az - \frac{b}{2}z^2 - \left(\frac{G}{2}(z - \bar{e})^2 \right) \right]. \end{aligned} \quad (3.6)$$

Under quotas the regulator chooses emissions, e , rather than expected emissions, z . Because of the assumption that the quota is binding in all states of nature, the single period expected benefits and damages under the quota e are

$$\begin{aligned} \text{expected benefits: } \tilde{f} + ae - \frac{b}{2}e^2 \\ \text{damages: } \frac{G}{2}(e - \bar{e})^2, \end{aligned}$$

and the regulator's problem is

$$\begin{aligned} \max_e \left[\tilde{f} + ae - \frac{b}{2}e^2 - \left(\frac{G}{2}(e - \bar{e})^2 \right) \right] \\ \tilde{f} + \max_e \left[ae - \frac{b}{2}e^2 - \left(\frac{G}{2}(e - \bar{e})^2 \right) \right]. \end{aligned} \quad (3.7)$$

The maximization problems 3.6 and 3.7 are exactly the same, with z replaced by e . Therefore, the optimal values of these two variables, e and z , are the same. This fact is an example of a more general principle, which we define here.

Definition 1 *The “Principle of Certainty Equivalence” states that the optimal decision in a stochastic problem is identical to the optimal decision in the corresponding deterministic problem in which we replace the random term by its expected value. Thus, the optimal decision (but not the value of the program) is independent of the variance and all higher moments of the random term.*

The Principle of Certainty Equivalence does not hold for general problems, and for that reason it is perhaps unfortunate that the result has become known as a “Principle”. However, the Principle of Certainty Equivalence does hold when the objective is a quadratic function of the control variable, the constraint is linear, and the random term enters the problem additively. All of these features hold in this static problem, and also in a dynamic generalization that we consider later.

The equality of the optimal e and the optimal z means that the expected level of emissions under the optimal tax equals the deterministic level of emissions under the optimal quota. Of course, the actual levels differ. Under quotas, the regulator chooses emissions, so this variable is deterministic. Under taxes, the regulator chooses the expected level of emissions, and emissions are random.

The difference between the expected social welfare under taxes and under quotas, obtained by subtracting the last line of equation 3.7 from the last line of equation 3.6, and using the fact that the functions after the two max operators are equal, is

$$\text{Welfare}^{\text{taxes}} - \text{Welfare}^{\text{quotas}} = \frac{\sigma^2}{2b^2} (b - G).$$

Recall that the Principle of Certainty Equivalence states that optimal actions, not the values of the program, are independent of the higher moments – here, the variance of θ . The payoff under taxes is greater than the payoff under quotas if and only if $b > G$.

The magnitude of the difference is proportional to σ^2 , but the sign of the difference does not depend on this parameter. This fact is fortunate, because it may be difficult to measure σ^2 . However, the fact that σ^2 is proportional to the variance of BAU emissions, does provide an avenue for calibrating this model. It would be difficult to obtain data on the level of benefits, but in many cases it is practical to estimate a function like equation 3.3. If we can also obtain an estimate of b and G , it is then possible to construct an estimate of σ^2 and thus estimate the magnitude of the gain from using the the welfare-maximizing policy. In any case, it is convenient to have a qualitative result, such as the ranking of taxes and quotas, depend on only two parameters.

If we measure costs and damages in dollars and e in units of tons, then the units of both b and G are

$$\frac{\text{dollars}}{(\text{tons})^2}.$$

Since b and G have the same units, it sensible to compare them. This observation becomes important when we move to the dynamic context, where the units of b and G are different.

Figure 1 shows the *expected* marginal benefits of emissions and the marginal damages of emissions, the solid lines. The horizontal and vertical lines show the optimal tax and the optimal quota. The dashed line shows the marginal benefit of emissions for a realization of the cost shock $\theta > 0$. In this figure $b \gg G$. The small heavily shaded triangle denoted D is the deadweight loss under the tax when θ takes the particular positive value shown in the figure. Under the tax, the firm increases its emissions from the expected level (“quota”) to a level slightly too high (from the standpoint of social welfare), resulting in a small deadweight loss. Under the optimal quota, the level of emissions is fixed at much too low a level (from the standpoint of social optimality) for this particular value of θ , and the deadweight loss is the large triangle E. As the relative magnitudes of b and G change, the relative size of the triangles, and thus the ranking of policies, can change.

3.1.1 A different perspective

We can obtain slightly different intuition about this problem by writing the firm’s payoff in terms of abatement costs rather than benefits. Abatement

is the difference between BAU emissions and actual emissions. Using the fact that $e^{\text{BAU}} = \frac{a+\theta}{b}$, if actual emissions are e , then abatement, A , equals $\frac{a+\theta}{b} - e$. Table 1 summarizes the levels of abatement and emissions under the two policies. The table shows that under taxes emissions are random and abatement is deterministic, whereas under quotas emissions are deterministic and abatement is random. Abatement is deterministic under the tax because there emissions and abatement costs are perfectly positively correlated; therefore, the difference between the two is independent of the cost shock. In contrast, under the quota emissions are independent of the cost shock; therefore the difference between BAU emissions and actual emissions is random.

	tax = $a - zb$	quota = e
emissions	$z + \frac{\theta}{b}$	e
abatement	$\frac{a}{b} - z$	$\frac{a+\theta}{b} - e$

Table 1: emissions and abatement under taxes and quotas

The social cost of abatement is simply the difference between the firm's benefits under BAU, given by equation 3.2, and its (tax-exclusive) benefits under regulation:

$$\frac{1}{2} \frac{2\tilde{f}b + a^2 + 2a\theta + \theta^2}{b} - \left[\tilde{f} + (a + \theta)e - \frac{b}{2}e^2 \right] = \frac{1}{2}bA^2.$$

With taxes, abatement and abatement cost are non-stochastic, but emissions and environmental damages are random. Under quotas, abatement and abatement costs are random, but emissions and environmental damages are non-stochastic. Moreover, by the Principle of Certainty Equivalence, the expected levels of emissions (and abatement) are the same under the two policies.

The policy ranking depends on whether the regulator prefers society to have random abatement costs (under quotas) or random damages (under taxes). Both abatement costs and damages are convex quadratic functions of abatement. By Jensen's Inequality, greater convexity increases the expectation of such a function. If $b > G$ then abatement costs are more convex than damages; in this case, society is willing to accept random damages in order to obtain certain abatement costs, so the regulator prefers taxes. If $b < G$ the ranking is reversed.

3.1.2 Cost variations

Most of the economic analysis focuses on the relative efficiency of the two policies, as above. However, some proponents of the tax point out that taxes have an additional advantage in that they render the price of emissions constant. In contrast, under tradable quotas, the price of emissions, equal to the marginal cost of abatement, is a random variable, equal to $bA = a + \theta - be$. Thus, under quotas the variance of price equals the variance of θ , σ^2 . To the extent that firms – either polluters or those firms investing in abatement technology – dislike price variation, taxes have an advantage apart from efficiency considerations.

Firms are likely more concerned about their average than their marginal abatement costs. The fact that the marginal abatement cost is steeper than the average abatement cost (when costs are convex) means that average abatement costs vary less than marginal abatement costs. In the linear model, the marginal abatement cost curve is twice as steep as the average abatement cost, causing the variance of marginal costs to be four times the magnitude of the variance of the average cost. The focus on the price of emissions permits (equal to marginal, not average costs) can give an exaggerated view of the importance of cost variation.

The ability, under the tax, to arbitrage abatement over states of nature means that expected *tax exclusive* abatement costs are lower under the tax than under the quota; this difference is the economic advantage of taxes, as discussed above. Economists are typically interested in the tax exclusive investment costs, because the emissions tax revenue is a transfer from tax-paying firms to society; such transfers have no effect on efficiency, and therefore are, for good reason, usually ignored in economic analyses.⁴

Firms, however, care about *tax inclusive* abatement costs; for this reason, firms prefer a cap and trade policy, with freely distributed permits, rather than a tax policy. Economic discussions that focus on marginal costs (which vary under the quota and are constant under the tax) provide a misleading comparison of the cost variability. The focus on marginal costs incorrectly suggests that the cost variation is zero under taxes and high under quotas.

⁴However, emissions tax revenue or auction revenue can replace tax revenue raised by distortion-inducing taxes (e.g. taxes that increase labor costs). This “double dividend” creates an additional argument for taxes, and for the auctioning of permits under cap and trade.

Business people understand average costs, but may be vague about marginal cost. For the linear model above, the variance of average (*tax inclusive*) abatement costs under the tax policy is four times as large as the variance of average abatement costs under the freely distributed quota (Problem 1.a). The larger variance is due to the variance of the tax payments, which are non-existent under the freely distributed quota. Of course, this comparison may not be the relevant one: to the extent that quota rights are auctioned, the variability of the emissions price increases the variability of average abatement costs (inclusive of payments for permits).

Although firms care more about average than marginal costs – because they understand the former better – what they really care about is total (tax inclusive) abatement costs. If, in addition to the assumptions of the linear model, we also assume that the cost shock is normally distributed,⁵ then we can obtain a formula for the ratio between the variance of costs under the freely distributed quota, to the variance under the tax (Problem 1.b). Using this formula with plausible numerical values, it appears that the two variances do not differ by much.

In summary, using the linear model that provides the intuition behind most of the “tax versus quotas” literature, we find (not surprisingly) that the tax inclusive expected abatement costs are higher than the expected abatement costs under the quota with freely distributed permits, and that the variances of abatement costs are essentially the same under the two policies. Therefore, unless firms are extraordinarily risk averse, they always prefer the freely distributed quota instead of the tax.

This result is important because much of the political economy analysis favoring taxes does so on the basis that these lead to less variable costs. That claim is correct, if we understand “costs” to mean “economic costs”, i.e. costs exclusive of tax payments. But that is *not* how most business people would understand the term. To them, costs include tax costs. Since the point about cost variability is often made in the context of a discussion of the political economy of the policies, it probably makes more sense to use the term “costs” as business people understand it. With that usage, the

⁵We assumed that the quota is binding in every state of nature. With normally distributed costs, there will be some realizations of θ for which any quota is not binding. Thus, the assumption of normally distributed cost shocks is not consistent with the model, but it nevertheless might be useful as a means of obtaining an approximation to assess the magnitude of cost variation under different policies.

political economy claim that taxes are better than quotas because they lead to greater cost certainty, is overstated.

3.2 The LQ control problem

Here we study a LQ problem in which both the state variable and the control variable are scalars. This specialization makes the calculations transparent, and helps in understanding the nature of the problem. The generalization where the state variable and/or control variable are multi-dimensional uses matrix algebra. Denote x as the state variable and u as the control variable. We begin with the following rather obvious observation:

Remark 1 *If $Q(x, u)$ is a concave quadratic function of x and u then*

$$Q^*(x) = \max_u Q(x, u)$$

is a quadratic function of x and

$$u^*(x) = \arg \max Q(x, u)$$

is a linear function of x .

Thus, we know that for the LQ problem, the value of the optimal program is quadratic in the state variable and the optimal control variable is a linear function of the state variable. We normalize time so that the initial period is $t = 0$, and study the following LQ problem

$$\max E_0 \sum_{\tau=0}^T \beta^\tau \left(\frac{ax_\tau^2 + bu_\tau^2}{2} \right) \quad (3.8)$$

subject to

$$x_{\tau+1} = cx_\tau + nu_\tau + v_\tau \text{ for } \tau \geq 0 \text{ and } x_0 \text{ given.} \quad (3.9)$$

where v is an iid r.v. with mean 0 and variance σ^2 and $a < 0$, $b < 0$, so the problem is concave in x, u . We make no assumptions about the random variable, except that it has a zero mean and finite variance. The expectation operator in expression 3.8 is conditioned on information at time the beginning of the problem, time $t = 0$.

We note the following features of this problem.

1. The random variable enters the equation of motion additively. This additivity is an essential feature of the problem. Section 3.3 discusses an alternative.
2. This problem could be generalized by including a bilinear term, xu , in the objective, a constant and terms that are linear in x and u in the objective, and a constant in the equation of motion, with little increase in complexity. In the interests of simplicity we do not include them, but we later discuss how their inclusion changes the solution. Note that by including a constant in the equation of motion we can relax the assumption that the random variable has mean zero.
3. Time enters the model only via the constant discount rate, β , and the number of periods, at the initial time, remaining in the problem, T . The other parameters of the problem, a, b, c, n, β and σ are time-invariant. These assumptions mean that in the limit as $T \rightarrow \infty$ the problem becomes autonomous. Recall that an autonomous problem is one in which calendar time enters explicitly only via constant discounting.

For an autonomous problem, we obtain a stationary control rule, defined as a function that depends on the state variable but not on calendar time, and which maps the state variable into the control variable. In a steady state, the value of the state variable, and thus the value of the control variable, are constant. Outside of a steady state, the state variable changes over time, and thus the optimal control also changes over time; but the function mapping the state variable into the control variable does not change over time. Our use of time-invariant parameters is an important specialization, but it makes the notation simpler. Our use, initially, of a finite value of T , makes it simple to construct the optimal solution, and also makes apparent the kind of notational changes that would be needed if the parameters depended on calendar time.

It is important to understand the distinction between current value and present value payoffs. Suppose that we have a stream of single-period payoffs, $\{y_t, y_{t+1}, \dots, y_{t+T}\}$, where T is the length of the sequence and t is an arbitrary calendar time. With a constant discount rate β , the current value of this stream of payoffs at the time the sequence begins, time t , is $\sum_{\tau=0}^T \beta^\tau y_{t+\tau}$. The present value at time $t = 0$ of this payoff is $\beta^t \sum_{\tau=0}^T \beta^\tau y_{t+\tau}$. The two

payoffs differ by the factor β^t , which discounts a payoff at time t back to time 0. For our purposes it is simpler to work only with the current value of the stream of payoffs.

The dynamic programming equation (DPE) states that the value of the program (a function of the current state variable and the number of periods remaining in the problem) equals the maximum, over the control variable in the current period, of the expectation of the sum of the payoff in the current period, and the discounted continuation payoff. As in Chapter 1, we use a superscript on a function to denote the number of periods remaining in the problem. Thus, $J^{s+1}(x)$ is the current value of the program that begins with a state variable x when there are $s+1$ periods left to go. This function does not depend explicitly on calendar time, so we have do not include calendar time as an argument. For arbitrary calendar time t , when x denotes the value of x_t , we use x' to denote the value of x_{t+1} .

The current value DPE is⁶

$$\begin{aligned} J^{s+1}(x) &= \max_u E_v \left\{ \left(\frac{ax^2 + bu^2}{2} \right) + \beta [J^s(x')] \right\} \\ &= \max_u \left\{ \left(\frac{ax^2 + bu^2}{2} \right) + \beta E_v [J^s(x')] \right\}. \end{aligned}$$

Note that the index s for time-to-go decreases as we move forward in calendar time. If in the current period we have $s+1$ decisions remaining, then in the subsequent period we will have s decisions remaining. In this problem, the expectation is taken over the random variable v . The assumption that this variable is independently and identically distributed means that it does not require a time index.

We want to show that the value of the program is quadratic in the state, $J^s(x) = \psi_s + \rho_s \frac{x^2}{2}$, and in the process find formulae for the endogenous parameters $\{\psi_s, \rho_s\}$.⁷ We use an inductive proof. The first step of the proof

⁶If we wanted to use present values, we would need to include calendar time. For example, we could denote $V(x_t, t; T+1)$ as the present value at time 0 of the program that begins at calendar time t with initial condition x_t and ends at calendar time $T+1$. With this notation, $\beta^t J^{T+1-t}(x) = V(x_t, t; T+1)$.

⁷In order to be consistent with our convention of using superscripts to denote time-to-go, the reader might prefer to see μ^s and ρ^s rather than μ_s and ρ_s . However, the parameter ρ enters subsequent formulae as a quadratic, and it might be confusing to have both the superscript represent a time-to-go index and an exponent. We therefore use subscripts on these variables.

begins in the final period where $s = 1$ and confirms that the hypothesis is true for that value of s . The second step confirms that if the hypothesis is true when there are s periods to go, then it is also true when there are $s + 1$ periods to go. In the process of confirming the hypothesis, we also obtain the optimal control rule.

In the final period, where $s = 1$, given the current value of the state variable x , the problem is static and deterministic:

$$\max \left(\frac{ax^2 + bu^2}{2} \right).$$

The solution in the final period is $u = 0$ (because $b < 0$) and the value of the program is $\frac{ax^2}{2}$. Thus, we see that the hypothesis is true at $s = 1$ with

$$\psi_1 = 0 \quad (3.10)$$

$$\rho_1 = a. \quad (3.11)$$

Under our hypothesis, the DPE when there are $s + 1$ periods left to go is

$$\psi_{s+1} + \rho_{s+1} \frac{x^2}{2} = \max_u \left(\left(\frac{ax^2 + bu^2}{2} \right) + \beta E_v \left(\psi_s + \rho_s \frac{(x')^2}{2} \right) \right). \quad (3.12)$$

Use the relation $x' = cx + nu + v$ to write the term after the expectations operator as

$$\psi_s + \rho_s \frac{(x')^2}{2} = \psi_s + \rho_s \frac{(cx + nu + v)^2}{2}.$$

Expanding $(cx + nu + v)^2$ we have

$$(cx + nu + v)^2 = c^2x^2 + 2cxnu + 2cxv + n^2u^2 + 2nuv + v^2.$$

Now take expectations with respect to v to write

$$E_v (cx + nu + v)^2 = c^2x^2 + 2cxnu + n^2u^2 + \sigma^2.$$

We use this expression to eliminate the expectation and rewrite the DPE, equation 3.12, as

$$\begin{aligned} & \psi_{s+1} + \rho_{s+1} \frac{x^2}{2} \\ &= \max_u \left(\left(\frac{ax^2 + bu^2}{2} \right) + \beta \left(\psi_s + \rho_s \frac{c^2x^2 + 2cxnu + n^2u^2 + \sigma^2}{2} \right) \right) \\ &= \max_u \left[\left(\frac{1}{2}b + \frac{1}{2}\beta\rho_s n^2 \right) u^2 + \beta\rho_s cnux + \left(\frac{1}{2}a + \frac{1}{2}\beta\rho_s c^2 \right) x^2 + R \right] \end{aligned}$$

which uses the definition

$$R = \beta \left(\psi_s + \frac{1}{2} \rho_s \sigma^2 \right).$$

The function R collects terms that do not involve either the state variable or the control variable. To obtain the last line of the DPE we collected terms in powers of x, u . Note that the right hand side of the DPE is a quadratic in u, x . The first order condition for a maximum is sufficient if and only if the maximand is concave in u . Concavity in u is equivalent to

$$b + \beta \rho_s n^2 < 0. \quad (3.13)$$

For the time being we merely assume that this inequality holds. Later we confirm it. Given concavity of the maximand, by Remark 1 we know that the left hand side (the current value function) is a quadratic in x and the control rule is linear in x .

We perform the maximization to obtain the control rule

$$u^{s+1}(x) = \frac{-\beta \rho_s c n}{b + \beta \rho_s n^2} x. \quad (3.14)$$

The left side of this equation contains the superscript $s + 1$ to indicate that this is the control rule when there are $s + 1$ periods to go. The argument of the control rule is the value of the state variable at the time the decision is made, x . The only non-constant parameter of this control rule, ρ_s , is a coefficient of the value function in the next period. Thus, the value of the control variable in an arbitrary period depends on the value of the state variable in that period, and on a parameter that describes the value function in the next period.

The maximized DPE is

$$\begin{aligned} & \psi_{s+1} + \rho_{s+1} \frac{x^2}{2} \\ &= \frac{1}{2} \frac{ab + a\beta\rho_s n^2 + \beta\rho_s c^2 b}{b + \beta\rho_s n^2} x^2 \\ &+ \frac{1}{2} \frac{2\beta\psi_s b + 2\psi_s \beta^2 \rho_s n^2 + \beta\rho_s \sigma^2 b + \rho_s^2 \sigma^2 \beta^2 n^2}{b + \beta\rho_s n^2} \end{aligned} \quad (3.15)$$

In order for this relation to hold for all values of x , it must be the case that the term multiplying x on the left side equals the term multiplying x on the right side, and that the term independent of x on the left side equals the term independent of x on the right side. We therefore “equate coefficients of x ” in equation 3.15 to obtain the two difference equations

$$\rho_{s+1} = a + \frac{\beta\rho_s c^2 b}{b + \beta\rho_s n^2} \quad (3.16)$$

$$\psi_{s+1} = \frac{1}{2} \frac{2\beta\psi_s b + 2\psi_s \beta^2 \rho_s n^2 + \beta\rho_s \sigma^2 b + \rho_{t+1}^2 \sigma^2 \beta^2 n^2}{b + \beta\rho_s n^2}. \quad (3.17)$$

Putting aside the question of concavity – to be confirmed shortly – we have established that the value function is quadratic in the state variable x , and we have obtained formulae for the coefficients of that function, ψ_s, ρ_s , equations; in the process we also showed that the optimal decision is a linear function of the state variable, and we obtained an expression for the optimal control rule. The formulae for ψ_s, ρ_s take the form of a pair of differential equations, 3.16 and 3.17, together with their boundary conditions, equations 3.10 and 3.11. Equation 3.16 is known as a “Riccati difference equation”.

We emphasize the following points:

Remark 2 *As the index s , time-to-go, increases, we move backwards in calendar time, away from the last period, when $s = 1$. In this sense, we determine the parameters of the control rule and the value function by solving the problem backwards in time. At the beginning of the program we know all of the values of these parameters. When we have $s + 1$ periods to go, we know that value of the state variable in that period, x , and we use the control rule, equation 3.14, to obtain the optimal value of the control in that period. This decision and the subsequent realization of the current random variable determine x' , the value of the state variable in the next period. That is, we obtain the endogenous parameters of the value function and the control rule by solving backwards in time and then we solve for the optimal control and the value of the state variable on the optimal trajectory forwards in time. This backward sweep followed by a forward sweep is typical in dynamic programming. It distinguishes dynamic programming from non-linear programming, because the latter attempts to solve for everything in one fell swoop.*

Remark 3 *The difference equations 3.16 and 3.17 are recursive. The equation for ρ is not a function of ψ , but the equation for ψ is a function of ρ . Therefore, we can solve for ρ first and then use that solution to solve for ψ . The fact that we can solve for these two endogenous variables recursively rather than simultaneously obviously simplifies the solution.*

Remark 4 *The control rule, equation 3.14, is a function of ρ , but not of ψ . In addition, from equation 3.16, ρ is not a function of σ . Therefore, the control rule is not a function of σ . This is an example of the “Principle of Certainty Equivalence”, which states that we obtain the same control rule if we set the random term equal to its expected value (here 0). The optimal control is independent of the variance and all higher moments of the random variable. This Principle of Certainty Equivalence depends on the fact that the objective is quadratic, the equation of motion is linear, there are no binding constraints (e.g. non-negativity constraints), and the random variable enters the problem additively.*

Remark 5 *The endogenous parameter ψ_s is a function of σ . Therefore, the value of the program (the value function) does depend on the variance – even though the control rule is independent of the variance.*

Remark 6 *The control rule is affine, not just linear, in the state variable, and the value function does not contain a term that is linear in the state variable. These features are a result of the fact that in our example the objective does not include terms that are linear in x, u and the equation of motion does not include a constant. Had we included such terms, the optimal decision rule would have included a constant, and the value function would have included a term that is linear in x . That is, we would have $J^s(x) = \psi_s + \gamma_s x + \frac{1}{2}\rho_s x^2$. The recursive structure of the problem remains: we first solve the difference equation for ρ , which would be unchanged by the presence of the new terms; we use that solution to solve the difference equation for γ , which does depend on the new terms; we then use these two solutions to find the solution for ψ .*

Remark 7 *The example above generalizes in a straightforward way to the case of time-dependent parameters. With time-dependent parameters, there is no particular advantage to using time-to-go rather than calendar time as*

the index. However, it would be straightforward to modify the formulae above to include time dependence. For example, suppose that we replace b by a time-varying parameter b_t , where t is calendar time. Using our definition of time-to-go, at calendar time t there are $s = T - t$ periods to go. We can define $\tilde{b}_s = b_t$. For the problem above, b appears on the right side of the formulae 3.14, 3.16 and 3.17. In each of these formulae we replace b by \tilde{b}_s . With time-dependent parameters, the problem is not autonomous even if $T = \infty$.

Our final task here is to confirm that the second order condition, inequality 3.13, is satisfied. We begin by rewriting the Riccati difference equation 3.16 as

$$\rho_{s+1} = \frac{ab + a\beta\rho_s n^2 + \beta\rho_s c^2 b}{b + \beta\rho_s n^2} = a + k(\rho_s) \quad (3.18)$$

with

$$k(\rho_s) \equiv \frac{\beta\rho_s c^2 b}{b + \beta\rho_s n^2}.$$

We have

$$k(0) = 0 \quad \text{and} \quad \frac{dk}{d\rho} = \beta c^2 \frac{b^2}{(b + \beta\rho n^2)^2} > 0. \quad (3.19)$$

Equations 3.19 imply that $\rho_s \leq 0$ is sufficient to insure $\rho_{s+1} < 0$. Our boundary condition states that $\rho_1 = a < 0$. Therefore, by induction, $\rho_s \leq 0$ for all $s \geq 1$. This fact together with $b < 0$ implies that $b + \beta\rho_s n^2 < 0$, so the sufficient condition for optimality is satisfied.

3.2.1 The infinite horizon limit

In the case where the terminal time, T , is finite, so that time-to-go s remains finite, there is not much more to be said about this problem. However, the solution simplifies tremendously if we take the limit as $T \rightarrow \infty$, which is equivalent to letting $s \rightarrow \infty$. In the limit we obtain a stationary control rule: ρ_s and ψ_s converge to constants. Remark 2 notes that we obtain ρ_s and ψ_s by moving away from the final time period, i.e. by solving the problem backwards in time, beginning with the terminal time. When T grows, the horizon, $T - t$, increases for any finite calendar time t . Our choice of writing the solution in terms of time to go rather than calendar time makes it easy to take the limit as the horizon becomes infinite.

A steady state value of the Riccati difference equation for ρ satisfies

$$\rho = a + \frac{\beta\rho c^2 b}{b + \beta\rho n^2}. \quad (3.20)$$

This equation is known as an algebraic Riccati equation. To emphasize that this equation is quadratic in ρ , we rewrite it as

$$0 = \beta\rho^2 n^2 + (b - a\beta n^2 - \beta c^2 b)\rho - ab \equiv h(\rho). \quad (3.21)$$

By inspection, $h(0) = -ab < 0$ and $h \rightarrow \infty$ as $\rho \rightarrow \infty$, so there is a positive and a negative root of the algebraic Riccati equation; denote these as ρ_+ and ρ_- , respectively. We noted above that any negative value of ρ_s is mapped into a negative value of ρ_{s+1} ; because we start with a negative value of ρ_1 , we know that every value of ρ must be negative. The steady state must therefore also be negative. We consequently have only one candidate for a stable steady state, the negative root ρ_- .

We now need to confirm that the negative root, ρ_- , is a stable steady state. Recall from Chapter 1 that a steady state ρ_∞ to equation 3.18 is stable if and only if

$$-1 < \frac{d(a + k(\rho_\infty))}{d\rho} = k'(\rho_\infty) < 1.$$

From inequality 3.19 we know that the first inequality is always satisfied, so we need only confirm that the second is satisfied at the candidate $\rho = \rho_-$. Rather than confirming this algebraically, it is easier to use graphs. Figure 2 graphs the right side of equation 3.20 for $\rho < \frac{-b}{\beta n^2}$. (The function is discontinuous at $\frac{-b}{\beta n^2} > 0$, but we are not interested in positive values of ρ .) This graph intersects the 45 degree line at the two roots, ρ_- and ρ_+ . Using $h(\rho)$, defined in equation 3.21, we see that the graph of the right side of equation 3.20 is everywhere below the 45 degree line between the two roots, and above the 45 degree line on either side of the roots (for $\rho < \frac{-b}{\beta n^2}$).⁸ Therefore the graph must cut the 45 degree line from above at ρ_- . Figure 2 shows these characteristics. Consequently, it must be the case that $k'(\rho_\infty) < 1$, i.e. ρ_- is a stable steady state. Examining the dynamics of ρ , using the construction described in Chapter 1 (for the capital

⁸A straightforward calculation shows that $h\left(\frac{-b}{\beta n^2}\right) = c^2 \frac{b^2}{n^2} > 0$. Therefore, the point of discontinuity of $g(\rho)$, $\frac{-b}{\beta n^2}$, lies to the right of ρ_+ .

accumulation problem), we see that the sequence of ρ_s converges to ρ_- for any initial condition $\rho_1 < \rho_+$. In particular, the sequence converges for $\rho_1 = a < 0$.

There are two other ways to see that the negative root of the algebraic Riccati equation is the correct root. The first alternative uses the fact that a necessary condition for the solution to the dynamic programming equation to be the correct solution to the infinite horizon optimization problem is $\lim_{t \rightarrow \infty} E\beta^t J(x_t) = 0$. Some tedious calculation shows that if we substitute the positive steady state ρ_+ into the stationary version of the control rule, equation 3.14, to obtain

$$u = -\beta\rho_+cn \frac{x}{b + \beta\rho_+n^2},$$

the expectation of the controlled state is explosive. Moreover, it increases at a sufficiently fast rate so that $\lim_{t \rightarrow \infty} E\beta^t J(x_t) = -\infty$. Therefore, the positive root does not satisfy a necessary condition for the solution of the problem. By substituting the negative root into the stationary control rule, the reader can confirm that the resulting controlled state converges and that $\lim_{t \rightarrow \infty} E\beta^t J(x_t) = 0$, as required. Thus, the negative root is the correct root.

An even simpler means of identifying the correct root notes that because $a < 0$ and $b < 0$, the value of the program must be negative for all values of $x \neq 0$. However, with the positive root, ρ_+ , the value of the program is positive for sufficiently large values of x . Therefore, the positive root is not the correct root.

To complete the solution to the infinite horizon problem, we find the steady state for ψ_t using equation 3.17. Solving the algebraic equation obtained by removing the time subscripts in this difference equation, we obtain the steady state

$$\psi = \frac{1}{2}\beta\rho \frac{\sigma^2}{1 - \beta},$$

which we evaluate at $\rho = \rho_-$, the steady state of ρ_t .

Note that in the autonomous control problem, the value function does not depend on calendar time or on time to go. For the autonomous problem, we have $J(x) = \psi + \frac{\rho}{2}x^2$, where ψ and ρ are the steady states obtained above.

In Chapter 1 we showed how to approximate the control rule for a general problem (not LQ) in the neighborhood of a steady state. The material in this chapter provides an alternative procedure. We can find the steady state to the general problem and then take a second order approximation of the single period payoff, evaluated at the steady state, and a first order approximation of the equation of motion, evaluated at the steady state. We then have a linear-quadratic approximation of the original general control problem. We can apply the methods of this chapter to obtain the solution, i.e. the control rule and the value function, to this linear quadratic problem. That solution is an approximation to the solution of the original general problem, where the approximation is evaluated at a steady state.

3.3 Two variations

This section consider two variations of the LQ control problem studied above. The first variation allows the noise to be multiplicative and the second introduces risk aversion. Both of these variations admit a closed form solution, and in both cases the optimal control variable is a linear function of the state variable. However, the Principle of Certainty Equivalence does not hold: the optimal control rule is a function of the variance of the stochastic term.

3.3.1 Multiplicative noise

Suppose that we replace equation 3.9 with

$$x_{\tau+1} = cx_{\tau} + n(1 + v_{\tau})u_{\tau}$$

or

$$x_{\tau+1} = c(1 + v_{\tau})x_{\tau} + nu_{\tau}.$$

In both of these variations the noise enters multiplicatively rather than additively.

This form of the problem is appropriate for describing the situation where parameters of the state equation are random variables with known mean. This form of the control problem also arises in a variation of the taxes versus quotas problem with asymmetric information, in which the cost shock affects the slope rather than the intercept of marginal abatement costs.

It is still the case that the value function is quadratic and the control variable linear in the state. However, the Principle of Certainty Equivalence no longer holds. In particular, the optimal control rule, in addition to the value of the program, depends on the variance of the random term.

3.3.2 Risk sensitivity

It is possible to introduce constant absolute risk aversion with respect to the entire stream of payoffs into the model with additive uncertainty. In this case, however, it is necessary to assume that the noise is Gaussian. Previously we assumed only that the random variable has mean zero and a finite variance. The maximand in expression 3.8 is replaced by

$$\max E_t \phi \exp \left[\sum_{\tau=0}^T \frac{\phi}{2} \left(\frac{ax_\tau^2 + bu_\tau^2}{2} \right) \right]$$

and the equation of motion remains equation 3.9. The constant ϕ determines the level of risk sensitivity; $\phi < 0$ implies that the regulator is risk averse and $\phi > 0$ implies risk preference. Note that this problem does not contain a positive discount rate (i.e. $\beta = 1$).⁹ Even with constant discounting, the solution to the problem with a positive discount rate leads to a time-inconsistent control rule. We discuss time-inconsistency in Chapter xx.

This problem is called the “Linear quadratic exponential Gaussian” control problem. The Principle of Certainty Equivalence does not hold. The optimal control rule in this case depends both on the variance of noise and on the level of risk aversion/preference ϕ . Variations of this problem include the situation where the summand is linear rather than quadratic in x .

The problem has an interesting relation to a dynamic game. The solution to the optimization problem is identical to the solution to the zero sum dynamic game in which the DM chooses the control rule for u to maximize

$$\sum_{\tau=0}^T \left(\frac{ax_\tau^2 + bu_\tau^2 + \phi\sigma^2v_\tau^2}{2} \right)$$

⁹Remember the difference between a discount rate and a discount factor. If the interest rate, or discount rate, for a single period is r then the discount factor for a period is $\beta = \frac{1}{1+r}$. A zero discount rate corresponds to a discount factor equal to 1.

and nature chooses the control rule for v to minimize this expression.

Whittle [??] provides an early treatment of this optimization problem and Jacobsen [??] provides further analysis. Applications of this control problem in economics include several papers by Karp [1987, 1988, 1992] (“Methods for Selecting the Optimal Dynamic Hedge When Production is Stochastic.” *American Journal of Agricultural Economics*, Vol. 69, No. 3 (1987) pp. 647-657; “Dynamic Hedging with Uncertain Production.” *International Economic Review*, Vol. 29, No. 4 (1988) pp. 621-637; “The Endogenous Stability of Economic Systems: The Case of Many Agents.” *Journal of Economic Dynamics and Control*, Vol. 16 (1992) pp. 117-138.) and by Hansen and ?? [??]

3.4 Taxes versus quotas with stock pollutants

Here we study the question of ranking taxes and quotas when firms create a stock pollutant. We begin by writing down the single period payoff, which equals the difference between the benefit of emissions and the damage resulting from the pollution stock; we then write down the equation of motion for the stock. The objective of the problem is to maximize the expectation of the present discounted value of the stream of net benefits (the benefits from emissions minus the damages from the stock); the equation of motion is the constraint.

This model differs from the LQ problem from Chapter 3.2 in one respect. Here the random shock enters the payoff directly, and also enters the equation of motion. In the LQ problem that we studied above, the random term enters only the equation of motion. This difference leads to a minor modification in the problem.

The benefit of emissions The dynamic model is a straightforward variation of the static model, but in the dynamic setting damages depend on a stock rather than a flow. As above, abatement equals the difference between the actual level of emissions and the Business as Usual (BAU) level. The abatement costs are quadratic in abatement, so the benefit of emissions is a quadratic function of emissions. Again, the intercept of the marginal benefit function equals a constant a plus a mean-zero random variable θ_t with a constant and known variance σ^2 . In the dynamic setting, time subscripts

denote the fact that some variables, including the cost shock and the level of emissions, change over time. The slope of marginal benefits is a known constant b . In period t the firm, but not the regulator, knows the value of θ_t . The benefit of emissions in period t is

$$\tilde{f} + (a + \theta_t) e_t - \frac{b}{2} e_t^2. \quad (3.22)$$

When the regulator sets a tax p_t per unit of emissions, the firm maximizes the benefit of emissions minus the cost of tax. Its problem is

$$\max_e \tilde{f} + (a + \theta_t) e_t - \frac{b}{2} e_t^2 - p_t e_t.$$

The first order condition to this problem implies that the level of emissions is

$$e_t^* = \frac{a - p_t}{b} + \frac{\theta_t}{b} \equiv z_t + \frac{\theta_t}{b}. \quad (3.23)$$

As above, we model the tax-setting regulator as choosing z_t , the expected level of emissions under a tax. Substituting e_t^* into the firm's benefit function (3.22) and taking expectations, gives the expected benefit of emissions under the tax policy z_t :

$$\tilde{f} + a z_t + \frac{\sigma^2}{2b} - \frac{b}{2} z_t^2. \quad (3.24)$$

The quota-setting regulator chooses e_t , which by assumption is binding with probability 1. Thus, the expected benefit of emissions under the quota policy e_t is simply

$$\tilde{f} + a e_t - \frac{b}{2} e_t^2. \quad (3.25)$$

Exactly as in the static model, the tax-setting regulator determines only the expected level of emissions, whereas the quota-setting regulator chooses actual emissions.

Environmental damages Let S_t be the stock of pollutants, and e_t be the flow of emissions in period t . All time dependent variables are constant within a period. The fraction $0 \leq \Delta \leq 1$ of the pollutant stock lasts into the next period, so the growth equation for S_t is:

$$S_{t+1} = \Delta S_t + e_t. \quad (3.26)$$

With taxes, the flow of emissions and thus the next period pollutant stock, S_{t+1} , is stochastic since it depends on the cost shock. With quotas, the regulator is able to exactly determine the change in pollution stock.

The environmental damage in period t is

$$D(S_t) = \frac{G}{2} (S_t - \bar{S})^2. \quad (3.27)$$

The single period payoff We continue to treat the regulator's decision variable under taxes as z , the expected level of emissions given a tax of p . In an attempt to reduce notational complexity, and to emphasize the similarity of the problems under taxes and quotas, we hereafter replace the quota e with z . The reader must keep in mind the context in which we use this variable. Under taxes, if the regulator chooses z , actual emissions are $z + \frac{\theta_t}{b}$, a random variable. Under quotas, if the regulator chooses z , actual emissions are z , which is not random.

The expected payoff in a period equals the expected benefits of emissions minus the damages. The expectation is taken with respect to the cost shock, θ . Under *taxes*, where emissions are given by equation (3.23), the expected single period payoff is

$$f + az_t - \frac{bz_t^2}{2} + \frac{\sigma^2}{2b} - cS_t - \frac{G}{2}S_t^2 \quad (3.28)$$

with $f \equiv \tilde{f} - \frac{G}{2}\bar{S}^2$, $c \equiv -G\bar{S}$.

In view of the notational change described above, we write the expected single period payoff under quotas as

$$f + az_t - \frac{bz_t^2}{2} - cS_t - \frac{G}{2}S_t^2. \quad (3.29)$$

(Prior to the notational change, the expression 3.29 would have the argument e rather than z .) The presence of the term $\frac{\sigma^2}{2b}$ in the expected single period payoff under taxes and its absence under quotas reflects the fact that with taxes the firm can adjust to the cost shock – a possibility not available under quotas. Thus, for the same expected level of emissions, the expected payoff is higher under taxes.

Comparison between the static and the dynamic problems We noted that in the static problem, the units of b and G are the same, so a direct comparison of the two parameters is sensible. The units of these parameters differ in the dynamic setting, so it is no surprise that the ranking of policies cannot depend merely on a comparison of their magnitudes.

Some economists claim that it is “obvious” from conventional parameter estimates that taxes dominate quotas for the control of greenhouse gasses, because most estimates suggest that the marginal damage function is quite flat, i.e. G is small relative to b . This observation contains useful intuition, but the fact that the two parameters do not even have the same units means that the intuition cannot possibly be exactly correct. We have to solve the dynamic problem in order to know how to use the slopes to compare the two policies.

We continue to measure costs and damages in dollars, as in the static setting, and here we measure the pollution stock in tons. Emissions in the dynamic problem are *flow*. If we choose a unit of time to be a year, then the units of e are $\frac{\text{tons}}{\text{year}}$. Because the units of GS^2 and be^2 are both dollars, the units of G and b are

$$\text{units of } G: \frac{\text{dollars}}{(\text{tons})^2}$$

and

$$\text{units of } b: \frac{\text{dollars}}{\left(\frac{\text{tons}}{\text{year}}\right)^2} = \frac{\text{dollars times year}^2}{(\text{tons})^2}.$$

In the dynamic problem, the units of b and G are not the same. Hoel and Karp (2002) show that the comparison of taxes and quotas depend on the length of a period, a parameter they denote as h . We ignore that complication here, setting $h = 1$ and suppressing the parameter.

The dynamic optimization problems The dynamic problems under taxes and under quotas have only two differences. First, the expected single period payoff under taxes has the positive term $\frac{\sigma^2}{2b}$. Given an infinite horizon and a constant discount factor, this term contributes $\frac{\sigma^2}{2b(1-\beta)}$ to the present discounted value of the expected payoff under taxes, an amount absent from the problem with quotas. This term is the source of the advantage of taxes. However, under quotas the equation of motion is $S_{t+1} = \Delta S_t + z_t$, in view of

equation 3.26. Under taxes, the equation of motion is $S_{t+1} = \Delta S_t + z_t + \frac{\theta_t}{b}$. The regulator chooses the actual future stock under quotas, but chooses only the expectation of future stock under taxes. This difference is the source of the advantage of quotas. Ranking of the two policies depends on the relative magnitudes of these two effects.

With a discount factor β , the *tax-setting* regulator's maximized expected payoff at time $t = 0$ is

$$J^{\text{tax}}(S_0) = \max E_0 \sum_{\tau=0}^{\infty} \beta^{\tau} \left\{ f + az_{\tau} - \frac{bz_{\tau}^2}{2} + \frac{\theta_{\tau}^2}{2b} - cS_{\tau} - \frac{G}{2} S_{\tau}^2 \right\}. \quad (3.30)$$

The dynamic programming equation is:

$$\begin{aligned} J^{\text{tax}}(S_t) &= \max_z E_{\theta_t} \left\{ f + az - \frac{bz^2}{2} + \frac{\theta_t^2}{2b} - cS_t - \frac{G}{2} S_t^2 + \beta J^{\text{tax}}(S_{t+1}) \right\} \\ &= \max_z \left\{ f + az - \frac{bz^2}{2} + \frac{\sigma^2}{2b} - cS_t - \frac{G}{2} S_t^2 + E_{\theta_t} \beta J^{\text{tax}}(S_{t+1}) \right\} \\ &\quad \text{subject to} \quad S_{t+1} = \Delta S_t + z + \frac{\theta_t}{b}. \end{aligned} \quad (3.31)$$

In this problem, the random shock enters both the single period payoff and the equation of motion. We used $E_{\theta_t} \frac{\theta_t^2}{2b} = \frac{\sigma^2}{2b}$ to move the expectations operator past the current period payoff. The “tax” superscript on the function J denotes the value function under taxes.

The single period expected payoff under taxes contains the term $\frac{\sigma^2}{2b}$, absent from the expected payoff under quotas. This constant term affects the value of the program, but it does not affect the optimal control rule. The other difference between the problems under taxes and quotas is that the equation of motion contains the random shock $\frac{\theta_t}{b}$ under taxes, whereas the equation of motion is deterministic under quotas. By the Principle of Certainty Equivalence, we know that the control rule under taxes is independent of σ^2 (and all higher moments). Therefore, the control rule is the same under taxes and under quotas. Recall that the random term in the equation of motion under taxes does affect the value of the program under taxes.

Therefore, there are two differences between the problems under taxes and under quotas: under taxes the variance enters the single period payoff directly, and it also affects the value of the program via the randomness of the equation of motion. The problem under taxes is more general than the problem under quotas, in the sense that we obtain the latter from the former by

setting $\sigma^2 = 0$. This observation means that we need only solve the problem under taxes. In the resulting value function, we set $\sigma^2 = 0$ to obtain the value function under quotas. We already know that the control rules are the same under the two problems, because of the Principle of Certainty Equivalence. For completeness, we show the DPE under quotas, even though we do not need to solve this problem; we obtain the solution merely by setting $\sigma^2 = 0$ in the solution to the problem under taxes. The DPE under quotas is

$$J^{\text{quota}}(S_t) = \max_z \left\{ f + az - \frac{bz^2}{2} - cS_t - \frac{G}{2}S_t^2 + \beta J^{\text{quota}}(S_{t+1}) \right\} \quad (3.32)$$

subject to $S_{t+1} = \Delta S_t + z.$

The superscript “quota” denotes the value function under quotas.

The fact that the control rules are the same under the two policies means that *given the same level of S* , the expected level of emissions under taxes equals the level of the (binding) quota. However, the fact that S is stochastic under taxes and deterministic under quotas means that the trajectory for S under the two types of policies will in general differ; therefore the realized trajectory (as distinct from the expected trajectory) of the quota and the expected emissions under taxes also differ. At the initial period, however, the level of S is given, and it is the same under taxes and under quotas. Therefore, in expectation the trajectories under taxes and quotas are the same.

It is worth pausing to repeat the two differences between taxes and quotas. Taxes tend to increase the expected value of the program because in each period the firm can take advantage of its private information. Under taxes, emissions and marginal abatement costs are positively correlated. This flexibility increases expected cost saving, favoring taxes. In expectation, this ability is worth $\frac{\sigma^2}{2b}$ in each period. However, taxes result in a stochastic stock of pollution. In view of Jensen’s Inequality, stochastic stocks increase expected damages because damages are convex in stocks. This stochasticity favors quotas.

We can now apply the results of Section 3.2 to obtain the solution under taxes, and then by setting $\sigma^2 = 0$ we obtain the solution under quotas. The only modification we need is that here the single period payoff contains both a constant, $f + \frac{\sigma^2}{2b}$, and a term that is linear in z and S , $az - cS$. Remark

6 discussed this complication; it means that we have an extra term, linear in S , γS , in the value function, and we have an intercept in the optimal control rule (i.e. the control rule is linear rather than affine in the state). The value function under taxes has the form

$$J^{\text{tax}}(S) = \psi + \gamma S + \rho \frac{S^2}{2}.$$

We calculate the values of ψ , γ , and ρ following the same recipe as in Section 3.2. Problem 2 asks the reader to perform this calculation.

The algebraic equations that determine ψ , γ , and ρ are recursive. The equation for ρ is an algebraic Riccati equation; it does not involve ψ or γ . The equation for γ involves ρ but not ψ , and the equation for ψ involves both ρ and γ . In addition, the equation for ψ depends on σ^2 but both γ and ρ are independent of σ^2 . (Problem 2 asks the student to verify these claims.) Consequently, the values of γ and ρ are the same under taxes and quotas. We already noted that the decision rules are the same under the two policies.

Therefore, for given S , the difference in expected payoffs under taxes and quotas depends only on the difference in the values of ψ . The sign of the difference depends on the relative magnitudes of b and G , and also on the parameters Δ and β (and on the length of the time period, which we set equal to 1). The fact that ψ but not γ or ρ depend on σ^2 means that the policy ranking does not depend on the level of the stock, S .

Variations and results The dynamic model above considers the “feedback” solution, where the regulator understands that all future policies are conditioned on information available at the current time. In the “open-loop” solution, the regulator acts as if she is able to choose, today, values of all future policies. In this setting, where there are no strategic considerations, there is no advantage to this kind of commitment, so the feedback equilibrium provides a (weakly) higher expected payoff. Under quotas, where the stock is deterministic, the open-loop and feedback solutions are identical. However, under taxes, where the stock evolves stochastically, the two solutions differ. Under taxes, the payoff under the feedback solution is strictly higher than under the open-loop solution. The regulator obviously does better under taxes if she can condition the time t tax on the time t value of the pollution stock, rather than conditioning this tax on the expectation at time 0 of time

t pollution stock. The open-loop equilibrium is implausible in a dynamic setting (unless stochastics are absent, as under quotas), but the open-loop model is useful because in some cases it leads to simpler results.

Hoel and Karp (2002) show that in the open-loop equilibrium taxes lead to higher expected welfare than quotas if and only if

$$\frac{G}{b} < \frac{1 - \beta\Delta^2}{\beta h^2}$$

where h is the length of a period (suppressed in the discussion above). In the feedback equilibrium, taxes dominate quotas if and only if

$$\frac{G}{b} < \frac{1 - \beta\Delta^2}{\beta h^2} \left(\frac{2 - \beta\Delta^2}{2(1 - \beta\Delta^2)} \right).$$

The right sides of both of these inequalities are decreasing in the discount factor β and the persistence parameter Δ . Thus, a greater concern for future damages (higher β) or a more persistent stock (higher Δ) decrease the critical ratio, and in that respect make it “more likely” that quotas are preferred to tax. (By saying that a change makes an event “more likely” we mean that the change increases the set of parameter space over which this event occurs.) We noted above that the advantage of quotas is that they reduce expected damages; recall Jensen’s Inequality, the fact that damages are convex in the stock, and the fact that the stock is stochastic with taxes and deterministic with quotas. An increase in β or Δ both increase the importance of future damages, thus favoring quotas.

The fact that $\frac{2 - \beta\Delta^2}{2(1 - \beta\Delta^2)} > 1$ means that taxes are more likely to be preferred to quotas under the feedback relative to the open-loop setting. This result is obvious because we saw that the payoff under quotas is the same under the feedback and the open-loop policies, but the payoff under the tax is strictly higher under the feedback policy. The formulae above also show that a shorter length of time between decisions (smaller h) favors taxes.

Newell and Pizer, JEEM 2003, allow the cost shock to follow an AR(1) process, i.e.

$$\theta_t = \alpha\theta_{t-1} + v_t, \nu_t \sim iid.$$

They obtain the criteria for policy ranking when the regulator uses an open-loop policy. The open loop assumption is particularly unrealistic when cost shocks are serially correlated, because the regulator’s knowledge about the

previous level of θ , gained by observing firms' response to the previous tax, provides useful information about the current cost shock. The open-loop assumption means that the regulator discards that information. However, the open-loop assumption leads to simple comparative statics, and in particular shows that positive correlation promotes the use of quotas.

Karp and Zhang (2005) consider serially correlated cost shocks under both open-loop and feedback policies. They show that the direction of the comparative statics is the same under the feedback policy: positive correlation of shocks favors quotas. However, numerical simulations show that the magnitude of the effect is much smaller in the feedback, compared to the open-loop setting. The change in the stock over several periods approximately equals the sum of flows during that time. (A positive decay rate means that the stock change does not exactly equal the sum of flows.) Recall that under taxes, emissions and cost shocks are negatively correlated. Thus, if cost shocks are negatively autocorrelated ($\alpha < 0$), then emissions are also negatively autocorrelated. Other things equal, the variation in the stock is smaller when the flows are negatively autocorrelated – as occurs under taxes when costs are negatively correlated. Thus, negative autocorrelation of costs reduces the characteristic (stochasticity of stocks) that tends to make taxes unattractive. Similarly, positive autocorrelation of costs increases the characteristic that tends to make taxes unattractive. This relation – “the stock correlation effect” – explains why the preference for quotas is monotonically increasing in ρ under an open-loop policy. In most economic settings it appears more likely that costs shocks are positively correlated, so this feature tends to promote quotas.

The stock correlation effect exists but is less important under the feedback tax policy, because the regulator is able to adjust the policy in every period to accommodate the previous shock. In choosing the current tax she need only consider next-period stock variability. All of these papers find that taxes are likely to dominate quotas for plausible parameter estimates.

Hoel and Karp, *J Pub Econ* 2001 consider the case where there is asymmetric information about the slope rather than the intercept of the firm's marginal cost. In this case, the results of Section 3.3.1 apply. The Principle of Certainty Equivalence does not hold. The control rules under taxes and quotas are different, and the expected level of emissions under taxes does not equal the level of emissions under quotas, for given stock S . Therefore, the expected trajectories of the stock are not equal under the two policies.

In addition, the ranking of taxes and quotas in general depends on the level of S . The comparison of policies is therefore much more complicated when the slope rather than the intercept of marginal costs is random. It also appears easier to find plausible parameter values for which quotas dominate taxes.

3.5 Dynamic investment decisions

[To be completed] The models above assume that firms' only decision is emissions. Firms solve a sequence of static optimization problems under taxes, and they merely carry out orders under quotas. This section discusses a model in which firms make dynamic investment decisions that affect future abatement costs. Firms are non-strategic but they have rational expectations, so their current equilibrium investment depends on their beliefs about the level and type of future environmental policy. In this case, the comparison of taxes and quotas is much more complicated.

3.6 Further discussion

We close by discussing several other views of the relative efficiency of taxes and quotas. One view is that the risk of extreme environmental damages, associated with high GHG stocks, means that over some range damages are likely to be very convex in stocks, i.e. the slope of marginal damages is actually very large. In addition, over a long enough time span, given the opportunities for the development and adoption of new technologies, the marginal abatement cost curve is actually rather flat. Based on these observations, and reasoning from the standard static model, Dietz and Stern (2007) conclude that quantity restrictions are more efficient than taxes for climate policy.

There are three reasons for doubting this conclusion. First, the use of the static framework (or the open loop assumption in a dynamic setting) is not appropriate for studying climate policy, because the current policymaker cannot choose policy levels decades into the future. More rapid adjustment of policy, i.e. a decrease in the length of period between policy adjustments, favors the use of taxes. Second, even if the possibility of extreme events makes the marginal damage function much steeper than current estimates suggest,

the magnitude of the slope of damages would have to be implausibly large to favor quotas. Third, the model in Section 3.5 suggests that endogenous investment in abatement capital is likely to increase the advantage of taxes, given the linear quadratic framework.

A second view is that the existing models inaccurately describe the abatement problem and are therefore simply inappropriate for comparing policies. The objection is that firms will first undertake the cheapest abatement opportunities, which will not be available in the future. There are (at least) two ways to respond to this objection. First, a stationary upward sloping marginal abatement cost curve (used in most previous analyses) is obviously consistent with the claim that firms first use the cheapest way of reducing emissions, and then use more expensive means when regulation becomes stricter. However, because abatement is a flow decision, the fact that the cheap abatement opportunities were used early in the program does not mean that they are unavailable later in the program. The firms move up their marginal abatement curves as the policy becomes stricter. A second response interprets the objection as a call to use a model in which abatement is a stock rather than a flow decision – specifically, a model with endogenous investment in abatement capital, in which there is a sequence of increasingly expensive technologies that reduce emissions. It would be fairly straightforward to produce that kind of model, using a slight modification of the model in Section 3.5. That model assumes that the cost of investment is a function of gross investment. To address the objection, we could create a model of investment cost for which the cost of an additional unit of capital increases with the current level of capital. With this formulation, the firms’s level of capital is a proxy for it’s stage of technology. Because it first adopts the cheapest (most efficient) technologies, it becomes increasingly expensive to make further reductions in abatement costs. It is not clear how this change affects the policy ranking, but it would require a fairly simple variation to the model in Section 3.5.

There are several other model variations that would address other interesting questions. For example, network externalities may cause the productivity of a firm’s capital to increase with the level of aggregate capital. There may be intra-firm increasing returns to scale. There might also be learning by doing, so that an increase in cumulative abatement decreases abatement costs. The inclusion of intertemporal trade (banking and borrowing) under quantity restrictions is also potentially interesting. Because GHGs are a

stock pollutant, the stream of damages can be sensitive to the cumulative emissions over a long period of time without being sensitive to the precise timing of emissions. Intertemporal trading allows firms to optimally allocate over time a given cumulative level of emissions. The introduction of banking and borrowing (under the quantity restriction) would likely significantly erode the advantage of taxes. The effect of banking and borrowing on the incentive to invest is not clear.

3.7 Problem set

The purpose of problem 1 is to give the student practice in working with the static model. The purpose of the next two problems is to give students practice in solving two versions of the linear-quadratic control problem, the first with additive and the second with multiplicative disturbances. This exercise develops mechanical skills and helps to reinforce understanding of the Principle of Certainty Equivalence.

Students are encouraged to tackle these problems using a program for symbolic calculations, such as MuPad (with ScientificWorkplace) or Maple or Mathematica. The time needed to learn how to use one of these programs well enough to do this problem set is probably less than the time needed to do the problem set "by hand".

1. (a) For the static model in Section 3.1, compare the magnitudes of the variance of *average* tax inclusive abatement costs (under the tax) and under the quota. (b) Now consider the case where $\theta \sim N(0, \sigma^2)$.
2. (Hoel and Karp 2001, Taxes versus Quotas for a Stock Pollutant) Equation 16 in Appendix A is the Dynamic Programming equation for the stock pollution problem with additive errors. There is a typo above equation 16. The expression for λ should be

$$\lambda = f + az - \frac{bz^2}{2} + \frac{\sigma^2}{2b} - cS - \frac{gS^2}{2}.$$

Using the algorithm in described in Section 3.2, solve the dynamic programming equation for this problem. The solution requires that you find the control rule and the unknown parameters ρ_0, ρ_1, ρ_2 . The paper contains the expressions for ρ_0 and ρ_2 so you need only confirm these equations and find the expression for ρ_1 , and figure out the relation between these coefficients and the optimal control rule. (The optimal control rule is a linear function of the state. The coefficients of this linear function depend on ρ_i .)

3. (Hoel and Karp 2000, Taxes and Quotas for a Stock Pollutant with Multiplicative Uncertainty) Equation 20 in section 5.1 is the Dynamic Programming Equation for a linear-quadratic control problem with multiplicative errors. Confirm that the control rule equals the expression

in equation (21) (equivalently, equation (6)), and that the unknown parameters ρ_0, ρ_1, ρ_2 solve equations (7) - (9). (This paper was published in J Pub Econ in 2001, but the published version contains typos in a couple of equations, so you should use the link to the paper beneath this problem set link.)

Note that the stochastic term plays a fundamentally different role in the problem with multiplicative rather than additive shocks. In particular, the control rule depends on the variance, and the effect of the variance (on the value of the program) depends on the state. In this problem, the Principle of Certainty Equivalence does not hold.

4 Reactions to Risk in a Simple Resource Exploitation Model

This chapter defines the meaning of an increase or decrease in risk, and it shows how to analyze the effect of a change in risk on optimal decisions. We also define and explain the role of “prudence”.

4.1 A Simple Model

We discuss risk in the context of a three-period model. One interpretation of this model is that a social planner maximizes the welfare obtained from consuming an exhaustible resource. The per period welfare obtained from consuming x_t units of the resource in period $t \in 1, 2, 3$ is $u(x_t)$, with $u' > 0, u'' < 0$. In the first period, i.e. at $t = 1$, the planner does not know the total stock of the resource. We model the stock as a random variable $\tilde{\theta}$ with non-degenerate support, e.g. continuous support $[\underline{\theta}, \bar{\theta}]$ with $\underline{\theta} < \bar{\theta}$, or discrete support $\{\theta_1, \dots, \theta_N\}$ with $N > 1$. It is optimal to consume all of the remaining resource stock in the last period. Therefore, we can write overall welfare (a random variable) as

$$v(x_1, x_2, \tilde{\theta}) = u(x_1) + u(x_2) + u(\tilde{\theta} - x_1 - x_2). \quad (4.1)$$

In order to focus on uncertainty and learning we ignore discounting.

We can also interpret equation (4.1) as a two-period stock pollutant model, e.g. for greenhouse gas (GHG) emissions. With this interpretation, $u(x_1)$ denotes welfare derived from emissions in the current period – that is from the underlying consumption to which the emissions are tied – and $u(x_2)$ denotes welfare derived from emissions in the future. The adverse effect of the GHGs is negligible (or neglected) in the current period, but emissions create the GHG stock $x_1 + x_2$ in the future where they cause the damage $-u(\tilde{\theta} - x_1 - x_2)$.

With the stock pollutant interpretation, we can write the objective as

$$v(x_1, x_2, \tilde{\theta}) = u(x_1) + u(x_2) - D(\tilde{\theta}, x_1, x_2)$$

with the damage function $D(\tilde{\theta}, x_1, x_2) = -u(\tilde{\theta} - x_1 - x_2)$, satisfying $D', D'' > 0$. We can think of this model as representing a two-phase simplification of

an infinite time horizon model with discounting. Let the current period be $[0, T]$ and ‘the future’ be $[T, \infty]$. The decision maker plans to consume x_1 in the first phase, i.e. for all $t \in [0, T]$, and x_2 from there on, i.e. for all $t \in [T, \infty]$. The benefits from emissions are $\int_0^\infty u(x_t)e^{-\rho t} dt = \int_0^T u(x_1)e^{-\rho t} dt + \int_T^\infty e^{-\rho t} u(x_2) dt = u(x_1) \int_0^T e^{-\rho t} dt + u(x_2) \int_T^\infty e^{-\rho t} dt$. A simple calculation ($\int_0^T e^{-\rho t} dt \stackrel{!}{=} \int_T^\infty e^{-\rho t} dt$) shows that $u(x_1)$ and $u(x_2)$ obtain equal weights (as in equation 4.1) if $T = \frac{\ln 2}{\rho}$. For a pure rate of time preference $\rho = .03$ we find $T = 23$. With this interpretation, the planner emits level x_1 for the first 23 years without seeing much damage from GHG emissions. Around year 23 the damage sets in and she changes emissions to level x_2 .

For the resource extraction interpretation of the model we have to assume that $x_1 + x_2 < \theta$ for all possible realizations of $\tilde{\theta}$: we cannot extract a resource that has not yet been discovered and might not exist. We assume that the utility function u and the distribution of $\tilde{\theta}$ are such that the optimal choices of x_1 and x_2 are interior solutions so that we can ignore this constraint. Given particular functional forms we can check that this assumption holds. For the GHG emission interpretation of the model such a restriction is not needed because $\tilde{\theta}$ only denotes the distribution of some damage parameter.

Our goal is to determine how the fact that the future resource stock (or damage) is uncertain changes the decision maker’s welfare and consumption plan. With uncertainty she maximizes the expression

$$\begin{aligned} & \max_{x_1} \max_{x_2} \mathbb{E} v(x_1, x_2, \tilde{\theta}) \\ &= \max_{x_1} \max_{x_2} u(x_1) + u(x_2) + \mathbb{E} u(\tilde{\theta} - x_1 - x_2). \end{aligned}$$

Without uncertainty, assuming that $\theta = E\tilde{\theta}$, she maximizes

$$\begin{aligned} & \max_{x_1} \max_{x_2} v(x_1, x_2, \mathbb{E}\tilde{\theta}) \\ &= \max_{x_1} \max_{x_2} u(x_1) + u(x_2) + u(\mathbb{E}\tilde{\theta} - x_1 - x_2) \end{aligned}$$

instead. We first consider the effect of uncertainty on welfare. To this end, we recall Jensen’s inequality, which is the basis for many results involving risk aversion and decision-making under uncertainty. This inequality states that for any concave function ϕ

$$\mathbb{E} \phi(\tilde{\theta}) \leq \phi(\mathbb{E}\tilde{\theta}).$$

The concavity of u implies that the introduction of risk reduces welfare: the value of $E v(x_1, x_2, \tilde{\theta})$ is less than the value of $v(x_1, x_2, E\tilde{\theta})$, for all values of x_1 and x_2 . Of course, the optimal levels of x_1 and x_2 , generally differ across the two scenarios. Let x_1^u and x_2^u denote the optimal consumption levels for the uncertain scenario and x_1^c and x_2^c denote the optimal consumption levels in the certain scenario. Then we have

$$E v(x_1^u, x_2^u, \tilde{\theta}) \leq v(x_1^u, x_2^u, E\tilde{\theta}) \leq v(x_1^c, x_2^c, E\tilde{\theta}),$$

where the first inequality follows from Jensen's Inequality and the second from the fact that x_1^u and x_2^u are feasible in the certain scenario but may not be optimal. Therefore, if the decision maker chooses x_1^c and x_2^c instead they must yield at least as much welfare. Thus, *the introduction of uncertainty reduces overall (ex-ante) welfare.*

4.2 Defining a reduction or increase in risk

This section explains what is meant by 'reducing or increasing risk'. The first formal model we discussed in section ?? illustrates the willingness to pay in order to reduce the risk of a health condition. Section 4.1 introduces risk over a resource stock or a damage parameter and compares the setting to one without uncertainty. Both models deal with a risk reduction or elimination. However, the reduction of risk has different meanings in the two models.

In a model with a binary outcome, x and $x - L$, where L is the loss following a bad outcome, a change of the probability p of the event has two implications. First – assuming that $p < \frac{1}{2}$ – reducing the probability p of the event decreases the variance of the distribution of outcomes: $\sigma^2 = L^2 p(1 - p)$ is single peaked and has its maximum in at $p = \frac{1}{2}$. Second, reducing p also increases the expected value of the outcome, $\mu = x - pL$. Hence, in this model a reduction in p both reduces the variance and increases the expectation of the outcome.

In the resource extraction or GHG emission model we compared the risky setting with a setting under certainty by exchanging the random stock (respectively damage parameter) with its expected value. There we kept the expected value of the random variable constant. The variance of a random variable may appear to be a good measure of risk, because setting the variance to 0 eliminates risk without changing the expected value. Although the variance is an adequate proxy if we are interested only in the effect of

eliminating risk, it is an inadequate measure for discussions of more general changes in risk. In particular, a risk averse decision maker (i.e. one with a concave utility function) might prefer a probability distribution with the same mean and a higher variance.

? provided the now widely accepted definition of ‘increasing risk’. Their definition corresponds to the following three statements, which they prove to be equivalent:¹⁰

Definition: A random variable \tilde{Y} is more risky than a random variable \tilde{X} with the same mean if and only if either of the following equivalent statements holds:

- Every risk averse expected utility maximizer prefers \tilde{X} to \tilde{Y} , i.e. $Eu(\tilde{X}) > Eu(\tilde{Y})$ for all u concave.
- \tilde{Y} equals \tilde{X} plus some noise. More precisely, \tilde{Y} has the same distribution as $\tilde{X} + \tilde{Z}$ for some \tilde{Z} satisfying $E(\tilde{Z}|\tilde{X}) = 0$ (uncorrelated noise).¹¹
- \tilde{Y} has more weight in the tails than \tilde{X} .

In particular, these conditions are met by a mean preserving spread which is characterized by the second bullet point. Today, this notion of increasing risk aversion is also known as second order stochastic dominance: the (less risky) random variable \tilde{X} in Definition 1 stochastically dominates the random variable \tilde{Y} . First order stochastic dominance is a statement about higher or lower payoff in a stochastic setting. A random variable \tilde{A} first order stochastically dominates a random variable \tilde{B} if every decision maker with a monotonically increasing utility function prefers lottery \tilde{A} over lottery \tilde{B} . Equivalently, the random variable \tilde{A} first order stochastically dominates the random variable \tilde{B} if the cumulative distribution function characterizing \tilde{A} takes on smaller values than the cumulative distribution function characterizing \tilde{B} at all consumption levels.

¹⁰? prove the equivalence result and argue that the statements correspond to a well defined notion of increasing risk. They do not actually define ‘more risky’.

¹¹ $E(\tilde{Z}|\tilde{X})$ is again a random variable as we only take expectation over \tilde{Z} . The statement means that the expected value of \tilde{Z} is zero for all realizations of \tilde{X} , i.e. $E(\tilde{Z}|\tilde{X} = x_i) = 0 \forall i$.

4.3 Reaction to Risk

We now return to the model of resource extraction, or of GHG emissions under uncertainty. Here we consider the case where the decision maker does not have the opportunity to insure against risk. We want to examine the effect, on the first period decision, of an increase in risk.

The first order conditions for the agent facing uncertainty are

$$u'(x_1) = u'(x_2) = \mathbb{E} u'(\tilde{\theta} - x_1 - x_2), \quad (4.2)$$

and the conditions for the agent who knows (or act as if) $\tilde{\theta}$ takes on the value $\mathbb{E}\theta$ with certainty are

$$u'(x_1) = u'(x_2) = u'(\mathbb{E}\tilde{\theta} - x_1 - x_2). \quad (4.3)$$

Denote the solutions of equation (4.2) by x_1^u and x_2^u and the solutions of equation (4.3) by x_1^c and x_2^c . We assume that the second order conditions are satisfied.

Inspection of equations (4.2) and (4.3) shows that the only difference between them is that in the first equation the expectation operator is outside the marginal utility function while for the second equation it is inside that function. This observation should bring to mind Jensen's Inequality. We first consider the case where marginal utility u' is strictly convex. Here Jensen's Inequality implies that

$$\mathbb{E} u'(\tilde{\theta} - x_1^c - x_2^c) > u'(\mathbb{E}\tilde{\theta} - x_1^c - x_2^c) \quad (4.4)$$

for any non-degenerate probability distribution of $\tilde{\theta}$ (that is any probability distribution that does not concentrate all the weight on a single point – which would correspond to certainty). Together with equation (4.3) we find

$$\mathbb{E} u'(\tilde{\theta} - x_1^c - x_2^c) > u'(x_1^c) = u'(x_2^c).$$

The only way to satisfy the first order conditions for the decision maker facing uncertainty (equation 4.2) is therefore to lower x_1 and x_2 compared to the optimal values x_1^c and x_2^c for the decision maker facing no uncertainty; this conclusion uses the fact that marginal utility is decreasing in consumption. In words, the marginal utility in the first period at the x^c level is relatively

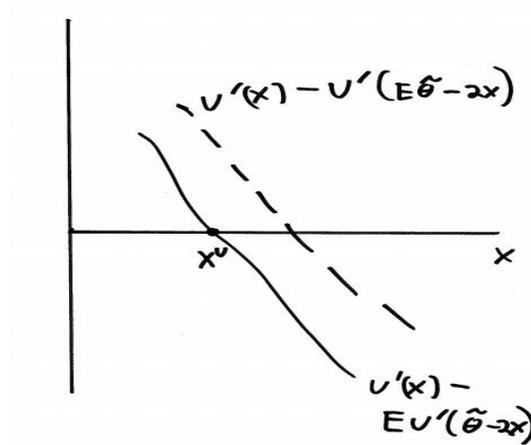


Figure 1 Reaction to Risk.

too low and we can increase overall welfare by decreasing first period welfare (increasing marginal welfare) and increasing second period consumption.

A graphical treatment illustrates the roles of both Jensen's Inequality and the second order condition in obtaining the conclusion that the introduction of uncertainty decreases first and second period consumption when marginal utility is convex, i.e. for $x_1^u < x_1^c$. Using the first part of equation (4.2) to write $x_1 = x_2 = x$, we can write the second part of the equation as

$$u'(x) - E u'(\tilde{\theta} - 2x) = 0. \quad (4.5)$$

The solid curve in Figure 1 shows the graph of the left side of equation (4.5). The curve decreases in x due to the second order condition. The intersection of the curve and the x axis shows the optimal level of first and second period consumption, x^u . Similarly, we can write equation (4.3) as

$$u'(x) - u'(E\tilde{\theta} - 2x) = 0.$$

By the second order condition, the graph of the left side of this equality (the dashed curve in Figure 1) is also downward sloping, and by equation (4.4) this graph lies above the solid graph. Therefore, the intersection of the dashed graph and the x axis, shown as x^c , is greater than x^u .

Analogous reasoning shows that the introduction of uncertainty increases initial consumption if and only if marginal utility is concave. For a utility

function that is three times differentiable these conditions are equivalent to $u''' > 0$ for the case of convex marginal utility and decreasing initial consumption, and to $u''' < 0$ for the case of concave marginal utility and increasing initial consumption. Because a decision maker with $u''' > 0$ reduces consumption in the face of uncertainty this condition is also known as *prudence*.

4.4 A characterization of prudence

The following axiomatic characterization of prudence, i.e. the convexity of marginal utility, helps to understand the concept. Assume that the preference relation \succeq over consumption or wealth x has an expected utility representation with a three times differentiable utility function satisfying $u' > 0$ and $u'' < 0$. Let $k > 0$ be a sure loss (of consumption or wealth) and let $\tilde{\epsilon}$ be a random variable with $E\tilde{\epsilon} = 0$. A decision maker is said to be prudent, i.e. $u''' > 0$, if and only if

$$\left(\frac{1}{2}, x - k; \frac{1}{2}, x + \tilde{\epsilon}\right) \succeq \left(\frac{1}{2}, x; \frac{1}{2}, x - k + \tilde{\epsilon}\right) \quad (4.6)$$

for all $k > 0, x, \tilde{\epsilon}$. The meaning of the brackets is “(probability 1, outcome 1; probability 2, outcome 2)”. That is, on the left we have a lottery that yields with probability $\frac{1}{2}$ the outcome $x - k$ (a sure loss) and with probability $\frac{1}{2}$ it yields $x + \tilde{\epsilon}$. This lottery is preferred (\succeq) to a lottery that yields, with equal probability, x or $x - k + \tilde{\epsilon}$. We can also write these lotteries as probability trees. Then condition (4.6) becomes

$$\begin{array}{c} \frac{1}{2} \swarrow \\ x + \tilde{\epsilon} \\ \frac{1}{2} \searrow \\ x - k \end{array} \succeq \begin{array}{c} \frac{1}{2} \swarrow \\ x \\ \frac{1}{2} \searrow \\ x - k + \tilde{\epsilon} \end{array}$$

for all $\tilde{\epsilon}, k > 0, x$.

The sure loss k and the randomness of ϵ both harm the risk averse decision maker. A prudent decision maker prefers to have one harm (determined by the lottery) with certainty rather than facing a lottery that causes either no harm or both harms at the same time. The two harms are ‘mutually aggravating’ to the decision maker.

Another way to interpret prudence is that the decision maker prefers to face the random term ϵ at the higher consumption level rather than the lower consumption level. This idea is linked to a point discussed in section 4.7, the fact that risk aversion might change in wealth. See e.g. ? for details and for similar characterizations of (all) higher order derivatives of the utility function. Also, combine equation (4.2) with Jensen's Inequality to find that, indeed, an optimal allocation plan under uncertainty shifts consumption from the first two periods to the third so that risk strikes the agent at a higher welfare level, where it does less harm to a prudent agent.

4.5 Reaction to Increasing Risk*

We now consider the more general case: the agent begins with some risk over the resource stock or the damage parameter and then experiences an increase or decrease in risk. The equivalence of the various expressions of (?) definition of increasing risk makes this generalization a straightforward exercise that follows almost line by line the reasoning in section 4.3.

We introduce a new random variable $\tilde{\theta}^*$ and assume that $\tilde{\theta}$ is more risky than $\tilde{\theta}^*$. That is, we assume that $\tilde{\theta}$ is equivalent to $\tilde{\theta}^*$ plus some (uncorrelated) noise or – equivalently – that $\tilde{\theta}$ has more weight in the tails than $\tilde{\theta}^*$. The first order conditions for the $\tilde{\theta}^*$ scenario are analogous to those for $\tilde{\theta}$:

$$u'(x_1) = u'(x_2) = E u'(\tilde{\theta}^* - x_1 - x_2). \quad (4.7)$$

To conserve notation we denote the solutions of the more risky $\tilde{\theta}$ scenario (equation 4.2) by x_1^u and x_2^u and those of the less risky $\tilde{\theta}^*$ scenario (equation 4.7) by x_1^c and x_2^c .

For u' weakly concave (rather than strictly convex), equation (4.4) is replaced by

$$E u'(\tilde{\theta} - x_1^c - x_2^c) \leq E u'(\tilde{\theta}^* - x_1^c - x_2^c).$$

It is worth emphasizing that here the inequality is not a consequence of Jensen's Inequality, but because of (?) proof that more weight in the tails (or more noise) is equivalent to the fact that $E v(\tilde{\theta}) < E v(\tilde{\theta}^*)$ for all v

concave.¹² Here we set $v = u'$. Together with equation (4.7) we obtain

$$\mathbb{E} u'(\tilde{\theta} - x_1^c - x_2^c) \leq (u'(x_1^c) + u'(x_2^c)).$$

The rest of the proof is exactly the same as in section 4.3, except that here the inequality is reversed because of our assumption that marginal utility is concave. The only way the decision maker facing the $\tilde{\theta}$ scenario can satisfy his first order condition (equation 4.2) is to raise x_1 and x_2 relative to the optimal choices x_1^c and x_2^c of the decision maker facing the θ^* scenario.

Thus, increasing risk in the sense of ? has the same effect as does introducing risk to a risk-free environment. A decision maker with a concave marginal utility function ($u''' < 0$) increases her first (and second) period consumption when uncertainty increases. Again, the analogous reasoning implies greater risk reduces initial consumption if and only if marginal utility is convex ($u''' > 0$).¹³ Obviously, if we move from the $\tilde{\theta}$ scenario to the θ^* scenario we have a decrease in risk and the opposite reactions of the decision maker.

4.6 Insurance against Risk

Because uncertainty reduces welfare, the decision maker is willing to pay an insurance premium $w > 0$ to eliminate the risk. Assuming that the decision maker pays for the insurance in period 1 in units of x_1 she is willing to pay w to eliminate the risk provided that

$$\begin{aligned} & \max_{x_1, x_2} u(x_1 - w) + u(x_2) + u(\mathbb{E}\tilde{\theta} - x_1 - x_2) \\ & \geq \max_{x_1, x_2} u(x_1) + u(x_2) + \mathbb{E} u(\tilde{\theta} - x_1 - x_2). \end{aligned} \quad (4.8)$$

The first order conditions for the second maximization problem, describing the scenario with uncertainty, are

$$u'(x_1) = u'(x_2) = \mathbb{E} u'(\tilde{\theta} - x_1 - x_2). \quad (4.9)$$

¹²Subtracting a constant from both random variables does not alter the fact that they have the same mean and that the random variable obtained by adding noise has more weight in the tails.

¹³Reversing (?) inequality is not as straightforward as in the case of Jensen's Inequality. The intuition is that if every risk averse decision maker prefers the less risky scenario over the more risky scenario, then every risk loving decision maker prefers the more risky scenario over the less risky one (given that both have the same expected value).

In the first maximization problem there is no uncertainty but the decision maker pays the insurance premium. The first order conditions are

$$u'(x_1 - w) = u'(x_2) = u'(E\tilde{\theta} - x_1 - x_2). \quad (4.10)$$

For a given utility function and a given probability distribution over $\tilde{\theta}$ we can solve equation (4.9) for x_1 and x_2 and plug the values back into the overall welfare function to find the maximal (ex-ante) welfare that can be attained under uncertainty. In general, the calculation has to be performed *numerically*. We can calculate x_1 and x_2 in the insurance case as functions of w from equation (4.10) and find the maximized welfare given the insurance premium w by plugging these functions back into the welfare function. Requiring equality in equation (4.8) and solving for w gives the maximum premium the decision maker will pay to eliminate the risk.

4.7 Common functional forms*

This subsection discusses common functional forms used to characterize utility and, thus, capture risk aversion. The function

$$u(x) = \frac{1 - e^{-\alpha x}}{\alpha} \quad (4.11)$$

gives rise to a constant coefficient of absolute risk aversion

$$A(x) \equiv -\frac{u''(x)}{u'(x)} = \alpha.$$

This functional form is generally referred to as CARA utility.

It is widely believed that the degree of *absolute* risk aversion decreases in wealth, as is the case with the class of constant *relative* risk aversion (CRRA) utility:

$$u(x) = \frac{x^{1-\gamma} - 1}{1-\gamma}, \quad (4.12)$$

This function is defined only for $x \geq 0$. The coefficient of relative risk aversion is

$$R(x) \equiv -\frac{u''(x)}{u'(x)}x = \gamma,$$

a constant. For this function the coefficient of absolute risk aversion $A(x) = \frac{R(x)}{x} = \frac{\gamma}{x}$ decreases in wealth.

A positive affine transformation of utility u (i.e. a positive multiplicative or an additive constant) does not alter decisions, so the denominators and the ± 1 in $u(x) = \frac{1-e^{-\alpha x}}{\alpha}$ and $u(x) = \frac{x^{1-\gamma}-1}{1-\gamma}$ are redundant. Nevertheless, there are three reasons for keeping these terms. First, their presence means that we do not need to change the function in order to maintain positive marginal utility when α or $1-\gamma$ are smaller than zero. Second, the denominators simplify the expressions for marginal utility, a function that is often of more interest than u itself. Third, the constants are necessary if we want to extend the domain of the parameters α and γ to include 0 respectively 1 in a meaningful way. These normalizing constants imply, using l'Hospital's rule, that $\lim_{\alpha \rightarrow 0} \frac{1-e^{-\alpha x}}{\alpha} = x$ and $\lim_{\gamma \rightarrow 1} \frac{x^{1-\gamma}-1}{1-\gamma} = \ln x$.

To illustrate the use of the CRRA functional form, consider the case where expected income is y and actual income is $y \pm \epsilon$, each with probability 0.5. Define r as the risk premium, the amount that society would spend to stabilize income at its mean value. The income shock is ϵ . Define $w = r/\epsilon$, the risk premium relative to the income shock. Using a second order Taylor expansion of the CRRA utility function, evaluated at $\epsilon = 0$, we obtain an approximation for the maximum risk premium that society would pay to eliminate the risk, relative to the income shock:

$$w = \frac{r}{\epsilon} \approx \frac{1}{\gamma} \left(-\frac{y}{\epsilon} + \sqrt{\left(\frac{y}{\epsilon}\right)^2 + \gamma^2} \right) \quad (4.13)$$

(See exercise xx.) This formula shows that the risk premium, relative to the size of the risk, depends on both the income level, y , and the aversion to risk, γ . The ratio y/ϵ is an inverse measure of the amount of risk relative to baseline income. As this risk becomes small (i.e. as $y/\epsilon \rightarrow \infty$), $w \rightarrow 0$. As the risk approaches its maximum, (i.e. as $y/\epsilon \rightarrow 1$), $w \rightarrow \frac{1}{\gamma} \left(-1 + \sqrt{1 + \gamma^2} \right)$, a quantity that varies between 0 and 1 as γ increases from 0 to ∞ . For example, if $\gamma = 2$ (a value sometimes proposed for climate policy models) then the approximation of w given by equation (reference here) exceeds 0.4 provided that $\epsilon > y/2.1$. Thus, moderate levels of risk aversion correspond to large values of w when the relevant risk, ϵ , is large.

A more general utility function, with harmonic absolute risk aversion (HARA),

subsumes both the CARA and CRRA functions:

$$u(x) = \xi \left(\eta + \frac{x}{\gamma} \right)^{1-\gamma} \quad (4.14)$$

implying

$$A(x) = \left(\eta + \frac{x}{\gamma} \right)^{-1}, \text{ and} \quad (4.15)$$

$$R(x) = x \left(\eta + \frac{x}{\gamma} \right)^{-1}.$$

The name HARA arises from the fact that absolute risk tolerance – the inverse of absolute risk aversion (i.e. the denominator in equation 4.15) – is linear. We obtain the CARA special case by letting γ approach infinity and defining $\alpha = \frac{1}{\eta}$. We obtain the CRRA special case by setting $\eta = 0$. Another frequently used functional form also contained in the HARA class (for $\gamma = -1$) is the quadratic utility function $u(x) = ax - bx^2$. For this functional form, utility is decreasing in wealth for $x > \frac{a}{2b}$. The simplicity of this functional form explains its frequent use. If we are interested in only small changes around a fixed wealth level we can think of the quadratic form as a second order approximation; risk aversion is a second order characteristic of the utility function. Also, in situations where too much of a good is just as bad as too little of it, quadratic utility might be a reasonable assumption. With quadratic utility, preferences can be rewritten as a linear function of the mean and the variance of the underlying random variable.

4.8 Anticipated Learning

This section discusses how anticipated learning affects optimal decisions. We consider two scenarios: complete learning, and a generalization, partial learning. With complete learning, a planner knows, at the time of making a first decision, that she will learn the value of a unknown parameter before having to make a second decision. Our objective is to understand how this anticipated learning changes the first period decision, relative to the case where the planner does not expect to learn the value of the parameter before having to make the second decision. With partial learning, the planner acquires information about the unknown parameter, without learning its exact value. Here our objective is to compare the effect, on first period decisions, of different levels of learning.

The preceding sections study the problem of a planner who chooses an optimal consumption path facing uncertainty about the size of the resource stock, or in the alternative interpretation, about the damage caused by GHG emissions. This optimal plan does not incorporate the possibility that the planner might learn about the unknown parameter as time evolves. That neglect is rational if the planner either knows that she will learn the total resource stock only after making her final decision, or if she has to commit to decisions before observing the stock/damage. Most of the following analysis as well as extensions can be found in ? and ?. ? also extend this model to analyze effects of ambiguous uncertainty. Cite Gollier, Jullien and Treich) provide the analysis of the model of partial learning.

4.8.1 Complete learning

Here we analyze the situation where the decision maker anticipates learning the value of the uncertain parameter after the first period but before her second period choice. In the resource extraction model we can think of this change as an *earlier resolution of uncertainty*, i.e. uncertainty resolves at the beginning of the second period rather than at the beginning of the third period. In the GHG model we can think of moving to a more *flexible decision process* that allows the social planner to react to observations.

The role of uncertainty in the climate change debate is sometimes modeled using the setting in Chapter 4, where there is uncertainty but not anticipated learning. It matters not only that the decision maker learns, but also that

she anticipates this in making early decisions. Without the anticipation of learning, she would solve the same problem as in Chapter 4 in the first period. Then, when she learns about the actual size of the resource stock or the damage parameter, she will make a new plan for second period consumption. In contrast, in the setting here, the decision maker takes into account her reaction to the resolution of uncertainty in the second period when she picks first period consumption.

Climate policy will unfold over a significant period of time, at least several decades. During this period, we are likely to get information both about abatement costs and about climate-related damages. It is important that future decision makers take their new information into account, and that current decision makers understand that future decision makers will do so.

The formal difference in the setting with anticipated learning is that the planner maximizes over the second period consumption before she takes the expected value. She understands that by the time she picks x_2 she will know the realization θ of $\tilde{\theta}$. At period 1 she therefore takes the expectation over welfare given an optimal adjustment of x_2 to the realization of $\tilde{\theta}$. At period 1 the optimization problem is

$$\begin{aligned} & \max_{x_1} \mathbb{E} \max_{x_2} v(x_1, x_2, \tilde{\theta}) \\ & = \max_{x_1} [u(x_1) + \mathbb{E} \max_{x_2} u(x_2) + u(\tilde{\theta} - x_1 - x_2)]. \end{aligned}$$

We solve this problem recursively. The first order condition for x_2 given x_1 and a realization θ is

$$u'(x_2) = u'(\tilde{\theta} - x_1 - x_2). \quad (4.16)$$

By solving this equation we find a function $x_2(x_1, \theta)$. Because u' is strictly monotonic, it is optimal to perfectly smooth consumption over periods 2 and 3 in the resource model, respectively to pick x_2 such that marginal damage equals marginal benefits, i.e.

$$\begin{aligned} x_2 & = \theta - x_1 - x_2 \\ \Leftrightarrow x_2 & = \frac{\theta - x_1}{2}. \end{aligned}$$

We substitute this decision rule into the first period objective, take the expected value, and then maximize with respect to x_1 using the period-2 first

order condition. The resulting first period first order condition is

$$\begin{aligned}
u'(x_1) &= -\mathbb{E} \left\{ u' \left(x_2(x_1, \tilde{\theta}) \right) \frac{dx_2}{dx_1} - u' \left(\tilde{\theta} - x_1 - x_2(x_1, \tilde{\theta}) \right) \left(1 + \frac{dx_2}{dx_1} \right) \right\} \\
\Leftrightarrow u'(x_1) &= -\mathbb{E} \left\{ -u' \left(x_2(x_1, \tilde{\theta}) \right) \right\} \\
\Leftrightarrow u'(x_1) &= \mathbb{E} u' \left(\frac{\tilde{\theta} - x_1}{2} \right). \tag{4.17}
\end{aligned}$$

We obtain the second line by using equation (4.16) to cancel equal terms. Comparing the first order conditions here with those of Chapter 4 enables us to determine how the anticipation of learning (or the early resolution of uncertainty) changes consumption (extraction/emissions) in the first period.

We denote the optimal consumption levels under uncertainty without learning as x_1^u and x_2^u . We initially assume that the decision maker is prudent, i.e. $u''' > 0$. By equation (4.9) we know that the optimal decisions satisfy

$$u'(x_1^u) = u'(x_2^u) = \mathbb{E} u'(\tilde{\theta} - x_1^u - x_2^u). \tag{4.18}$$

Therefore the first equality in the following calculation has to hold

$$\begin{aligned}
u'(x_1^u) &= \mathbb{E} u'(\tilde{\theta} - x_1^u - x_2^u) = \mathbb{E} \left[\frac{1}{2} u'(\tilde{\theta} - x_1^u - x_2^u) + \frac{1}{2} u'(x_2^u) \right] \tag{4.19} \\
&> \mathbb{E} u' \left(\frac{1}{2}(\tilde{\theta} - x_1^u - x_2^u) + \frac{1}{2} x_2^u \right) = \mathbb{E} u' \left(\frac{\tilde{\theta} - x_1^u}{2} \right).
\end{aligned}$$

The inequality sign holds by the assumed convexity of u' . The only way to satisfy the first order condition (4.17) for the learning scenario is by increasing x_1 above the level x_1^u ; this change decreases $u'(x_1)$ and increases $\mathbb{E} u'(\frac{\tilde{\theta} - x_1}{2})$ until both of them are equal for some $x_1^l > x_1^u$.

It is worth remembering how this argument implicitly uses the second order condition. We relate back to Figure 1. The first order condition under uncertainty without learning implies

$$\begin{aligned}
u'(x_1^u) - \mathbb{E} u'(\tilde{\theta} - x_1^u - x_2^u) &= \tag{4.20} \\
u'(x_1^u) - \mathbb{E} \left[\frac{1}{2} u'(\tilde{\theta} - x_1^u - x_2^u) + \frac{1}{2} u'(x_2^u) \right] &= 0.
\end{aligned}$$

The assumed convexity of u' implies

$$\begin{aligned} & \mathbb{E} \left[\frac{1}{2} u'(\tilde{\theta} - x_1^u - x_2^u) + \frac{1}{2} u'(x_2^u) \right] \\ & > \mathbb{E} u' \left(\frac{1}{2} (\tilde{\theta} - x_1^u - x_2^u) + \frac{1}{2} x_2^u \right). \end{aligned}$$

The second order condition for maximization implies that the graph of the first order condition under uncertainty without learning (the solid graph in Figure 1), as a function of x_1 , has a negative slope, at least in the neighborhood of the equilibrium. Moving from the scenario with uncertainty and no learning to the scenario with learning causes this graph to shift upward, thus increasing the intersection of the graph and the x_1 axis, increasing the equilibrium value of x_1 .

At the optimal levels for a decision maker who does not anticipate learning, the expected marginal utility in the second period is relatively too low and the marginal utility in the first period is relatively too high. Increasing x_1 at the cost of decreasing x_2 brings the marginal conditions into balance. Thus, the anticipation of an early resolution of uncertainty makes the prudent decision maker want to consume more in the initial period than in a scenario where uncertainty is resolved late, or where the early resolution is not anticipated. Again, for a decision maker who is not prudent (i.e. with $u''' < 0$) the inequalities in equations (4.18) and (4.20) flip. Thus, anticipated learning causes a decision maker who is not prudent to consume less in the first period, relative to the same decision maker who does not anticipate learning.

Exercise xx asks the student to show that for a prudent decision maker the first period consumption in the face of uncertainty with anticipated learning is below the level of the agent who faces no uncertainty, i.e. $x_1^u < x_1^l < x_1^c$. The argument is similar to the one used in deriving the other two inequalities. Simply rewrite one of the two first order conditions in a suitable way to apply Jensen's Inequality.

Learning has a similar effect as a decrease in risk. The fact that under learning the decision maker observes the random variable $\tilde{\theta}$ before making her choice of x_2 permits her to partially offset the possible low level of the resource or the high level of the damage. That ability to make adjustments effectively reduces her risk over outcome fluctuation by smoothing the joint volatility of the terms $u(x_2)$ and $u(\tilde{\theta} - x_1 - x_2)$. Thus, we can think about learning

as effectively decreasing risk. This perspective also gives a straightforward intuition for why $x_1^u < x_1^l < x_1^c$.

4.8.2 Remark on Discounting*

Our model assumes that the agent does not discount, or that there is a two-phase setting with a suitable time horizon. Discounting might overturn some of the earlier results. It is easiest to think of the resource extraction interpretation of the model. We consider the case of a prudent decision maker. A discounting decision maker values the future less than the present. In our model, where the decision maker does not receive interest payments, discounting causes her to prefer to consume more in period 2 than in period 3.

If there is uncertainty that resolves in period 3, volatility strikes the utility function around the level that the decision maker expects to consume in period 3. If uncertainty instead resolves at the end of period 1, second period consumption adjusts to the new information. Therefore, learning makes second period consumption volatile, causing second period consumption to absorb some of the third period volatility. With learning, the volatility strikes the utility function at the period 2 consumption level, which in expectation can be higher than expected third period consumption, relative to the previous scenario of late resolution. In particular, if prudence is much higher at the higher consumption level, the (prudent) decision maker might want to decrease consumption when uncertainty resolves earlier. ? construct such an example. They use a quite special utility function and a discount factor $\beta < \frac{1}{2}$ (huge!) to get this result. The authors also prove that for HARA utility functions (see section 4.7) the earlier result always holds: a prudent decision maker increases initial consumption if uncertainty resolves early rather than late. Their model corresponds to our model enriched by a discount rate (and it permits the agent to receive interest payments).

4.8.3 Partial Learning*

Here we study the generalization where the decision maker anticipates learning something, rather than everything, about the unknown parameter, θ . A decision maker obviously obtains more information in the scenario where she learns everything rather than nothing about the unknown parameter. In a more general setting, we need a way of ranking information structures, so that we can say that one structure is more informative than another. Given such a definition, we can examine how the anticipation of the receipt of better information affects first period decisions. We assume here that the unknown parameter $\tilde{\theta}$ has a discrete subjective distribution, with possible realizations θ_i , $i = 1, 2, \dots, m$.

Suppose that it is possible to conduct an experiment that teaches us something about the true value of $\tilde{\theta}$. The outcome of the experiment, before it takes place, is a unknown, so we regard the outcome as a random variable. The fact that the outcome tells us something about the true value of $\tilde{\theta}$ means that the two are correlated. Thus, we speak of the experiment as a random variable \tilde{y} that is correlated with $\tilde{\theta}$. The conditional probability of $\tilde{\theta}$ given that the outcome of the experiment is y is $\pi_y(\theta_i)$:

$$\pi_y(\theta_i) = \Pr(\tilde{\theta} = \theta_i \mid \tilde{y} = y).$$

Just as there are different but equivalent ways of expressing the meaning of “more risky” or “more risk averse”, there are different but equivalent ways to compare the information content of an experiment. We need some notation. Define π_y as the m -dimensional vector whose i 'th element is $\pi_y(\theta_i)$ and define the simplex

$$\Delta_\pi = \left\{ \pi_y \in R_+^m \mid \sum_i \pi_y(\theta_i) = 1 \right\}.$$

Consider two possible experiments, equivalently, two random variables, \tilde{y} and \tilde{y}' . Experiment \tilde{y} is said to be more informative than experiment \tilde{y}' if and only if, for any convex function $\phi(\pi_y)$ defined on Δ_π

$$E_{\tilde{y}}[\phi(\pi_{\tilde{y}})] \geq E_{\tilde{y}'}[\phi(\pi_{\tilde{y}'})]. \quad (4.21)$$

Each of the experiments (the random variables \tilde{y} and \tilde{y}') has various possible outcomes, and each outcome tells us something about the unknown parameter θ . That information leads to a conditional distribution π_y , for which

there is a corresponding payoff $\phi(\pi_y)$. The information structure \tilde{y} is more informative than the information structure \tilde{y}' , i.e. the decision maker prefers the experiment \tilde{y} to \tilde{y}' , if and only if inequality 4.21 holds for all convex functions $\phi(\pi_y)$.

This definition is not intuitive. In an attempt to help develop intuition, we digress to consider the case where $n = 2$, i.e. where θ can take on only two values. This binary setting enables us to use a two-dimensional graph to illustrate the role of convexity. To this end, suppose that the two possible realizations are $\theta_1 = 2$ and $\theta_2 = 4$. In this binary case, a single number, $p = \Pr\{\tilde{\theta} = 2\}$, (rather than a vector) describes the prior beliefs. For this example the ex ante (before the receipt of information) mean of $\tilde{\theta}$ is $4 - 2p$ and the variance is $4p(1 - p)$. For concreteness, suppose that $p = 0.5$, the value that maximizes the ex ante variance. If there was complete learning, then after the resolution of uncertainty the decision maker would know the true value of θ with certainty. In this context, that certainty means that either $p = 0$ or $p = 1$, depending on the result of the “experiment”.

Now consider two experiments, represented by the random variables \tilde{y} and \tilde{y}' , which both lead to partial learning. In this binary example, there is no loss of generality in assuming that both experiments are binary random variables, yielding a low outcome with probability α and a high outcome with probability $1 - \alpha$. Under experiment \tilde{y} the conditional probability after observing the low signal is p'_1 ; this conditional probability occurs with probability α . The conditional probability after observing the high signal is p'_2 ; this conditional probability occurs with probability $1 - \alpha$. The expectation of the conditional probabilities must equal the ex ante probability, i.e.

$$\alpha p'_1 + (1 - \alpha) p'_2 = p = 0.5.$$

This equality states that the decision maker’s expectation, before observing the signal, of his ex post belief equals his ex ante belief. Figure 2 shows the conditional probabilities, p'_1 and p'_2 . These probabilities lie on either side of the ex ante probability $p = 0.5$.

If the experiment \tilde{y}' led to complete information, then we would have $p'_1 = 0$ and $p'_2 = 1$. If the experiment provided no information, then we would have $p'_1 = 0.5 = p'_2$. The fact that the experiment leads to only partial information explains why the values p'_1 and p'_2 lie between 0 and 1, on either side of the ex ante probability 0.5. As the experiment becomes more informative, it

should seem at least plausible to the reader that the values p'_1 and p'_2 move away from $p = 0.5$ toward the states of complete certainty, $p = 0$ and $p = 1$. Denote the ex post probabilities under the more informative experiment, \tilde{y} , as p_1 and p_2 . Figure 2 shows these values, with

$$0 < p_1 < p'_1 < 0.5 < p'_2 < p_2 < 1. \quad (4.22)$$

These inequalities imply that both experiments provide some learning, but not complete learning, and that experiment \tilde{y} is more informative than \tilde{y}' .

Now we ask “Under what conditions does a decision maker prefer the more informative experiment?” Recall that in this example, with the experiment \tilde{y}' the decision maker faces p'_1 with probability α and he faces p'_2 with probability $1 - \alpha$. With the experiment \tilde{y} he faces p_1 with probability α and he faces p_2 with probability $1 - \alpha$. Figure 2 shows a solid convex graph and a dashed concave graph. The relative positions of these two graphs is immaterial; all that matters is that one is concave and the other convex. If the convex graph describes the decision maker’s payoff, as a function of the ex post probability, then he prefers the more informative signal. If the concave function were to describe his payoff, as a function of the ex post probability, then he would prefer the less informative signal. Again, the purpose of this example and this graph is to explain why the statement that one experiment is more informative than a second experiment involves the expectation of a convex function.

We can push this example a bit further in order to make a second point. Before learning, the decision maker views his ex post belief as a random variable. Denote the random variable “ex post belief” as \tilde{p} . Under experiment \tilde{y} the realizations of this random variable are p_1 and p_2 , and under the experiment \tilde{y}' the realizations are p'_1 and p'_2 . Both of these random variables have the same expectation, because in both cases the expectation of the ex post belief must equal the ex ante belief. In the binary example it should be clear that the variance of \tilde{p} is greater under the more informative signal. To demonstrate this claim formally, note that the variance under the more informative signal is

$$\sigma_{\tilde{y}}^2 = \alpha (p_1 - 0.5)^2 + (1 - \alpha) (p_2 - 0.5)^2$$

and the variance under the less informative signal is

$$\sigma_{\tilde{y}'}^2 = \alpha (p'_1 - 0.5)^2 + (1 - \alpha) (p'_2 - 0.5)^2.$$

The ranking

$$\sigma_{\tilde{y}}^2 > \sigma_{y'}^2$$

is a consequence of inequality 4.22.

The statement that better information leads to a higher variance strikes many readers as counter-intuitive. However, it is important to keep in mind what the variance here refers to: it is the variance of what we think today our beliefs will be tomorrow. If we do not expect to obtain new information, then we think that our beliefs tomorrow will be the same as our beliefs today. If we expect to obtain significant information, then we think that our beliefs tomorrow will be quite different from our beliefs today. However, since we do not yet know what that information will be, we regard tomorrow's beliefs as highly variable.

Although better information is associated with a larger variance of our posterior beliefs, for this binary example (but not in general) the conditional variance of the outcome of the *underlying random variable*, $\tilde{\theta}$, is lower under better information. For example, if the signal is low, then the conditional variance of the outcome under the more informative signal is $4p_1(1-p_1)$ and the conditional variance of the outcome under the less informative signal is $4p'_1(1-p'_1)$. In view of inequality 4.22, $4p_1(1-p_1) < 4p'_1(1-p'_1)$. We have the same ranking of the conditional variance given a high signal. Therefore, the ex ante expectation of these conditional variances of the underlying random variable is lower under the more informative signal.

We summarize the three main lessons of this binary example as follows:

- A more informative signal leads to a higher spread of ex post probabilities; the decision maker prefers a more informative signal if and only if his ex post expected payoff is a convex function of the probability that defines his ex post belief.
- A more informative experiment leads to a higher ex ante variance of the random variable “ex post belief” .
- A more informative experiment leads to a lower expectation of the conditional (ex post) variance the underlying random variable $\tilde{\theta}$.

The statement that one experiment is more informative than another implies that the unconditional distribution of θ (before the result of an experiment

is known) is the same under the two experiments. (We used this result above in discussing the binary example.) To verify this claim, note that the linear function $\phi(\pi_y) = \gamma\pi_y$, where γ is a vector of constants, is convex, as is the negative of this function, $\phi(\pi_y) = -\gamma\pi_y$. Therefore, if \tilde{y} is more informative than \tilde{y}' , inequality 4.21 implies that

$$E_{\tilde{y}} \{\gamma\pi_{\tilde{y}}\} = E_{\tilde{y}'} \{\gamma\pi_{\tilde{y}'}\}.$$

Because this equality holds for all γ , it must be the case that $E_{\tilde{y}}\pi_{\tilde{y}} = E_{\tilde{y}'}\pi_{\tilde{y}'}$, i.e. the unconditional distributions are the same. Thus, comparison of the two experiments is “fair”, because we start out with the same beliefs before learning the results of either experiment.

We now apply this definition to the resource/climate change problem. Suppose that the decision maker faces the information structure defined by the random variable \tilde{y} . Her optimization problem in the first period is:

$$\max_{x_1} \left[u(x_1) + E_{\tilde{y}} \max_{x_2} E_{\tilde{\theta}|\tilde{y}} \left(u(x_2) + u(\tilde{\theta} - x_1 - x_2) \right) \right]. \quad (4.23)$$

In choosing first period consumption, the decision maker knows that she will observe the result of the experiment and update her subjective distribution of the unknown parameter θ . Consider the consequence of a particular realization, y , and resulting subjective probability distribution π_y . Given this realization, the second period conditional payoff is

$$\rho(\pi_y; x_1) = \max_{x_2} \sum_{i=1}^m (u(x_2) + u(\theta_i - x_1 - x_2)) \pi_y(\theta_i). \quad (4.24)$$

The function $\rho(\pi_y; x_1)$ is convex in π . This fact, and the definition of “more informative” means that the unconditional (i.e. before the result of the experiment is known) expected maximized value of the second period payoff is higher when the decision maker expects to have better information. To confirm that the function $\rho(\pi_y; x_1)$ is convex in the vector π_y , consider two possible values for this vector, π^1 and π^2 , and a convex combination $\pi^\lambda = \lambda\pi^1 + (1 - \lambda)\pi^2$, with $0 \leq \lambda \leq 1$. Convexity requires showing that

$$\lambda\rho(\pi^1; x_1) + (1 - \lambda)\rho(\pi^2; x_1) \geq \rho(\pi^\lambda; x_1).$$

To establish this inequality, denote the optimal second period action under π^1 , π^2 , and π^λ as, respectively, x^1 , x^2 and x^λ . Now use the fact that x^λ

is feasible, but typically not optimal, under the distributions π^1 and π^2 , to establish that the left side of the inequality is always at least as great as the right side. Exercise xx asks the student to fill in the details of this demonstration.

Substituting the definition in equation 4.24 into the optimization problem 4.23 allows us to write that problem as

$$\max_{x_1} [u(x_1) + E_{\tilde{y}} \rho(\pi_y; x_1)],$$

with first order condition

$$u'(x_1) + E_{\tilde{y}} \frac{d\rho(\pi_y; x_1)}{dx_1} = 0 \quad (4.25)$$

and second order condition

$$u''(x_1) + E_{\tilde{y}} \frac{d^2\rho(\pi_y; x_1)}{dx_1^2} < 0.$$

By the envelope theorem we have

$$\begin{aligned} \frac{d\rho(\pi_y; x_1)}{dx_1} &= \frac{d[\sum_{i=1}^m (u(x_2^*) + u(\theta_i - x_1 - x_2^*)) \pi_y(\theta_i)]}{dx_1} \\ &= \left[\sum_{i=1}^m \frac{du(\theta_i - x_1 - x_2^*)}{dx_1} \pi_y(\theta_i) \right] \end{aligned} \quad (4.26)$$

where x_2^* is the optimal value of x_2 as a function of x_1 and the information y (the result of the experiment).

The comparative statics of the effect of information on the first period decision depends on the curvature of a derivative, $\frac{d\rho(\pi_y; x_1)}{dx_1}$, not on the curvature of $\rho(\pi_y; x_1)$. We encountered a similar result in considering the effect of complete information; there the comparative statics depends on the curvature of $u'(x)$. The function $\frac{d\rho(\pi_y; x_1)}{dx_1}$ might be either convex or concave in π_y . Suppose that this function is concave. In that case, the definition of “more informative” implies that if \tilde{y} is more informative than \tilde{y}'

$$E_{\tilde{y}} \frac{d\rho(\pi_y; x_1)}{dx_1} \leq E_{\tilde{y}'} \frac{d\rho(\pi_y; x_1)}{dx_1}. \quad (4.27)$$

In this chain of reasoning, the function $\frac{d\rho(\pi_y; x_1)}{dx_1}$ plays the role of $\phi(\pi_y)$ that appears in the definition of “more informative”.

Again, it may help to visualize the graph of the left side of the first order condition as a function of x_1 , which by the second order condition has a negative slope at least in the neighborhood of the optimum value of x_1 . Inequality 4.27 implies that the graph of the first order condition under the experiment \tilde{y} lies below the corresponding graph under the less informative experiment \tilde{y}' . Therefore the intersection of the former graph and the x_1 axis is less than the intersection under the latter graph. That is, the anticipation of better information (having the experiment \tilde{y} rather than the experiment \tilde{y}') reduces the first period action x_1 . This result is reversed if $\frac{d\rho(\pi_y; x_1)}{dx_1}$ is convex rather than concave.

The comparative static question considered here is more difficult than the question that arises in comparing the effect of complete information, because here we need to determine the curvature of the endogenous function $\frac{d\rho(\pi_y; x_1)}{dx_1}$ rather than the exogenous function $u'(x)$. These two functions are related, using equation 4.26. For the HARA utility function, $u(x) = \xi \left(\eta + \frac{x}{\gamma} \right)^{1-\gamma}$, better information decreases the optimal x_1 if and only if $0 < \gamma < 1$; for $\gamma < 0$ or $\gamma > 1$ better information increases x_1 . If the utility function is not HARA, the effect of better information is ambiguous.

Jones and Ostroy (1984) {Robert Jones and Joseph Ostroy "Flexibility and uncertainty" Review of Economic Studies 51(1) 13 - 32, 1984} provide an alternative characterization of the effect of better information on first period actions. Define the expected second period value of information as

$$\Lambda(x_1) \equiv E_{\tilde{y}} \rho(\pi_y; x_1) - E_{\tilde{y}'} \rho(\pi_y; x_1).$$

The anticipation of better information decreases the first period action if and only if $\Lambda(x_1)$ is a decreasing function.

In the context of the climate change problem, where x_1 is emissions during the first phase of a decision problem, the slope of $\Lambda(x_1)$ could be either positive or negative. For example, it might be the case that if we allow GHG concentrations to reach a high level (x_1 is large), society will have painted itself into a corner and better information will not be of much use; in contrast, if we keep concentrations relatively low, we will have the flexibility to respond to better information about damages. In this case, $\Lambda(x_1)$ is a decreasing function, so the anticipation that we will obtain better information reduces the optimal level of current emissions. Alternatively, it might be the case that better information will be especially important if future GHG

concentrations are high. For example, with high GHG concentrations the optimal period 2 action might be very sensitive to our belief that the world is near a tipping point; in that case the expected future payoff would likely be sensitive to the receipt of better information. In contrast, with low GHG concentrations we might be quite sure that the world is far from a tipping point, and future actions might then be insensitive to the type of information we receive. In this scenario, the value of information is likely to be increasing in x_1 , so the anticipation of better information increases current emissions. This example illustrates why casual reasoning is not much help in determining how society should respond to the anticipation of future improvements in climate science. The optimal response can be quite sensitive to modeling assumptions. The role of formal analysis is to make these assumptions explicit and then to understand their effect on model predictions.

4.8.4 Learning under risk neutrality

Here we examine the effect of anticipated learning when the unknown parameter enters the payoff linearly. Learning can increase or decrease the first period action. This problem is intrinsically important because in some situations it is natural to assume that the unknown parameter enters the problem linearly. It is also pedagogically useful as a means of showing how to use the machinery developed above to study the effect of anticipated learning. In an important special case, where the primitive functions are quadratic, learning increases the first period action. We can also show the relation between prudence and the effect of anticipated learning. We discuss this problem in the context of a climate change model.

Let emissions in the first and second period be x_1 and x_2 . In the absence of decay, the stock of pollution at the end of the second period is $x_1 + x_2$. The benefit of emissions is the concave increasing function $u(x)$ and the cost of pollution is $\theta v(x_1 + x_2)$, where the parameter θ is unknown and v is convex and increasing. Neglecting discounting, the payoff is

$$u(x_1) + u(x_2) - \theta v(x_1 + x_2). \quad (4.28)$$

The decision maker obtains information about θ between making the first period and the second period decisions. We assume that the non-negativity constraints, $x_i \geq 0$, are not binding, so we ignore those constraints.

The key feature of this problem is that the unknown parameter, θ , enters the payoff linearly. In the absence of learning, the risk neutral decision maker maximizes expression 4.28, replacing θ by its subjective expectation. Thus, in the absence of anticipated learning, uncertainty about θ has no interesting effect on the problem; it merely requires that we replace a random variable by its mean.

Ulph and Ulph (1997?) studied this problem for the case where the primitive functions u and v are quadratic. By obtaining an explicit solution for the first period decision rule, they showed that anticipated learning increases the first period emissions, i.e. the opposite of the Precautionary Principle holds. In most settings, an explicit solution to the first period decision cannot be obtained; in those cases, we can nevertheless obtain some intuition about the problem by using the machinery developed above. Consideration of the more general formulation, where we require only that u be increasing and concave and v be increasing and convex, illustrates the use of this machinery. It leads to a somewhat simpler derivation of the result for the quadratic case, and it also shows the effect of allowing more general functional forms. For example, the opposite of the Precautionary Principle also holds if u is quadratic and $v''' < 0$ or if v is quadratic and $u''' > 0$. (The signs of the third derivatives are opposite in these sufficient conditions, because there is a negative sign in front of v in the objective.)

There are many ways to model learning, but perhaps the simplest is to assume that

$$\theta = \bar{\theta} + \sum_{i=1}^n y_i, \text{ with } y_i \text{ iid } (0, \sigma^2).$$

That is, θ is the sum of a known constant, $\bar{\theta}$, and n unknown parameters, which we model as independently identically distributed random variables, each with mean 0 and variance σ^2 . The decision maker observes the first s of these random variables between making the first and the second decisions: a larger value of s corresponds to more learning.

Define the regulator's expectation of θ , after the receipt of information ("learning"), as $\theta_1 = \bar{\theta} + \sum_{i=1}^s y_i$. Prior to learning, the decision maker treats θ_1 as a random variable with mean $\bar{\theta}$ and variance $s\sigma^2$. In this context, increased learning corresponds to a greater variance of the subjective random variable describing our views in the first period of the information that we will have after learning. In the absence of learning $\theta_1 = \bar{\theta}$, a constant.

It is useful to recall the binary example above. There we saw that increased learning increases the variance of the random variable that defines the beliefs we hold tomorrow. In the binary example, tomorrow's beliefs are completely described by the realization of \tilde{p} . In the example under consideration here, the sufficient statistic to describe tomorrow's beliefs is θ_1 . From the perspective of today, the anticipation of better information (a more informative signal) increases the variance of the random variables that describe our future beliefs. However, more learning, equivalently, a larger value of s , decreases the ex post variance of the underlying random variable $\tilde{\theta}$. In the absence of learning, the second period variance of this random variable is $n\sigma^2$. With learning, the second period conditional variance of this random variable is $(n-s)\sigma^2$, which decreases with s . Again, we see that more learning increases the subjective variance of the random variable that describes our future information; more learning decreases the conditional (ex post) variance of the underlying random variable.

Due to the fact that θ enters the payoff linearly, only its posterior mean, θ_1 , and not the higher subjective moments of θ , enter the second period problem. The entire ex ante (before learning) distribution of θ , and not merely the prior mean ($\bar{\theta}$) affect the problem in the first period.

The regulator's problem after learning, is to maximize

$$E_{\theta} [u(x_2) - \theta v(x_1 + x_2)] = u(x_2) - \theta_1 v(x_1 + x_2).$$

The first order condition to this problem is

$$u'(x_2) - \theta_1 v'(x_1 + x_2) = 0.$$

This first order condition implicitly defines the optimal second period decision, $x_2 = x^*(\theta_1, x_1)$. Differentiating the first order condition once and then again, leads to the expressions for the first and second partial derivatives of the decision rule:

$$\frac{\partial x^*}{\partial \theta_1} = \frac{v'}{u'' - \theta_1 v''} < 0, \quad (4.29)$$

$$\begin{aligned} \frac{\partial^2 x^*}{\partial \theta_1^2} &= \frac{1}{(u'' - \theta_1 v'')^2} \left[(u'' - \theta_1 v'') v'' \frac{\partial x^*}{\partial \theta_1} - v' \left[(u''' - \theta_1 v''') \frac{\partial x^*}{\partial \theta_1} - v'' \right] \right] \\ &= \left(\frac{v'}{u'' - \theta_1 v''} \right)^2 \left[- \left(\frac{u''' - \theta_1 v'''}{u'' - \theta_1 v''} \right) - 2 \left(\frac{-v''}{v'} \right) \right]. \end{aligned} \quad (4.30)$$

The first term in square brackets in the last line of equation 4.30 equals the measure of absolute prudence of the continuation payoff, $u - \theta_1 v$, and the second term is twice the measure of absolute risk aversion of the damage function. Thus, the second period decision rule is concave in the information realization, θ_1 , if and only if absolute prudence with respect to the entire continuation payoff is less than twice the absolute risk aversion associated with the damage function.

We summarize the anticipated amount of information using s , the number of components of θ that the decision maker will learn before taking the second action. The continuation payoff, as a function of the first period decision and the anticipated amount of information, is

$$J(x_1, s) = E_{\theta_1} \{u(x^*(\theta_1, x_1)) - \theta_1 v(x_1 + x^*(\theta_1, x_1))\}.$$

The first order condition for x_1 is

$$u'(x_1) + J_{x_1}(x_1, s) = 0.$$

At the optimal x_1 , the graph of the left side of this equation, as a function of x_1 , has a negative slope, by the second order condition. Better information (a larger value of s) increases the first period level of emissions if and only if it shifts up the second term, i.e. if and only if $J_{x_1, s}(x_1, s) > 0$. We now show how to sign this partial derivative.

Using the envelope theorem, we have

$$J_{x_1}(x_1, s) = E_{\theta_1} \{-\theta_1 v'(x_1 + x^*(\theta_1, x_1))\}.$$

Define the function

$$f(\theta_1; x_1) = -\theta_1 v'(x_1 + x^*(\theta_1, x_1)),$$

so we have $J_{x_1}(x_1, s) = E_{\theta_1} f(\theta_1; x_1)$. We noted that an increase in information corresponds (in this setting) to a larger variance of θ_1 . Using the results from Chapter 3 we conclude that $J_{x_1, s}(x_1, s) > 0$ if and only if $f(\theta_1; x_1)$ is convex in θ_1 . We therefore consider the curvature of this function.

Using equations 4.29 and 4.30, we have

$$\begin{aligned}
f_{\theta_1}(\theta_1; x_1) &= - \left[v'(x_1 + x^*(\theta_1, x_1)) + \theta_1 v''(x_1 + x^*(\theta_1, x_1)) \frac{\partial x^*}{\partial \theta_1} \right] \\
f_{\theta_1 \theta_1}(\theta_1; x_1) &= - \left[2v'' \frac{\partial x^*}{\partial \theta_1} + \theta_1 \left[v''' \left(\frac{\partial x^*}{\partial \theta_1} \right)^2 + v'' \frac{\partial^2 x^*}{\partial \theta_1^2} \right] \right] \\
&= -v'' \left[2 \frac{\partial x^*}{\partial \theta_1} + \theta_1 \left[\frac{v'''}{v''} \left(\frac{\partial x^*}{\partial \theta_1} \right)^2 + \frac{\partial^2 x^*}{\partial \theta_1^2} \right] \right] \\
&= -v'' \left[2 \frac{v'}{u'' - \theta_1 v''} + \theta_1 \left[\frac{v'''}{v''} \left(\frac{v'}{u'' - \theta_1 v''} \right)^2 + \right. \right. \\
&\quad \left. \left. \frac{v'}{(u'' - \theta_1 v'')^2} \left(2v'' - \left((u''' - \theta_1 v''') \frac{v'}{u'' - \theta_1 v''} \right) \right) \right] \right] \\
&= -\frac{v'v''}{u'' - \theta_1 v''} \left[2 + \theta_1 \left[\frac{v'''}{v''} \left(\frac{v'}{u'' - \theta_1 v''} \right) + \right. \right. \\
&\quad \left. \left. \frac{1}{(u'' - \theta_1 v'')} \left(2v'' - \left((u''' - \theta_1 v''') \frac{v'}{u'' - \theta_1 v''} \right) \right) \right] \right] \\
&= -\frac{v'v''}{u'' - \theta_1 v''} [A + B]
\end{aligned}$$

where we define

$$\begin{aligned}
A &= 2 + \theta_1 \frac{2v''}{(u'' - \theta_1 v'')} = \frac{2\theta_1 u''}{(u'' - \theta_1 v'')} > 0 \\
B &= \theta_1 \left(\frac{v'}{u'' - \theta_1 v''} \right) \left[-\frac{(u''' - \theta_1 v''')}{(u'' - \theta_1 v'')} - \left(-\frac{v'''}{v''} \right) \right]. \quad (4.31)
\end{aligned}$$

The fact that $-\frac{v'v''}{u'' - \theta_1 v''} > 0$ implies that $f(\theta_1; x_1)$ is convex in θ_1 (so that learning increases the first period action) if and only if $A + B > 0$. The concavity of u and the convexity of v imply that $A > 0$. Thus, a sufficient condition for $A + B > 0$ is $B \geq 0$. If both u and v are quadratic, then $B = 0$, so learning increases the first period decision.

More generally, we see that $B > 0$ if and only if the term in square brackets in equation 4.31 is negative. This term is the difference between two measures of absolute prudence: the absolute prudence associated with the continuation payoff, $u - \theta_1 v$, (the first term) minus the absolute prudence

associated with the damage function, v (the second term). We conclude that if the decision maker is more prudent with respect to damages than with respect to the continuation payoff, then anticipated learning increases the first period action. If the ranking of prudence is reversed, the effect of anticipated learning on the first period action is ambiguous.

If $u''' = 0$ and $v''' \neq 0$, then $B > 0$ if and only if $v''' < 0$; if instead $v''' > 0$, then $B < 0$ and the sign of $A + B$ is ambiguous. Similarly, if $v''' = 0$ and $u''' \neq 0$ then $B > 0$ if and only if $u''' < 0$; if instead $u''' > 0$, then $B < 0$ and the sign of $A + B$ is again ambiguous.

5 Multiperiod Anticipated Learning

”Active learning” describes the situation where the decision maker manipulates a state variable in order to gain information. For example, a monopoly might believe that the stock of loyal customers (a state variable) changes as a result of advertising, but not know the exact relation between advertising and the change in the stock, i.e. the monopoly might not know a parameter value in the equation of motion. The monopoly might choose its advertising level (a control variable) over time in order to learn about the unknown but fixed parameter value. The model in the Clarke and Mangel, where an animal decides which patch to search for food, contains another example of active learning.

In some settings, even though the decision maker affects the future value of the state variable, which in turn affects the amount of information contained in a random signal, it is too costly to manipulate the state for the purpose of improving information. For example, in a macro economic setting, the magnitude of changes in the money supply might affect the amount of learning about the relation between money supply and inflation. In an environmental setting, the level of the stock of greenhouse gasses might affect the amount of learning about the relation between those stocks and damages. However, the cost of changing the money supply or the stock of greenhouse gasses is so large relative to the value of additional information produced by that change, that the possibility of acquiring information has negligible affect on the choice of the control rule. That is, even though the model does allow for active learning, and even though the anticipation of learning is important to the decision rule, that rule is not chosen for the purpose of acquiring information.

”Passive learning” describes the situation where learning is exogenous to the decision maker’s actions. In this setting, actions taken by the decision maker, such as those that alter the level of the state variable, have no effect on the amount of information. The model structure determines whether learning is active or passive. We explain this relation in the context of a particular problem below.

Modeling learning requires using a state variable to describe information. We discuss two types of approaches, using either a discrete or a continuous distribution. In both cases we use examples rather than a general frame-

work, in order to make the ideas as clear as possible and keep notation to a minimum.

Throughout this section, a central assumption is that the decision maker *anticipates* learning. If the decision maker happens to learn, but does not anticipate that he will learn in the future, the problem is very different (and much simpler and less interesting).

5.1 Discrete distribution

Suppose that the single period payoff is

$$U(c) - \Delta d(S)$$

Where S is the stock of GHG (a state variable) in the current period and c is emissions of GHG (the control variable) in a period. The benefit of emissions is U and the damages associated with the stock is Δd , where d is a known function and Δ is unknown. Section 5.2.1 considers the case where Δ is a random variable, and we learn about its distribution. Section 5.2.2 considers the case where Δ is a fixed parameter and we obtain information about the value of this parameter.

The stock of GHGs evolves according to

$$S' = f(S, c) \tag{5.1}$$

where f is the growth equation and S' is the stock in the next period. Recall our use of the convention that a variable without a time subscript is the value of the variable in an arbitrary period – in many cases, the “current period”, and the variable with a prime is the value of that variable in the next period.

5.2 Learning about Δ

We first discuss two ways to think about learning. We can either assume that Δ is a random variable and we learn about its distribution, or we can treat Δ as a fixed but unknown number, which we learn about over time. Then we consider two ways of solving the problem (using either model of learning): either dynamic programming or stochastic programming.

5.2.1 Δ is a random variable with unknown distribution

First, suppose that Δ is a random variable. For the purpose of generating a tractable model, we assume that Δ is a draw from one of two distributions. For example, we may have competing models of climate change that result in two distributions of the random variable. For modeling purposes, we assume that one of these distributions is correct; we obtain information over time about which of the two distributions is more likely to be the correct one. The problem is to choose the control rule for emissions, taking into account that in the future we will have better information than today concerning which of the two possible distributions is correct.

Denote these two distributions as x_1 and x_2 . Associated with each distribution there are two possible outcomes, G and B . Under the first distribution the probability is q that the realization of Δ is G ; under the second distribution the probability is r that the realization of Δ is G . G and B are numbers, with $G < B$. The two realizations correspond to low damage (G) and the high damage (B) outcomes.

Table 1 gives the outcomes and probabilities associated with these two distributions. These are *conditional probabilities*, the probability of an event conditional on x .

realization of Δ	$x = x_1$	$x = x_2$
G	q	r
B	$1 - q$	$1 - r$

Table 1: The conditional distributions of random variable Δ

If $q < r$ then GHG stocks present a greater danger if the "truth" is $x = x_1$ rather than $x = x_2$.

For the purpose of the model, q and r are taken to be objective probabilities. The decision maker is uncertain which distribution is correct, and at time t has subjective probability p_t that $x = x_1$. The subjective probability at time t that $x = x_2$ is therefore $1 - p_t$. Denote $P(p_t, \Delta_t)$ as the posterior probability, the subjective probability that $x = x_1$ when the prior is p_t and you observe $\Delta = \Delta_t$:

$$p_{t+1} = P(p_t, \Delta_t) = \Pr \{x = x_1 \mid \Delta_t, p_t\}. \quad (5.2)$$

Using Bayes' Rule we can write the values of $P(p_t, \Delta_t)$ for the two possible realizations of Δ :¹⁴

$$P(p_t, G) = \frac{p_t q}{p_t q + (1 - p_t) r} \quad (5.4)$$

$$P(p_t, B) = \frac{p_t (1 - q)}{p_t (1 - q) + (1 - p_t) (1 - r)}. \quad (5.5)$$

In this model, the subjective probability, p is a state variable, the equation of motion of which is given by equation 5.2. Note that the evolution of p is stochastic, since the evolution depends on the realization of a random variable, Δ . Also note that the evolution does not depend on actions that the decision maker takes; learning is passive.

Increasing the number of possible outcomes of Δ does not greatly increase the size of the problem – it just means that we have more possible outcomes, i.e. more equations of the form of equations 5.4 and 5.5. In contrast, increasing the number of possible distributions increases the dimensionality of the state variable, and significantly increases the size of the problem (which has important implications on the feasibility of obtaining a numerical solution). If there are n possible distributions we need $n - 1$ state variables to describe beliefs; each state variable is the subjective probability that a particular distribution describes the world. Since the probabilities sum to 1, we only need $n - 1$ numbers to keep track of the n probabilities.

5.2.2 The "star information structure"

Kolstad (JEEM 1996) uses an alternative called the "star information structure". In this setting Δ is a parameter (rather than a random variable) that takes a particular value, either G or B , but the decision maker does not know which value it takes. Let g be a signal that makes us think it is more likely that $\Delta = G$, and b be a signal that makes us think it is more likely

¹⁴To obtain equation 5.4 and 5.5, use the rule

$$P(A \cap B) = P(A | B) P(B) = P(B | A) P(A)$$

to write

$$P(B | A) = \frac{P(A | B) P(B)}{P(A)}. \quad (5.3)$$

Associate the event A with $\Delta_t = G$ and the event B with $x = x_1$. Equations 5.4 and 5.5 then follow directly from the formula 5.3.

Missing figure

that $\Delta = B$. If at time t our subjective probability that $\Delta = G$ is p_t , then if we observe the signal g we update our probability according to

$$p(p_t, g) = \lambda + (1 - \lambda)p_t.$$

If we observe the signal b we update my probability to

$$p(p_t, b) = 0\lambda + (1 - \lambda)p_t.$$

Our updated subjective probability is a convex combination of the previous probability and the state of being certain ($p = 1$ or $p = 0$). If $\lambda = 1$ then after one observation we know the true value of Δ . If $\lambda = 0$ our observation in the current period provides no information, and we never change our subjective beliefs. As λ ranges from 0 to 1, the signal becomes more informative.

Figure ?? represents the star information structure when there are three (rather than two) possible values of Δ , call them G, B, U and let p_j $j = 1$ for G , $j = 2$ for B , and $j = 3$ for U be the subjective prior probabilities of each of the outcomes. A point such as "A" in the simplex represents a particular prior distribution. If, for example, $\lambda = \frac{1}{2}$ as in the Figure, then if the signal in the present period is g then our subjective belief moves half way from the prior probability, point A , to being certain that $\Delta = G$. The point $(0,1)$ in the simplex represents this state of certainty. The point "a" in the figure is half way between the point A and $(0,1)$; point "a" is our posterior if $\lambda = \frac{1}{2}$, the prior is "A" and the signal is g . Similarly, if the signal in the current period is b then the posterior probability is point "b", and if the signal is u the posterior is point "c".

5.2.3 Comparing the two ways of modeling learning.

The star information structure has the advantage that it uses a single parameter, λ , to model the speed of learning. The disadvantage of that approach is that it does not permit a clear distinction between risk (the fact that there is objective randomness) and uncertainty (the fact that we do not know the probability distribution of a random variable). If Δ is really a parameter that takes a single value (rather than a r.v.), then one observation should be

enough to learn that value, unless there is some other randomness – but that additional randomness is not modeled. For example, as p approaches either 0 or 1, the amount of randomness goes to 0. It seems reasonable to think that the difficulty of learning arises because we receive noisy signals, i.e. because the relation between damages and the stock of GHGs really is stochastic, not merely uncertain. As our uncertainty about this relation diminishes, there remains the objective stochasticity. However, the star information structure assumes that as our uncertainty about the relation diminishes, the objective stochasticity also diminishes.

In contrast, the specification of underlying distributions recognizes that there is exogenous randomness; this fact makes it difficult to learn the true distribution. Even after the decision maker knows the true distribution, i.e. as p approaches 0 or 1, there is still randomness. The disadvantage of this approach is that it is less parsimonious – the modeler has to specify both q and r , instead of the single parameter λ . Nevertheless, this approach is more appealing than the star information structure, because the former is internally consistent. We therefore consider only that approach below.

5.3 Describing the optimization problem

There are two approaches to describing – and then solving – the optimization problem. We can write it as a dynamic programming (DP) or as a stochastic programming (SP) problem. We consider each of these approaches.

5.3.1 The DP problem

We study an autonomous control problem: there is an infinite horizon and no explicit time dependence, apart from discounting at a constant discount factor β . In this problem there are two state variables, S and p . We continue to assume that S evolves deterministically, but it is straightforward to add a random component to equation 5.1. This change requires that in solving the problem we take expectations with respect to the additional random component.

Denote the value function as $J(p, S)$ and write the DPE as

$$J(p, S) = \max_c E_{\Delta} \{U(c) - \Delta d(S) + \beta J(p', S')\}$$

subject to equations 5.1 and 5.2.

In taking expectations of Δ we use the subjective distribution:

$$\begin{aligned} \Pr(\Delta = G) &= pq + (1 - p)r \\ \Pr(\Delta = B) &= 1 - [pq + (1 - p)r]. \end{aligned}$$

The realization of Δ determines not only this period's damages, but also the subjective beliefs in the next period, p' .

Because Δ appears linearly in the current period payoff, we can write

$$E\Delta = G(pq + (1 - p)r) + (1 - [pq + (1 - p)r])B. \quad (5.6)$$

Given p , there are only two possible values of p' , depending on whether the current realization of Δ is G or B . We can list the two possible outcomes of p' , say p'_G and p'_B , corresponding to the outcomes G and B in the current period, and write

$$EJ(p', S') = [pq + (1 - p)r]J(p'_G, S') + [1 - [pq + (1 - p)r]]J(p'_B, S'). \quad (5.7)$$

Using equations 5.6 and 5.7 we can remove the expectations operator from the DPE.

Digression: anticipation Note that if the decision maker either does not learn, or does not anticipate learning, the problem becomes much simpler. In that case, we have a single state variable, S , and we treat p as a parameter. Now the DPE is

$$J(S) = \max_c \{U(c) - (E\Delta)d(S) + \beta J(S')\}$$

subject to equations 5.1, (5.8)

where equation 5.6 gives the formula for $E\Delta$. If the decision maker really never learns, equation 5.8 is the DPE in every period.

Perhaps the regulator learns about Δ but simply does not anticipate learning. In that case, he still solves the DPE 5.8, but (to the decision maker's surprise) he has a different value of p in each period, as he acquires information. Our point here is that it is not simply the fact of learning, but the anticipation of learning that determines the nature of the problem that the decision maker solves.

Digression: a different state variable Suppose that the problem is non-autonomous, so we have some reason to keep track of calendar time. For example, there may be a finite horizon, or maybe there is some exogenous (e.g. technical) change that we need to keep track of. In this situation, we have to add the argument t to the value function. Set the initial time equal to 0 so t gives both calendar time and the number of times we have observed the realization of Δ .

In this case, an alternative to using p as the state variable is to use n = number of times we observed the realization $\Delta = G$. If we began with subjective probability p_0 that $x = x_1$ and n times we observed $\Delta = G$, and $t - n$ times we observed $\Delta = B$, then we can use Bayes' Rule to calculate my current subjective probability p_t .

In other words, it does not matter whether we treat the arguments of the value function as S, p, t or S, n, t . In both cases we need to keep track of three arguments. (See the reading by Clarke and Mangel for an example of the latter approach.)

However, consider the following minor change in the problem. Suppose that there are 3 possible realizations of Δ , G, B, U ; as before, there are only two possible distributions – x takes one of only two values. If we treat p as the state variable, the addition of a third outcome to Δ means that we need one more equation for p' (like equations 5.4 and 5.5), so we need to do a few more calculations when to compute the expectations. However, we do not need to increase the number of state variables. We will see that the numerical complexity depends on the number of state variables. In contrast, if we decide to keep track of outcomes, then we need to keep track of n , the number of times we observed G , and m , the number of times I observed B (so that I can calculate $t - n - m$, the number of times I observed U). In this case, the minor change increases the dimension of the state variable – something that we would like to avoid in order to reduce the difficulty of solving the problem.

The point here is that often there are different ways to define the state variable. We want as parsimonious a representation as possible.

5.3.2 The SP problem

The SP problem requires keeping track of "histories". A history is a sequence of outcomes. Suppose the initial period is time $t = 0$. By definition, there are no histories at $t = 0$. At $t = 1$ there are two possible histories $\{G\}$ and $\{B\}$; at $t = 2$ there are 4 possible histories, $\{G, G\}$, $\{G, B\}$, $\{B, G\}$, and $\{B, B\}$. Define a particular history at time t as h_t .

At the initial time we can calculate the subjective probabilities of each of these histories occurring in the future. For example, at time 0 the subjective probability of history $\{G\}$ is simply p_0 , the initial subjective probability. The subjective probability of $\{G, G\}$ is $p_0 P(G, p_0)$, where we use equation 5.4 to calculate $P(G, p_0)$. Using calculations of this sort we can determine the subjective probability of any history occurring in the future, conditional on current information. Define this probability as $\mu(h_t, t)$. For each time t and history h_t we can also calculate our subjective probability that $\Delta = G$; denote this probability as $\pi(h_t, t)$. For example, we can calculate our posterior probability that $x = x_1$ after history h_t ; denote this posterior as $p_t = P(h_t, p_0)$ to reflect the fact that the posterior depends on both our initial prior p_0 and the history h_t . With this notation

$$\pi(h_t, t) = p_t q + (1 - p_t) r.$$

The SP is to maximize the expectation, over "future histories", of the present value of the sequence of future payoffs. The SP problem is

$$\max_{\{c(h_t, t)\}} \sum_{t=0}^{\infty} \beta^t \left[\sum_{h_t} \mu(h_t, t) \left(U(c) - \underline{[\pi(h_t, t) G + (1 - \pi(h_t, t) B) d(S_t)]} \right) \right]$$

subject to equation 5.1.

The underlined term gives expected damages in each period conditioned on the history. Note that the optimization chooses a level of emissions for each possible history, for each time.

In most cases we are really interested in behavior only during the next few periods (each of which might consist of a decade). For most problems, the decision rules are insensitive to the choice of final horizon, T , provided that T is sufficiently big; with a discount factor $\beta < 1$, the very distant future does not matter much. Therefore, there is little cost to realism in replacing

the infinite horizon with a finite horizon, so that we have a finite number of possible histories, and thus a finite number of choice variables $c(h_t, t)$. In this case, the SP problem can be solved as a non-linear programming problem, e.g. by using GAMS. Computational constraints may still limit the number of possible histories.

The advantage of SP over DP is that the former allows the researcher to solve a dynamic problem using a program like GAMS, whereas the DP generally requires some programming. As a practical matter, the DP approach can handle larger problems (more histories).

5.3.3 Digression: active and passive learning again

In the models above, the decision maker does not manipulate the state variable in order to learn; in that sense the learning process is exogenous, i.e. learning is passive. However, this model can be modified to introduce a different kind of active learning, one that may be more relevant to environmental problems.

For example, in the Star Information Structure, the parameter λ measures the speed of learning. It might be possible to increase λ by paying a cost. By solving the problem for different values of λ we can see how λ affects the value of the payoff, and in this way determine how much we should spend to increase λ .

In the model with explicit distributions, learning can occur more rapidly if we obtain better signals. For example, the outcome G might include "very good" (VG) and "etzi-ketzi" (EK). We can compare the value of the program when we observe only G and B with the value when we observe VG , EK , and B . We can compare the increase in the value of the program, due to the more precise signals (and faster learning) with the cost of obtaining more precise signals.

The point of these examples is that we can use models of passive learning to think about certain kinds of activities that change the speed of learning.

5.4 Continuous distribution

The model with discrete distributions is attractive because it permits a very general representation of the different possibilities. The disadvantage of

that model is that it is not parsimonious. As noted above, if we have n possible distributions, we need $n - 1$ state variables to represent our current information; each of these state variables is our subjective probability that one of the possible distributions is the correct distribution.

An alternative uses continuous distributions. This alternative is more flexible in that it allows infinitely many rather than several possibilities. However, it restricts each of these infinitely many possibilities to be a member of a particular family that the modeler chooses. In addition, this modeling approach requires that our prior distribution and likelihood function are *conjugates* – meaning that the posterior distribution has the same form as the prior.

5.4.1 Example 1: Poisson and gamma distributions

This example uses the climate change problem. We choose the distributions to illustrate the methods, not because they necessarily provide a good representation of the climate change problem.

Suppose, as above, that the single period payoff is

$$U(c) - \Delta d(S),$$

where c is emissions, S is the stock of GHGs, U and d are known functions, and Δ is a random variable. In each period we observe actual damages $\Delta d(S)$ and we know the function d , so we can calculate the realization of Δ in that period.

For the purpose of obtaining a tractable model, we assume that Δ has a Poisson distribution, with parameter λ :

$$\Pr(\Delta = s \mid \lambda) = \frac{e^{-\lambda} \lambda^s}{s!},$$

where s is an element of the set of positive integers. With this distribution,

$$E\Delta = \text{var}(\Delta) = \lambda.$$

We model uncertainty by assuming that we do not know the true value of λ . For example, if λ is a very small number, then the expected damage of S is small; but if λ is large, expected damages are large. We model uncertainty

about λ by choosing the prior distribution to be gamma with parameters r and t :

$$\Pr(\tilde{\lambda} = \lambda \mid r, t) = \frac{e^{-\lambda t} (\lambda t)^{r-1} t}{(r-1)!}.$$

More precisely, this is a special form of the gamma distribution, called the Erlang distribution. Using this density,

$$E\tilde{\lambda} = \frac{r}{t} \quad \text{and} \quad \text{var}(\tilde{\lambda}) = \frac{r}{t^2}.$$

The choice of initial values of r and t provides considerable flexibility in modeling our uncertainty about λ . Suppose that the current values of the parameters are r, t and in this period we observe a particular value of Δ ; we noted above that we can calculate this realization because we observe actual damages and know the function d and the value of S . Due to the fact that the gamma and Poisson are conjugates, our posterior on λ , after observing Δ , is also a gamma with parameters

$$\begin{aligned} r' &= r + \Delta \\ t' &= t + 1. \end{aligned} \tag{5.9}$$

Note that the evolution of the parameter r is stochastic, because it depends on the realization of Δ , a random variable. However, the evolution of t is deterministic. This fact is useful, because it means that we can include non-stationarity (e.g. technical change) in the problem without increasing the dimension of the state variable.

This problem has three state variables, S, t , and r . The expectation of the single period payoff, given r, t , and S is

$$U(c) - d(S)E_{\lambda} (E_{\Delta|\lambda}) \Delta = U(c) - d(S)\frac{r}{t}.$$

The fact that Δ enters the problem linearly leads to a particularly simple expression for the expected single period payoff. If Δ had entered the payoff non-linearly, the expression for the expected single period payoff would have required more calculation, but that is really a fairly minor issue, since we need to use numerical methods to solve this problem in any case.

The DPE is

$$J(S, r, t) = \max_c \left[U(c) - d(S)\frac{r}{t} + \beta E_{\Delta} J(S', r + \Delta, t + 1) \right]$$

subject to equation 5.1.

The formula for the last expectation is

$$E_{\Delta} J(S', r + \Delta, t + 1) \\ \int_0^{\infty} \left(\sum_{s=0}^{s=\infty} \frac{e^{-\lambda} \lambda^s}{s!} \right) \left(\frac{e^{-\lambda t} (\lambda t)^{r-1} t}{(r-1)!} \right) J(S', r + s, t + 1) d\lambda.$$

The numerical solution to this problem is complicated by the fact that r and t increase without bound. To solve it numerically one would have to truncate the state space. For example, suppose that learning occurs only a finite number of times; in this case, t is finite. Rather than letting r be any positive integer, approximate the Poisson distribution by assuming that it takes one of a finite number of values. These numerical issues are important, but not the focus of this chapter.

Once again, it is worth pointing out that a fundamental feature of this problem is the *anticipation* of learning, not simply the fact of learning. If we did not anticipate learning, then we treat $E\Delta$ as a parameter, rather than a function of the state variable. Without anticipated learning, the state variable is simply S and the DPE is

$$J(S) = \max_c [U(e) - d(S)E_{\Delta} + \beta J(S')] \\ \text{subject to equation 5.1.}$$

The fact that we anticipate learning greatly complicates the problem, because it means that we have to keep track of a larger state variable, one consisting of three elements instead of a single element.

5.4.2 Example 2: Linear quadratic functions and normal distributions

This example makes three points. First, we demonstrate how it is possible to make some headway towards an analytic solution using the LQ functional form. Second, we explain why the structure of the problem eliminates active learning. Third, we show how to calibrate a model if in addition to the LQ functions one additionally assumes normal likelihood and priors.

The LQ functional form Here we extend the LQ model considered in Chapter 2, by allowing the damage parameter to be uncertain and assuming that we anticipate learning about this parameter. We briefly explain how the model can be further extended to include asymmetric information about costs, as in Chapter 2. This extension shows how anticipated learning about the damage parameter affects the ranking of taxes and quotas. For the time being, we ignore the problem of asymmetric information.

In the absence of a cost shock (thus removing the possibility of asymmetric information) the benefit of emitting at rate x is

$$f + ax_t - \frac{b}{2}x_t^2.$$

The stock grows according to

$$S_{t+1} = \delta S_t + x_t$$

where $0 < \delta < 1$ is a known parameter. [The notation in this chapter is not consistent with the notation in Chapter 2 – something to be fixed later.]

The damages associated with a stock of S_t are

$$D(S_t, \omega_t; G^*) = \frac{G^*}{2} (S_t - \bar{S})^2 \omega_t.$$

Here $G^* > 0$ is a fixed but unknown number and ω_t is the realization of a random variable with known distribution. In each period we observe damages and we know the stock S_t and the parameter \bar{S} . This information enables us to calculate $G^* \omega_t$, but we are not able to infer the fixed value of G^* . For example, we cannot tell whether high damages were a result of a high realization of ω or a high value of G^* . However, after observing a succession of damages, we begin to get a better idea of the level of G^* .

In order to be able to use standard dynamic programming methods, we need to be able to describe the subjective distribution of G^* using a finite number of parameters. Those parameters are elements of the state vector. For simplicity only we assume here that the subjective distribution of G^* at time t is determined by two moments, the mean and variance, $\chi_t \equiv (G_t, \sigma_{G,t}^2)$. However, the reader can think of $\sigma_{G,t}^2$ as a vector of higher moments rather than a scalar (the variance). The regulator cannot predict his future subjective expectation, so his current subjective expectation is an unbiased estimator of its future value, i.e. $E_t G_{t+\tau} = G_t$ for $\tau \geq 0$.

The form of our prior and our likelihood function determine how we update the state χ . The prior and the likelihood function lead to a system of equations analogous to equation 5.9. At least one equation in this system must be stochastic, to reflect the fact that we do not know what information we will have in the future, so we do not know what our beliefs will be in the future.

For the time being, put aside the question of the exact updating system of equations for χ . Assume that there is some such a system. Write $\bar{\omega} = E_{\omega}\omega$. Since, by assumption, we know the distribution of ω , we treat $\bar{\omega}$ as a fixed parameter. The DPE is

$$J(S, \chi) = \max_x \left[f + ax - \frac{b}{2}x^2 - \left(E_{G^*|\chi} \frac{G^*}{2} \right) (S_t - \bar{S})^2 \bar{\omega} + \beta E_{\chi'|\chi} J(S', \chi') \right].$$

A nice feature of this problem is that the value function is linear-quadratic in S . That is, J has the form

$$J(S, \chi) = \xi(\chi) + \gamma(\chi)S + \frac{\rho(\chi)S^2}{2}.$$

In the LQ problem without learning, we saw that the value function also has this form, but in that case ξ , γ , and ρ are numbers. Here, with learning, they are functions of the information, χ . These functions can be calculated numerically.

However, we can obtain some insight even without performing this calculation. In particular, consider the problem with learning where our current subjective expectation of G^* is $G_t = G$, where G is just some number. Compare this to the problem in which we never expect to learn (perhaps because we are certain that we know the value, or because we are naive, or simply because we never do learn anything). In this problem without anticipated learning, our expectation of G^* is also G . Obviously, the two problems are identical except that we expect to change our mind about the value of G^* under anticipated learning, but not in the absence of such learning.

We can show analytically that for any value of G and S , the optimal level of emissions is higher when we anticipate learning, compared to when we expect never to learn. This result does not depend on the updating rule for information. Anticipated learning makes us less cautious because we

know that if we get bad news in the future (i.e., learn that G^* is probably greater than we originally thought) we can modify our actions in the future. This option to act on future information makes bad news "less bad". The result that anticipated learning increases emissions in this dynamic model is a generalization of the same result for the two-period model, discussed in Chapter 3.

If we modify this model by adding a cost shock that is private information to firms, then we have a generalization of the stock pollutant problem under asymmetric information, studied in Chapter 2. We can show that anticipated learning about the damage parameter favors taxes over quotas. This result is quite intuitive, in view of the earlier result that anticipated learning increases the optimal level of emissions. The increase in emissions under learning means that it is "as if" learning causes the damage parameter G^* to become smaller. We saw in Chapter 2 that a smaller value of the damage parameter favors the use of taxes over quotas.

This problem is also convenient for explaining why there is no possibility for active learning in this model. This feature is due to the fact that the uncertain parameter and the random term ω enter the problem multiplicatively – not to the linear quadratic structure of the problem. We write the "data" (or signal) at time t as

$$data_t = ((2D_t)/((S_t - S)\hat{A}^2)) = G^*\omega_t.$$

By Bayes's theorem, the posterior on G^* , $Pr(G^*|data_t)$, is proportional to the product of the likelihood function, $Pr(data_t|G^*)$, and the prior, $Pr(G^*)$. The numerical value of the data at t depends on G^* and ω_t , but not on S_t . A change in S_t causes an offsetting change in D_t , leaving unchanged the ratio $((2D_t)/((S_t - S)\hat{A}^2))$. Therefore, $Pr(data_t|G^*)$ is independent of S_t ; consequently, the posterior $Pr(G^*|data_t)$ is also independent of S_t . Changing S_t does not change the information (about G^*) that the regulator obtains from observing $G^*\omega_t$.

In order to make this point sharper, consider an alternative damage function $(G^*/2)(S_t - S)\hat{A}^2 + \omega_t$, where ω appears additively rather than multiplicatively. In that case, the data at time t still consists of the level of damages and the level of the stock, (D_t, S_t) ; but in this case, a larger value of $(S_t - S)^2$ causes G^* to explain a greater proportion of the variation in damages. The

possibility of active learning exists in this additive model, because by increasing the stock of pollution, the regulator learns more about the damage parameter.

The assumption of normality In order to calibrate a model for greenhouse gasses, we need an explicit learning rule. One alternative is to assume that the distribution of the damage shock is lognormal:

$$\omega_t \sim i.i.d. \text{ lognormal} \left(-\frac{\sigma_\omega^2}{2}, \sigma_\omega^2 \right). \quad (5.10)$$

We express the subjective moments in terms of $g \equiv \ln G$. The regulator begins in period t with normal priors on $g^* = \ln G^*$, with mean g_t and variance $\sigma_{g,t}^2$:

$$g^* \sim N(g_t, \sigma_{g,t}^2). \quad (5.11)$$

Given distribution (5.11), the subjective distribution of G^* is log-normal with

$$\begin{aligned} E_t G^* &\equiv G_t = \exp\left(g_t + \frac{1}{2}\sigma_{g,t}^2\right) \\ \sigma_{G,t}^2 &\equiv \text{var}_t(G^*) = \exp(2g_t + \sigma_{g,t}^2) (\exp(\sigma_{g,t}^2) - 1). \end{aligned} \quad (5.12)$$

Since damages are a product of independent log-normally distributed variables, the regulator has log-normal priors on damages. After observing damages and the current stock, the Bayesian regulator updates his belief about g^* . The moment estimator of g^* , denoted \hat{g}_t , is

$$\hat{g}_t = \ln \frac{2D_t}{(S_t - \bar{S})^2} + \frac{\sigma_\omega^2}{2} \quad (5.13)$$

with variance $\sigma_g^2 = \sigma_\omega^2$. The posterior for g^* is normally distributed with the posterior mean g_{t+1} and posterior variance $\sigma_{g,t+1}^2$:

$$g_{t+1} = \frac{\sigma_\omega^2}{\sigma_\omega^2 + \sigma_{g,t}^2} g_t + \frac{\sigma_{g,t}^2}{\sigma_\omega^2 + \sigma_{g,t}^2} \hat{g}_t, \quad (5.14)$$

$$\sigma_{g,t+1}^2 = \frac{\sigma_{g,t}^2 \sigma_\omega^2}{\sigma_\omega^2 + \sigma_{g,t}^2} \Rightarrow \sigma_{g,t}^2 = \frac{\sigma_{g,0}^2 \sigma_\omega^2}{\sigma_\omega^2 + t \sigma_{g,0}^2}, \quad (5.15)$$

where $\sigma_{g,0}^2$ is the prior at the beginning of the initial period, $t = 0$. (These formulae are available in Greene's text on Econometrics, pages 407-410 or in Wikipedia.)

A smaller value of σ_{ω}^2 is equivalent to greater precision of future information. Using equation (5.14), greater precision of information implies that this period's posterior mean, g_{t+1} , is more responsive to information obtained in the current period. Using equation (5.15), greater precision of information means that the posterior variance decreases over time more rapidly. Thus, greater precision of information increases the speed and amount of learning.

The subjective distribution for the unknown damage parameter G^* collapses to the true value of this parameter as the number of observations approaches infinity. Appendix B1, available through *JEEM*'s online archive for supplementary material at <http://www.aere.org/journal/index/html>, proves this result.

If the regulator begins with too optimistic a prior ($g_0 < g^*$), g_t increases over time, in expectation. This increase can be enough to offset the decrease in $\sigma_{g,t}^2$, leading to an increase in $var_t(G_t)$ (using equation (5.12)). In this case, during a phase of the learning process the regulator becomes less certain about the value of G^* , although he eventually learns the correct value with probability 1. It is also straightforward to show that the regulator's current expectation of G^* is an unbiased estimate of the future expectation: $E_t G_{t+\tau} = G_t, \forall \tau \geq 0$.

[Add material explaining how this model can be calibrated.]

In the absence of anticipated learning, the regulator solves the control problem treating G_t as a constant. In this case the constant $\bar{G} \equiv G_t = \exp\left(g_t + \frac{\sigma_{g,t}^2}{2}\right)$ is the certainty equivalent value of G^* .

5.4.3 The role of conjugates

Both of these examples rely on "conjugate priors", i.e. the prior and the likelihood function are chosen so that the posterior has the same form as the prior. This choice greatly simplifies the problem because it makes it possible to write, in closed form, the equations of motion for the parameters that define current beliefs.

In principle at least it is fairly straightforward to drop the assumption of conjugates, at the cost of some additional work on numerical approximation.

5.4.4 Criticism of these examples

These examples show how to model anticipated learning about climate change. Arguably, the examples miss an important feature of the real world problem, because they do not take into account abrupt and irreversible changes. Later in this course we will discuss two kinds of irreversibilities, in the sections on sudden, catastrophic events, and in the section on nonconvex problems. Both of these modeling frameworks – particularly the nonconvex environment – offer the possibility of studying anticipated learning in a much more interesting and realistic setting.

6 Numerical Methods

There are three principal approaches to studying dynamic problems. First, models with certain functions admit closed form solutions; the log utility paired with Cobb-Douglas production in Chapter xx and the linear-quadratic model in Chapter xx are prominent examples. Second, some models, especially those with a single state variable, are amenable to qualitative graphical analysis by means of a phase portrait, as we will discuss in Chapter xx. The third approach involves numerical methods and has become increasingly important during the last fifteen years. With ever more powerful computing and improved algorithms, numerical approaches are likely to become an essential tool for economists working on dynamic problems. This chapter introduces some of the ideas used to numerically solve dynamic problems. The successful application of numerical methods requires art (or experience) as well as science, and cannot be reduced to a series of menus. However, understanding the basic menus is a first step in using numerical methods. We focus on what we consider the most interesting case: a continuous control variable and a continuous state space. At the end of the chapter we discuss an application to the integrated assessment of climate change.

6.1 An abstract algorithm: function iteration

We begin with a one-state variable deterministic problem:

$$V^*(x, t) = \max_{\{a\}_{\tau=t}^T} \sum_{\tau=t}^T \beta^{\tau-t} U(a_\tau, x_\tau)$$

subject to $x_{\tau+1} = g(a_\tau, x_\tau)$, with $x_t = \bar{x}$ given. (6.1)

In this problem, time is not an argument of the single period payoff function, U , or the growth function, g , so the problem is autonomous if $T = \infty$. Given that it is deterministic, and assuming $T < \infty$, we can solve the problem using non-linear programming methods. Non-linear programming is impractical in stochastic settings, except for problems where there are a small number of possible realizations of a random variable, and thus a small number of decision trees. Dynamic programming is especially useful in complex stochastic

settings, but we begin with the deterministic problem in order to present the basic ideas simply. We discuss higher dimensional and stochastic problems below.

With t denoting calendar time, and defining s as time-to-go before the end of the problem, $s = T - t$, we write the value function as $V^s(x) = V^*(x, t)$. The dynamic programming equation, for $s \geq 1$, is

$$V^s(x) = \max_a \{U(a, x) + \beta V^{s-1}(g(a, x))\}, \quad (6.2)$$

with the boundary condition

$$V^0(x) = \max_a U(a, x). \quad (6.3)$$

Here the superscript denotes the number of decisions remaining after the current decision.

We discuss the algorithm known as *value function iteration*. The algorithm begins by finding the solution to the problem on the right side of equation (6.3) to obtain the function $V^0(x)$. Substituting that function into equation (6.2) for $s = 1$, we then solve the resulting problem to obtain $V^1(x)$. Proceeding iteratively, we solve the T one-stage problems. At each stage, s , we obtain two functions: the decision rule, denoted $a^s(x) = \arg \max U(a, x) + \beta V^{s-1}(g(a, x))$, and the value function $V^s(x)$. We use the value function for the “backward sweep”, increasing s , approaching the initial time period. We use the decision rule for the “forward sweep”. Given the initial condition, the value of $x_t = \bar{x}$, we can find the trajectory of the optimally controlled state variable. In the stochastic setting we can use the control rule for Monte Carlo simulations. In general, we can neither calculate nor store exact solutions for the value function or the control rule. We therefore approximate the value function in each step. Given the value function, we obtain the control rule from a quasi-static optimization problem. Hence, we will sometimes decide to only approximate and store the value function, and not the control rule. In order to approximate the value function we will have to decide on the intervals over which we approximate the function and the approximation method.

An abstract summary of the *value function iteration algorithm* is

I *Initialization:*

1. Choose how to approximate the value function. Usually this step involves the choice of:
 - basis functions to approximate the value function,
 - interval of the state space on which to approximate the value function,
 - interpolation nodes at which you evaluate the optimization problem.
 See section 6.2 for details.
2. For $s = 0$ in the case where
 - T finite: maximize the right hand side of equation (6.3) at the interpolation nodes. Proceed to step 4.
 - $T = \infty$: pick an initial guess for the approximate value function V^0 . Set $s = 1$ and proceed to step 3.

II *Iteration:*

3. Maximize the right hand side of the Bellman equation (6.2) for s .
4. Approximate the solution of the maximization step. Usually this step involves solving for the coefficients of the basis vectors. You obtain the value function V^s .
5. Increment time to go s by 1 and repeat steps 3 and 4 until for
 - T finite: you obtain the value function in the present V^T .
 - $T = \infty$: a break criterion is satisfied up to a given tolerance, usually related to the change of the value function or the basis coefficients from one iteration to the next.

III *Simulation:*

6. Simulate the system dynamics by solving the Bellman equation (6.2) iteratively forward in time. Starting with the initial condition $x_t = \bar{x}$, the simulation solves a sequence of quasi-static optimization problems, given the (approximate) value functions at every point in time. If we fitted the policy functions in the earlier steps, we can use these directly to simulate the system dynamics. In the stochastic case, we can simulate using expected draws as a proxy, and then also simulate large sets of truly random paths and determine distributional properties.

In the case where T is finite, we have to store all of the value function V^0, \dots, V^T . In the case where $T = \infty$, we have an autonomous problem and

we usually store only the current and the previous value function. We keep track of the last iteration's value function so that the break criterion can evaluate changes from one iteration to the next. In the case of an infinite time horizon, s is not the time to go, but simply counts the iterations starting from our initial guess.

In the one dimensional deterministic setting, simple numerical optimization algorithms solve the *static* optimization problem in step 3 (right side of equation 6.2), conditional on a given value of x and the function V^{s-1} . In a higher dimensional stochastic setting, this static optimization problem in itself may require sophisticated methods. We do not discuss those methods, merely assuming that each of the static problems can be solved, given x and V^{s-1} . In application, we would usually rely on a freely distributed or commercial solver for the maximization step.

6.2 Approximating a function

Even if we knew the function V^{s-1} , it would rarely be the case that we could find a closed form solution for function V^s . Assume we cannot find a closed form solution but want to solve the problem exactly (up to numerical precision). Then, we would have to solve the maximization problem for every $x \in X$ and store the optimization result, i.e. the value of the function V^s , for every point. Given $X \subset \mathbb{R}$ is uncountable, this procedure is infeasible. As you might know from algebra, we can approximate continuous functions on a compact support $X \subset \mathbb{R}$ arbitrarily closely by a (countable) sequence of polynomials. In general, there are different countable basis of function spaces. Let us denote a sequence of orthonormal¹⁵ basis functions by $\phi_0, \phi_1, \phi_2, \dots$. Then, every function f in the corresponding space can be written as

$$f = \sum_{i=0}^{\infty} c_i^* \phi_i$$

¹⁵An orthonormal basis satisfies $\int_X \phi_i(x)\phi_j(x)dx = \delta_{i,j}$, where $\delta_{i,j}$ is the Kronecker symbol: unity if $i = j$ and zero otherwise. It is a set of properly normalized orthogonal functions. The integral over the two functions (the inner product) can also carry a weighting function.

with coefficients in \mathbb{R} satisfying¹⁶

$$c_i^* = \int_X \phi_i(x) f(x) dx . \quad (6.4)$$

This simple formula for the coefficients relies on the orthonormality of the basis functions. A countably infinite series is still too much to keep track of. We therefore resort to a finite subset of the basis. In addition, we have to find an efficient way of dealing with the integration in equation (6.4). A simple but effective approach replaces the integral by a sum, evaluating both functions only at a finite set of points, the so-called interpolation nodes x_1, \dots, x_J . Then, we obtain the approximate formula

$$f \approx \sum_{i=0}^N c_i \phi_i \quad \text{with} \quad c_i = \sum_{j=1}^J \phi_i(x_j) f(x_j) = \Phi_i' \mathbf{f} , \quad (6.5)$$

where $\Phi_i \equiv (\phi_i(x_1), \dots, \phi_i(x_J))$ and $\mathbf{f} \equiv (f(x_1), \dots, f(x_J))$ are (column) vectors of numbers: they contain the value of the functions at the interpolation nodes. The equation for the coefficients c_i replaces the integral over a compact interval with a dot product of two vectors. While the integral is the inner product on the function space, the dot product is the inner product on our approximating vector space. However, for arbitrary choices of the interpolation nodes x_1, \dots, x_J , the vectors $\Phi_0, \Phi_1, \dots, \Phi_N$ will no longer be orthogonal. Section 6.4 discusses a particular set of basis functions (Chebychev polynomials) and the corresponding set of nodes (Chebychev nodes) for which the vectors $\Phi_0, \Phi_1, \dots, \Phi_N$ remain orthogonal. If they remain also orthonormal, we obtain the vector of basis coefficients as

$$\mathbf{c} = \Phi' \mathbf{f} , \quad (6.6)$$

where $\Phi = (\Phi_0, \Phi_1, \dots, \Phi_N)$. If the basis vectors are only orthogonal, but not normalized to unity, the correct formula for the coefficients will also include a weight.

¹⁶The * on the coefficients distinguishes the exact coefficients from the coefficients c_i that we will use in our numeric approximation below. More generally, we have to assure that the function we want to approximate lies in some Hilbert space (a Banach space with an inner product), and the integration is the inner product. In particular, the integration can involve an additional weighting function.

We pause to emphasize the relation between the functions ϕ_i and the vectors Φ_i . In the value function approximation, we fit the value function V^s to the solution of the maximization problem on the right hand side of the Bellman equation (6.2). Here, the vector \mathbf{f} corresponds to the solution of the Bellman equation at the interpolation nodes. We use the vectors Φ_i to find the coefficients of the basis functions ϕ_i , given only the finite vector of values at the interpolation nodes (of \mathbf{f} or the maximization problem). We use the functions ϕ_i whenever we need to evaluate the approximated function between different interpolation nodes.

We now discuss the value function approximation in more detail. We also show a parallel between the function fitting procedure and an OLS estimator. We begin by choosing a sequence of J points, our interpolation nodes, x_0, x_1, \dots, x_{J-1} on an interval $[a, b]$; $x_0 = a$ and $x_{J-1} = b$. The subscript here identifies the node; it is not a time index. We have, or we generate, the value of the function that we seek to approximate, at each of these nodes. We also choose N basis functions $\phi_i(x)$, $i = 0, 2, \dots, N - 1$. We approximate the function of interest as a linear combination of these basis functions, evaluated at the nodes. Figure ?? illustrates the procedure. The graph of the function that we wish to approximate is the solid curve, shown as $f(x)$. For the time being, we take as known the value of this function, $f(x_j)$, at each node $j = 0, \dots, J - 1$; the heavy dots represent those values. The objective is to fit a curve through those values. One such curve is shown as $\hat{f}(x)$.

Given the values of $f(x_j)$ and the basis functions $\phi_i(x)$, the curve-fitting problem is to minimize the distance between the estimated values, $\hat{f}(x_j)$, and the observed values, $f(x_j)$. Define

$$\begin{pmatrix} \hat{f}(x_0) \\ \hat{f}(x_1) \\ \vdots \\ \hat{f}(x_{J-2}) \\ \hat{f}(x_{J-1}) \end{pmatrix} = \begin{bmatrix} \phi_0(x_0) & \phi_1(x_0) & \cdots & \phi_{N-2}(x_0) & \phi_{N-1}(x_0) \\ \phi_0(x_1) & \phi_1(x_1) & \cdots & \phi_{N-2}(x_1) & \phi_{N-1}(x_1) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \phi_0(x_{J-2}) & \phi_1(x_{J-2}) & \cdots & \phi_{N-2}(x_{J-2}) & \phi_{N-1}(x_{J-2}) \\ \phi_0(x_{J-1}) & \phi_1(x_{J-1}) & \cdots & \phi_{N-1}(x_{J-1}) & \phi_{N-1}(x_{J-1}) \end{bmatrix} \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_{N-2} \\ c_{N-1} \end{pmatrix},$$

or using matrix notation

$$\hat{\mathbf{f}} = \Phi \mathbf{c}$$

where $\hat{\mathbf{f}}$ is $J \times 1$ vector of estimated values of f , evaluated at the J nodes; Φ is an $J \times N$ “interpolation matrix”; and \mathbf{c} is the $N \times 1$ vector of “basis

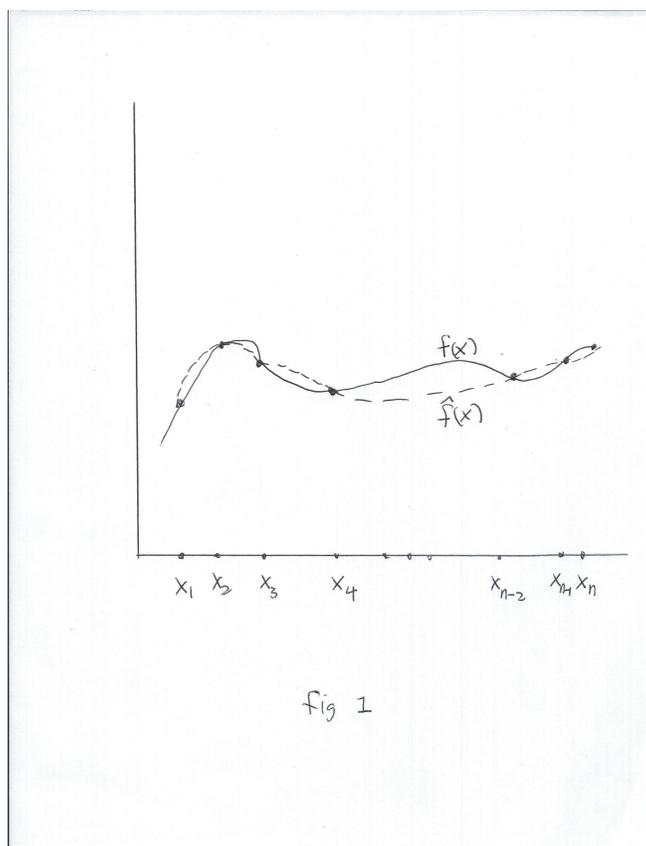


Figure 2 shows sketches the interpolation of a function f evaluated at nodes x_1 to x_n by a function \hat{f} that is spanned by a finite set of basis functions.

coefficients". Define \mathbf{f} as the $J \times 1$ vector with j 'th element the known value $f(x_j)$. For $J \geq N$, the vector \mathbf{c} that minimizes the Euclidean distance between $\hat{\mathbf{f}}$ and \mathbf{f} is the familiar Ordinary Least Squares estimator

$$\mathbf{c} = (\Phi' \Phi)^{-1} \Phi' \mathbf{f}.$$

The inverse exists because of the assumption that the basis functions ϕ_i are linearly independent, and $J \geq N$. Practitioners often choose $J = N$, so that $\mathbf{c} = \Phi^{-1} \mathbf{f}$. With $J = N$, $\hat{\mathbf{f}} = \mathbf{f}$, so the approximation equals the true value of the function at the nodes; in general, the approximation does not equal the true value of the function for a value of x other than a node.

Denote $\phi(x)$ as the N dimensional row vector of functions with i 'th element

$\phi_i(x)$. Note again the important distinction: $\phi(x)$ is a vector of functions, whereas the i 'th column of Φ is a vector Φ_i of values $\phi_i(x_j)$ for $j = 0, \dots, J-1$. The approximation of the function $f(x)$ is $\hat{f}(x) = \phi(x) \mathbf{c}$.

6.3 Approximation and value function iteration

For the application at hand, consider the problem once the nodes and the basis functions have been chosen. Here V^s , rather than f , is the function that we want to approximate. At $s = 0$ we solve the problem on the right side of equation (6.3) for each of the nodes, resulting in the J values (not functions) $V^0(x_i)$, $i = 0, 2, \dots, J-1$, which we denote in vector form as \mathbf{V}^0 . We obtain the basis coefficients as above, yielding $\mathbf{c}^0 = (\Phi' \Phi)^{-1} \Phi' \mathbf{V}^0$. The superscript 0 identifies these as the basis coefficient at the 0 iteration, i.e. at the final stage of the problem. Our estimate of the function $V^0(x)$ is $\hat{V}^0(x) = \phi(x) \mathbf{c}^0$. We now proceed iteratively, at each stage replacing the true but unknown function V^{s-1} with its approximation $\hat{V}^{s-1}(x) = \phi(x) \mathbf{c}^{s-1}$. For example, at stage s , given \mathbf{c}^{s-1} , we solve

$$\mathbf{V}_j^s \equiv \max_a \{U(a, x_j) + \beta \phi(g(a, x_j)) \mathbf{c}^{s-1}\}$$

for each each of the J nodes x_j , $j = 0, 2, \dots, J-1$. At stage $s = 0$ we can claim that we know the values of $V^0(x_j)$, subject to the limits of numerical accuracy. At stages $s > 0$, matters are slightly different. The state $s > 0$ problem is conditioned on the estimate, rather than the true value, of the $s-1$ value function. Therefore, we have only estimates $V^s(x_j)$, not the actual value $V^s(x_j)$. We write this estimate for the values at the nodes as the vector \mathbf{V}^s . The stage s basis coefficients are $\mathbf{c}^s = (\Phi' \Phi)^{-1} \Phi' \mathbf{V}^s$. Our approximation of the value function at stage s is $\hat{V}^s(x) = \phi(x) \mathbf{c}^s$; if $J = N$, then $\hat{V}^s(x_j) = \mathbf{V}_j^s$.

By using approximations, we replace the difficult problem, at each stage, of finding a function, with the considerably simpler problem of finding a vector of coefficients. Rather than having to store functions in memory, we only have to store vectors of coefficients. Note also that interpolation matrix, Φ , does not vary over stages. That matrix depends only on our choice of basis functions and of nodes, so we need to compute the matrix $(\Phi' \Phi)^{-1} \Phi'$ only once. Moreover, in the case where the basis vectors Φ_i are orthogonal, the

matrix $\Phi'\Phi$ is a diagonal matrix and we can do without matrix inversion. If the basis vectors are orthonormal, we are back to equation (6.6).

At each stage we obtain the J values

$$a_j^s = \arg \max \{U(a, x_j) + \beta \phi(g(a, x_j)) \mathbf{c}^{s-1}\}.$$

It is important to note that, even if the initial condition for x equals a node, optimal behavior likely causes the next-period state variable to lie between nodes. Therefore, we need the function approximation $\hat{V}^s(x) = \phi(x) \mathbf{c}^s$ rather than just the vector \mathbf{V}_j^s . Moreover, for implementing the forward sweep we apply the sequence of optimal control rules. The value of the state variable in a given period will generally not coincide with a node. Therefore, we need the optimal control off the grid points. We have two options. First, we estimate the control rules using the same procedure as above. At stage s , having calculated the J values of the optimal control, each corresponding to a node, we have the vector \mathbf{a}^s . If we choose to use the same interpolation matrix (as is often done), we approximate the stage s optimal control rule as $\phi(x) \mathbf{d}^s$, where $\mathbf{d}^s = (\Phi'\Phi)^{-1} \Phi' \mathbf{a}^s$. Second, we can simply rely on the stored value functions in the forward sweep and re-solve the optimization problem for the (single) exact value of the state in that period. In either of the two cases, we will generally pass information about the optimizing controls \mathbf{a}^s as initial conditions into the optimization problem of the next iteration $s + 1$.

6.4 Choosing a basis functions and interpolation nodes

A natural choice would be the “monomial basis” $\phi_i(x) = x^{i-1}$, $i = 0, \dots, N - 1$, and to space the nodes equally over the interval $[a, b]$. These choices lead to a poor approximation in many cases. The interpolation matrix Φ induced by the monomial basis is known as the Vandermonde matrix, a famously “ill-conditioned” matrix. This sensitivity to numerical mistakes is often measured by a so-called condition number.¹⁷ With $J = N$, the vector of basis coefficients \mathbf{c}^s is the solution to $\Phi \mathbf{c}^s = \tilde{\mathbf{V}}^s$. Because $\tilde{\mathbf{V}}^s$ is the result of approximations at previous stages, it contains “errors”, in the sense that

¹⁷For the linear system $Ay = b$, the condition number of the matrix A provides an upper bound of the elasticity of the norm of the solution, y , to the the norm of the data, b . For the Vandermonde matrix this condition number grows large with N , implying that a small error in b may lead to a large error in y .

it does not equal the true value of the program at a particular stage and particular vector of states. If Φ has a high condition number, as with the Vandermonde matrix, errors increase with iterations; the error in $\tilde{\mathbf{V}}^s$ leads to a larger “error” in \mathbf{c}^s , increasing the error in $\tilde{\mathbf{V}}^{s+1}$, and so on. In addition, the condition number of the Vandermonde matrix increases with N . It would seem that a higher degree polynomial approximation (increasing N) would increase the accuracy of the approximation. However, with the monomial basis, the higher condition number associated with larger N may cause the accuracy of the approximation to fall with N .

A more suitable polynomial bases is given by the Chebychev polynomials (of the “first kind”). It is convenient to change the units of the state variable, replacing x with

$$z = \frac{2(x-a)}{b-a} - 1,$$

so z ranges over $[-1, 1]$ as x ranges over $[a, b]$.¹⁸ The Chebyshev functions, Γ_i , $i = 0, 1, \dots, N-1$ are defined by¹⁹

$$\Gamma_i(z) = \cos[i \arccos(z)], \quad z \in [-1, 1].$$

We can generate these polynomials recursively by defining

$$\Gamma_0(z) = 1, \quad \Gamma_1(z) = z$$

and

$$\Gamma_i(z) = 2z\Gamma_{i-1}(z) - \Gamma_{i-2}(z).$$

For example, we find

$$\begin{aligned} \Gamma_2(z) &= 2z^2 - 1, & \Gamma_3(z) &= 4z^3 - 3z, \\ \Gamma_4(z) &= 8z^4 - 8z^2 + 1, & \Gamma_5(z) &= 15z^5 - 20z^3 + 5z. \end{aligned}$$

As these examples suggest, the even polynomials are symmetric around zero, while the uneven polynomials are antisymmetric. Figure 3 graphs the first five Chebychev polynomials. The polynomial Γ_i has i zeros which lie in the

¹⁸The so-called *shifted Chebychev polynomials* (of the first kind) are defined on $[0, 1]$ as $\Gamma_i^* = \Gamma_i(2z-1)$, $z \in [0, 1]$, where Γ_i are the Chebychev polynomials as we define them below.

¹⁹Observe the property $\Gamma_i(\cos \theta) = \cos(i\theta)$, $\theta \in [0, \Pi]$.

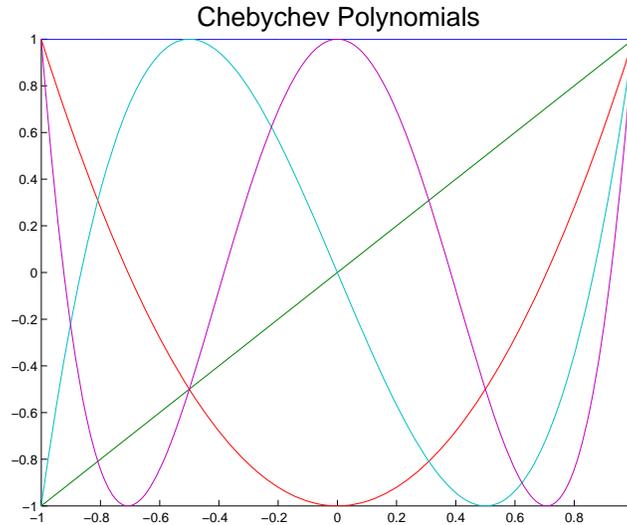


Figure 3 shows the first five Chebychev Polynomials $\Gamma_0, \dots, \Gamma_4$.

interval $(-1, 1)$. They are given by

$$z_j = \cos\left(\frac{j + \frac{1}{2}}{i} \pi\right), \quad j = 0, 1, \dots, i - 1. \quad (6.7)$$

It turns out that choosing these zeros as interpolations nodes minimizes the maximal distance between a function and its Chebychev approximation. Therefore, the zeros in equation (6.7) are called Chebychev nodes.

The left hand side of Figure 4 shows how badly Chebychev polynomials approximate functions when using equidistant interpolation nodes. We observe that the more basis functions and nodes we use, the further off is the approximation close to the bounds. Note that, for equidistant nodes, the basis functions $\Gamma_0, \Gamma_1, \dots, \Gamma_{N-1}$ are no longer orthogonal. In contrast, if we use the N zeros of the N th order polynomial given in equation (6.7) as interpolation nodes, we obtain an orthogonal set of basis functions and the approximations shown on the right hand side of Figure 4. These approximations seem to do particularly well when approximating a smooth function (top), and still somewhat alright when approximating a step function (bottom). The combination of Chebyshev nodes and Chebyshev polynomials produce a well-conditioned interpolation matrix Φ . This matrix is orthogonal, i.e. $\Phi' \Phi$ is diagonal,

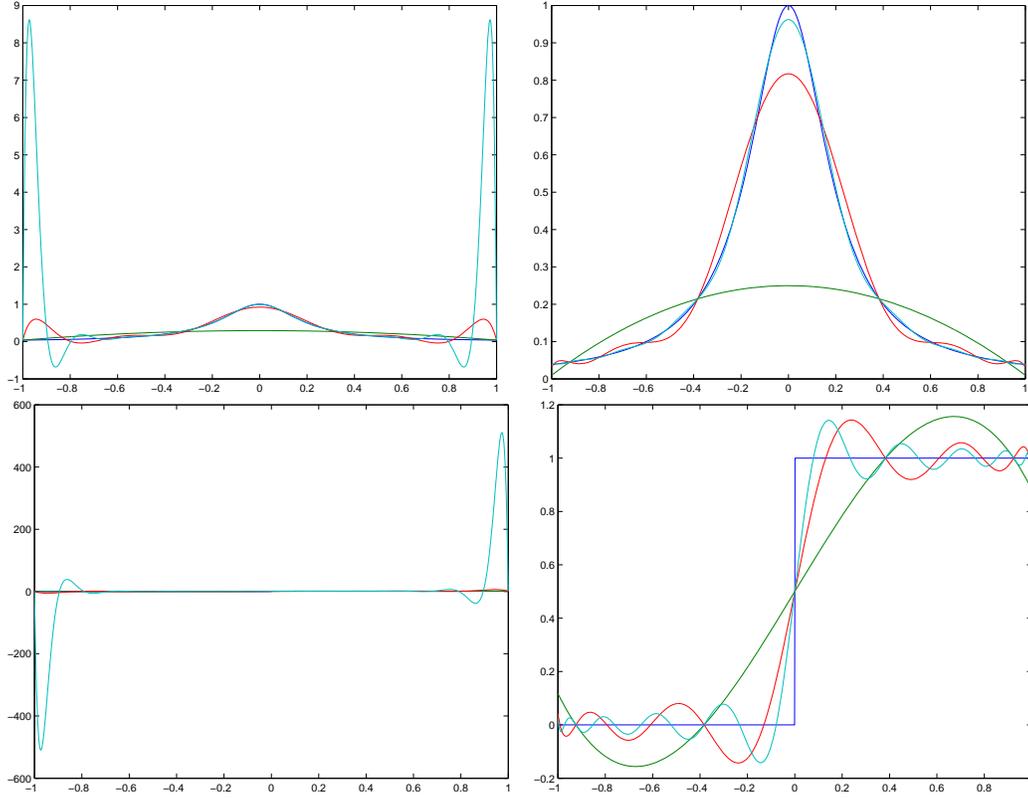


Figure 4 shows how the Chebyshev polynomials Γ_4 (dark green), Γ_{12} (orange), and Γ_{20} (cyan) approximate two different function (dark blue) for equidistant interpolation nodes (left) and Chebyshev interpolation nodes (right). In the upper row, the polynomial approximate the smooth Runge function $f(z) = \frac{1}{1+25z^2}$ (dark blue) and do better than when approximating the discontinuous Heaviside function in the lower row.

making it possible to accurately and cheaply compute $\mathbf{c}^s = (\Phi' \Phi)^{-1} \Phi' \tilde{\mathbf{V}}^s$. As we observed already in equations (6.5) and (6.6), the matrix inversion boils down do a simple matrix multiplication, and a normalization factor because Chebyshev polynomials are not normalized to unity.

We can employ the cosine representation of the Chebyshev polynomials (see observation in footnote 19) to calculate the matrix elements with $J = N$ as

$$\Phi_{ij} = \Phi_i(x_j) = \Gamma_{i-1}(z_j) = \cos\left(i \frac{j + \frac{1}{2}}{J} \pi\right) \quad i=0, \dots, N-1; j=1, \dots, J-1.$$

The diagonal matrix $\Phi' \Phi$ gives us the normalization weights in front of the

sum when calculation the coefficients²⁰

$$c_0 = \frac{1}{N} \sum_{j=1}^J \Phi_{ij} f(x_j) \quad \text{and} \quad c_i = \frac{2}{N} \sum_{j=1}^J \Phi_{ij} f(x_j).$$

6.5 Higher dimensional state space

The previous methods generalize to the multi-dimensional case. We use the basis functions and interpolation nodes discussed above in each dimension, and construct a tensor basis for the product space. Approximation techniques analogous to the one-dimensional case apply. Now we have to solve a multivariate maximization problem on the right hand side of the Bellman equation for every point on a multidimensional grid, which makes the problem computationally more expensive. We briefly discuss the case for two states.

Suppose that we want to approximate a function of two variables $f(x, y)$ where $x \in I_x \subset \mathbb{R}$ and $y \in I_y \subset \mathbb{R}$. We approximate the twodimensional function on the Cartesian product $I \equiv I_x \times I_y = \{(x, y) \mid x \in I_x \text{ and } y \in I_y\}$.²¹ We choose a set of univariate basis functions on each of the intervals, $\phi_{i_x}^x$, $i_x = 1, 2, \dots, N_x - 1$ and $\phi_{i_y}^y$, $i_y = 1, 2, \dots, N_y - 1$. We obtain a basis on the product space using the so-called tensor basis

$$\phi_{i_x, i_y}(x, y) = \phi_{i_x}^x(x) \otimes \phi_{i_y}^y(y) \text{ for all } i_x, i_y.$$

We can think of the tensor product as a bilinear multiplication of two vector space.²² Here, the multiplication is either that of two functions, i.e., Chebychev polynomials, or of values, i.e., Chebychev polynomials evaluated at particular values. The reason why we do not write the tensor product

²⁰Apart from the normalization factors, inserting the cosine back into the sum shows that we essentially face a discrete cosine transform of $f(\cos(\theta_j))$.

²¹In general, we might be more interested in particular combinations of the states than in others, e.g. low capital stock might not come with high CO₂ concentrations. However, for the described function fitting procedure we need a rectangular set. If we have to be parsimonious in the amount of nodes, it is worthwhile considering to approximate a set of normalized variables on the rectangular grid rather than the actual states.

²²Bilinearity is $a(\phi_{i_x}^x(x) \otimes \phi_{i_y}^y(y)) = (a\phi_{i_x}^x(x)) \otimes \phi_{i_y}^y(y) = \phi_{i_x}^x(x) \otimes (a\phi_{i_y}^y(y))$ for $r \in \mathbb{R}$, $(\phi_{i_x}^x(x) + \phi_{l_x}^x(x)) \otimes \phi_{i_y}^y(y) = \phi_{i_x}^x(x) \otimes \phi_{i_y}^y(y) + \phi_{l_x}^x(x) \otimes \phi_{i_y}^y(y)$, and the analogous relation for an addition of two Chebychev polynomials in the second space.

as a simple multiplication is because the two polynomials live on different spaces. However, for practical purposes we can think usually think of evaluating the tensor products point by point dealing simply with multiplication of numbers. We approximate the function f by the estimate $\hat{f}(x, y)$

$$\hat{f}(x, y) = \sum_{i_x, i_y} c_{i_x, i_y} \phi_{i_x}^x(x) \otimes \phi_{i_y}^y(y). \quad (6.8)$$

If we evaluate the function at a given point, then each summand is simply the product of the coefficient, the value of the basis function for x , and the value of the basis function for y . Using a particular tensor notation, equation (6.8) is sometimes written as

$$\hat{f}(x, y) = [\phi^x(x) \otimes \phi^y(y)] c, \quad (6.9)$$

where $\phi^x(x)$ and $\phi^y(y)$ are, respectively, N_y and N_x dimensional (row) vectors of basis functions, with elements $\phi_{i_x}^x$ and $\phi_{i_y}^y$; and c is a $N \equiv N_x N_y$ dimensional column vector of Chebychev coefficients.²³ Once we evaluate the Chebychev polynomials at a given set of nodes, equation (6.9) has a straight forward interpretation as the Kronecker product of two matrices (see below).

In order to estimate the coefficients, we evaluate the functions f and ϕ_{i_x, i_y} at a finite set of grippoints (x_{j_x}, y_{j_y}) , $j_x = 0, 1, \dots, J_x - 1$ and $j_y = 0, 1, \dots, J_y - 1$. If we use Chebychev functions for the basis, we also use Chebychev nodes on I_x and I_y to construct a rectangular grid on I . We denote $\Phi_{j_x, i_x}^x = \phi_{i_x}^x(x_{j_x})$ to obtain matrices Φ^x and Φ^y , interpolating the vectors of basis functions. We assume $N_x = J_x$ and $N_y = J_y$. Once we replace the basis functions by their values on the grid, we can write the expression in the square brackets in equation (6.9) using the Kronecker product. The Kronecker product of

²³Note that c is *not* of the dimension $N_y + N_x$. We do not simply apply one part of the c vector to the basis function in the x-space and the other part to the basis vectors in the y-space. General basis functions are combinations of basis functions in both dimensions; there is a coefficient in c for each such combination.

the matrices Φ^x and Φ^y is defined as²⁴

$$\begin{aligned} \Phi^x \otimes \Phi^y &= \begin{pmatrix} \Phi^x_{0,0} \Phi^y & \dots & \Phi^x_{0,N_y-1} \Phi^y \\ & \vdots & \vdots \\ \Phi^x_{N_x-1,0} \Phi^y & \dots & \Phi^x_{N_x-1,N_y-1} \Phi^y \end{pmatrix} \\ &= \begin{pmatrix} \Phi^x_{0,0} \Phi^y_{0,0} & \Phi^x_{0,0} \Phi^y_{0,1} & \dots & \Phi^x_{0,N_y-1} \Phi^y_{0,N_y-1} \\ \Phi^x_{0,0} \Phi^y_{1,0} & \Phi^x_{0,0} \Phi^y_{1,1} & \dots & \Phi^x_{0,N_y-1} \Phi^y_{1,N_y-1} \\ \vdots & \vdots & \vdots & \vdots \\ \Phi^x_{N_x-1,0} \Phi^y_{N_y-2,0} & \Phi^x_{N_x-1,0} \Phi^y_{N_y-2,1} & \dots & \Phi^x_{N_x-1,N_y-1} \Phi^y_{N_y-2,N_y-1} \\ \Phi^x_{N_x-1,0} \Phi^y_{N_y-1,0} & \Phi^x_{N_x-1,0} \Phi^y_{N_y-1,1} & \dots & \Phi^x_{N_x-1,N_y-1} \Phi^y_{N_y-1,N_y-1} \end{pmatrix} \end{aligned} \quad (6.10)$$

With this representation of the tensor product, equation (6.9) is a straight forward multiplication of an $N \times N$ matrix with an N dimensional column vector of basis coefficients. Check that this multiplication indeed reproduces the function values according to equation (6.8) at the gridpoints.²⁵

For an interpolation of f according to equation (6.8), we need to estimate the basis coefficients. Fortunately, we only have to invert the individual matrices in each dimension and not, e.g., matrix (6.10). We find the coefficients in the above tensor notation as

$$c = [\Phi^x^{-1} \otimes \Phi^y^{-1}] \mathbf{f}$$

or spelled out

$$c_{i_x, i_y} = w_{i_x, i_y} \sum_{j_x, j_y} f(x_{j_x}, y_{j_y}) \phi_{i_x}^x(x_{j_x}) \phi_{i_y}^y(y_{j_y}). \quad (6.11)$$

where w_{i_x, i_y} is the inverse of the norm of the corresponding basis vector. Let $w_{i_x}^x = 1 / \sum_{j_x} \Phi^x_{i_x, j_x} \Phi^x_{j_x, i_x} = 1 / (\Phi^x_{\cdot, i_x} \Phi^x_{i_x, \cdot})$ be the norm of the i_x th basis vector on I_x , and analogously for the basis vectors on I_z . Then $w_{i_x, i_y} =$

²⁴We use the same symbol for the tensor product and the Kronecker product. More precisely, the Kronecker product of the two matrices yields a particular representation of the tensor product. Similar to an outer product of two vectors producing a matrix, the Kronecker product produces an $N \times N$ dimensional tensor from the two original matrices.

²⁵The matrix representation (6.10) of the tensor product in equation (6.8) is very memory intensive for higher dimensional state spaces.

$w_{i_x}^x w_{i_y}^y$. For Chebychev polynomials we have

$$w_{i_x, i_y} = (2 - \delta_{i_x, 0})(2 - \delta_{i_y, 0}) / (N_x N_y) .$$

We can verify formula (6.11) for the Chebychev coefficients by checking

$$\begin{aligned} f(x_{j_x}, y_{j_y}) &= \sum_{i_x, i_y} c_{i_x, i_y} \phi_{i_x}^x(x_{j_x}) \phi_{i_y}^y(y_{j_y}) \\ &= \sum_{i_x, i_y} \left[w_{i_x, i_y} \sum_{l_x, l_y} f(x_{l_x}, y_{l_y}) \phi_{i_x}^x(x_{l_x}) \phi_{i_y}^y(y_{l_y}) \right] \phi_{i_x}^x(x_{j_x}) \phi_{i_y}^y(y_{j_y}) \\ &= \sum_{l_x, l_y} f(x_{l_x}, y_{l_y}) \sum_{i_x} w_{i_x} \phi_{i_x}^x(x_{l_x}) \phi_{i_x}^x(x_{j_x}) \sum_{i_y} w_{i_y} \phi_{i_y}^y(y_{l_y}) \phi_{i_y}^y(y_{j_y}) \\ &= \sum_{l_x, l_y} f(x_{l_x}, y_{l_y}) \delta_{l_x, j_x} \delta_{l_y, j_y} = f(x_{j_x}, y_{j_y}) , \end{aligned}$$

where we used orthogonality of the rows in the matrix of basis vectors to transform the sums into the Kronecker δ . Figure 5 illustrates the interpolation of a two dimensional function using the tensor basis as well as a variation discussed below.

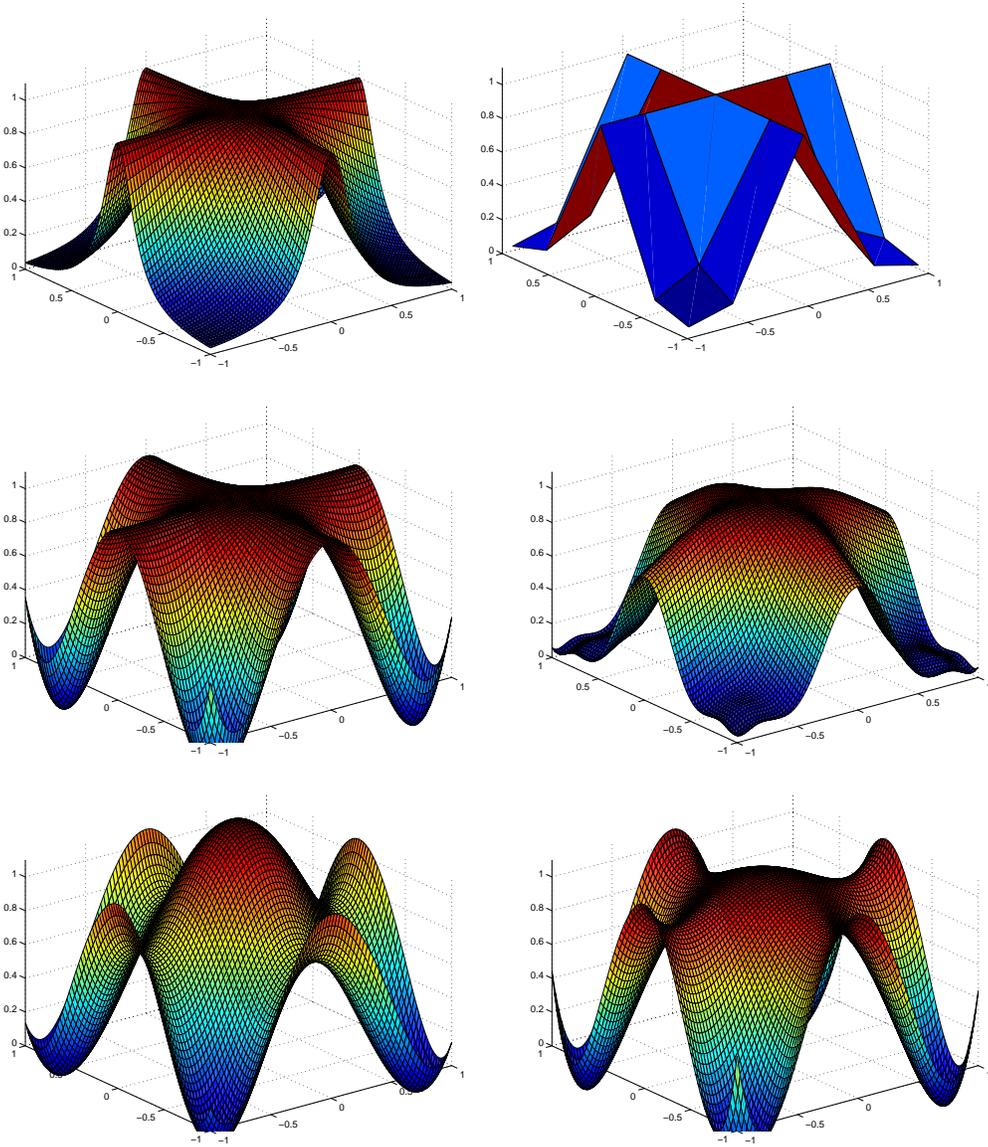


Figure 5 shows an interpolation of the 2-dimensional function $f(x, y) = \frac{1}{1+25(x*y)^2}$, depicted in the upper left corner. In the upper right corner we depict the function evaluated over the grid of 5×5 Chebyshev nodes, and linearly interpolated between these nodes. The middle row shows the approximation using 5 Chebyshev polynomials (and nodes) in each dimension (left), and using 10 Chebyshev polynomials (and nodes) in each dimension (right). Both interpolations use the tensor basis. In the lower row, the left graph uses 5 Chebyshev polynomials in each dimension together with a complete basis, i.e., dropping mixed orders higher than 5 (we kept the full set of interpolation nodes). The right graph uses 5 Chebyshev polynomials in each dimension, but dropped the highest order term proportional to $\Gamma_4 \otimes \Gamma_4$ from the tensor basis.

An analogously constructed tensor basis with a Chebychev interpolation scheme works for any dimension of the state space. However, we have to solve the maximization problem at every point on the grid of interpolating Chebychev nodes. Using 10 nodes in n dimensions implies that we solve the maximization problem at 10^n nodes. A function iteration with three dimensions and a thousand nodes seriously challenge a PC. Doubling the dimensions to six makes the problem intractable without very efficient programming, node reductions, or heavy parallelization over nodes. The algorithm is particularly suitable for parallelization over multiple processors because we can independently solve the maximization problem at different gridpoints. However, given the exponential growth of processor time in the dimension, even parallelization does not solve what is known as the curse of dimensionality. The tensor basis contains the product of the highest order polynomials in all dimensions. Dropping all products of basis functions that have a joint order higher than that of the highest order univariate Chebychev polynomial, we obtain what is known as a complete Chebychev basis. Some algorithms, in particular Smolyak's algorithm, suggest the use of sparser grid that only grow polynomially in the dimension of the state space.

6.6 Other considerations

Here we consider a series of related issues.

6.6.1 Uncertainty

As noted above, the power of dynamic programming, relative to an alternative such as nonlinear programming, is most evident with stochastic problems. Suppose that the stochastic equation of motion

$$x_{\tau+1} = g(a_\tau, x_\tau, \varepsilon_\tau)$$

replaces the deterministic equation of motion in the second line of equation (6.1). Here, ε_τ is the time τ realization of an independently identically distributed random variable with known distribution. The dynamic programming equation is now

$$V^s(x) = \max_a E_\varepsilon \{ U(a, x) + \beta V^{s-1}(g(a, x, \varepsilon)) \}.$$

The approximation proceeds as above, except now we have to take expectations at every stage. If ϵ is distributed continuously, Gaussian quadrature presents an efficient approximation to the expectation integral. Gauss-Legendre quadrature once more approximates a probability weighted integral by a weighted sum. The L quadrature nodes and the L weights in the sum are selected to match first $2L$ moments of the distribution

$$\int_Z z^k p(z) dz = \sum_{l=1}^L w_l x_l^k \quad \text{for } k = 0, \dots, 2L - 1 .$$

Then we approximate the expectation by

$$E_\epsilon \{U(a, x) + \beta V^{s-1}(g(a, x, \epsilon))\} \approx U(a, x) + \beta \sum_{l=1}^L w_l V^{s-1}(g(a, x, \epsilon_l)) .$$

Given an estimate of the value function at stage $s - 1$, $\hat{V}^{s-1}(x) = \phi(x) \mathbf{c}^{s-1}$, at stage s we obtain

$$\mathbf{V}_j^s = \max_a U(a, x) + \beta \sum_{l=1}^L w_l \phi(g(a, x_j, \epsilon_l)) \mathbf{c}^{s-1} . \quad (6.12)$$

We calculate the stage s basis coefficients \mathbf{c}^s as above, to obtain an estimate of the stage s value function, $\hat{V}^s(x) = \phi(x) \mathbf{c}^s$, and proceed to stage $s + 1$. A useful trick to increase the efficiency of the value function iteration is first run the iteration with only a few Gaussian quadrature nodes, e.g. three. The right hand side of equation (6.12) is evaluated many times for the maximization at every node. Therefore, a low number of nodes increases the iteration significantly. Taking the resulting value functions as an initial guess for a second run with more nodes usually converges very quickly. Moreover, this way you can evaluate whether your paths or control rules still depend on the number of Gauss-Legendre nodes in your approximation of the stochastic distribution.

6.6.2 Approximation interval

The choice of the approximation interval $[a, b]$ depends on the model specifics. For example, in a renewable resource problem, the modeler probably knows

the initial value of the stock, and should choose $[a, b]$ to include that value. If the horizon, T , is long enough that the researcher thinks that the stock is likely to approach its steady state, then $[a, b]$ should also include that value. The steady state can be obtained without solving the dynamic problem, merely by solving (typically numerically) two algebraic equations. The steady state condition for the state variable, $x = g(a, x)$, and the steady state of the Euler equation, discussed in Chapter 1, comprise these two algebraic equations. Sensitivity studies with respect to a and b help to determine whether the interval of approximation has been chosen well.

A tighter interval usually gives a better approximation with less nodes. This reasoning becomes important in a higher dimensional state space where the number of nodes is expensive. With a finite time horizon, we can then also decide to choose different intervals (and basis functions) for every time step, tightly bounding the intervals around the relevant path. In a deterministic problem, a tight interval around the relevant path can be very efficient and is generally less problematic. In a stochastic problem, however, we have to be careful not to tighten the interval too tightly around an expected path. In general, stochasticity can lead us along many path in the state space and we want to have an interval that is large enough to capture a reasonable large set of random paths. We have to be particularly careful with tightening the bounds, if the expected value might be influenced significantly by payoffs for low probability events that are further astray from the expected path. Again, a sensitivity analysis with respect to the interval bounds is generally a good idea.

Finally, convergence in real life is, unfortunately, more troublesome than the theory generally suggests. Trying different intervals can also help to overcome convergence problems, sometimes merely because it makes the finite series of Chebychev polynomials better suited to approximate your actual value function. In other cases, your equation of motion might imply that the evaluation of $\phi(g(a, x_j, \varepsilon_l))$ lies outside of the actual interval bounds. This situation can happen under uncertainty and is very likely under uncertainty for some of the far out Gauss-Legendre nodes. Then, a careful choice of the intervals can help to reduce the amount of evaluations that take place outside of the approximated interval, where the quality of the approximation generally deteriorates rapidly. Lastly, in the multi-dimensional setting, different states interact in the economic problem as well in there equation of motions. In the value function iteration algorithm we evaluate all combina-

tions of states within the bounds of our approximation intervals. Sometimes these combinations can lead to situation that are economically implausible or highly unlikely and we should check that these situations might not imply extreme controls that can potentially harm convergence of the algorithm.

6.6.3 More on infinite horizon problems

For an infinite time horizon, we start with an arbitrary initial guess for the value function. However, when do we stop the iteration? Usually, we use a break criterion that is either based on the change of the coefficients estimated in the function approximation, or the maximal change of the value function at a grid point. Once we found a value function approximation satisfying the break criterion, we can derive the control rules and the time paths. Then, we generally want to repeat the iteration taking the solution as an initial guess but reducing the tolerance. Indicator for a reasonable tolerance is when we find that control rules and time paths no longer change, or only change marginally when reducing the tolerance further.²⁶ A different test for the quality of the approximation is to iterate the Bellman equation once more on a much finer grid, only for a single step, check again the maximal deviation of the value function. Finally, observe that for a given number of basis function and interpolation nodes we cannot satisfy an arbitrarily low tolerance criterion. In general, the actual value function, which is the true fixpoint of the Bellman equation, lies outside of the space of functions we can approximate using our basis functions.

We can accelerate the infinite time horizon algorithm significantly by using what is known as *modified policy function iteration* or “*Howard’s method*”. The most costly step in every value function iteration is the maximization. We can think of the value function iteration as follows. For a given value function approximation, we solve for the optimal policies. For a given policy-solution to the maximization problem, we fit a new value function to the right hand side of the Bellman equation. Now assume that the policy choice a^* would indeed be optimal. Would the new fit of the value function also give us the correct value function? No. Assume you would indeed know the correct policy rule (or value at every gridpoint). If you start with an arbitrary guess of the value function, but fix the policy to its optimum, you would face

²⁶Note that time path often stop changing before the control rules do.

the following iteration in s

$$V^{s+1}(x) = U(a^*, x) + \beta V^s(g(a^*, x)). \quad (6.13)$$

In general, if the actual Bellman iteration contracts to its true solution V , then also equation (6.13) will contract to V . Even if we are already close to the optimal policy, we generally still need many iterations to get close to the optimal value function. Moreover, an iteration according to equation (6.13) with a (not necessarily optimal but) fix policy a^* is computationally cheap because we avoid the maximization. Policy iteration suggests that whenever we solved the maximization step for a new policy a^* , we first solve for the corresponding value function before we engage in another optimization step. Howard’s method, or modified policy iteration, suggests to iterate the inexpensive version of the “Bellman equation” (6.13) a few times after every true iteration of the maximizing Bellman equation. That way we contract the value function closer to our new policy before we invest into optimizing our policy variables again.²⁷ A related method suggests to only re-optimize the policy for a random draw of nodes in every iteration.

6.6.4 Various

Non-autonomous problems. So far our payoffs and our equation of motion did not depend on time explicitly, i.e. the problem was autonomous. For a problem with a finite time horizon, the exact same solution algorithm applies to non-autonomous problems, only that the transition equation or the payoff function now depends on s . With an planning horizon, we can solve a non-autonomous problem by making time a state variable. Then every iteration solves the problem for all times, or more precisely at the interpolation nodes in the time dimension. Once the function iteration terminates, we have a value function that we can evaluate for any period and all (other) states. This value function interpolates time smoothly, even if we only have a solution for the time step that we applied to our equations of motions. If

²⁷Be aware that the break criterion has to incorporate the difference between “Howard loops” and maximizing loops. We will generally construct the break criterion based on differences only between maximization loops. But these differences can be larger now for the same approximation quality without using Howard’s method. Rather than using a fixed number of “Howard loops” we can also use a second break criterion for switching back from an iteration according to equation (6.13) to a maximizing iteration.

we make the time step a model parameter, we can still solve the problem for any desired time step. Usually, we expect that a larger time step increases the speed of convergence.

Convergence. Also in cases where theory proves a contraction mapping (see Appendix to chapter lxxx), the numerical algorithm might not converge. The theory generally does not take into account that we only use a finite subset of the basis functions that would span the true function space and that we use a finite approximation interval, even if stochastic equations of motion will usually carry us out of this interval.²⁸ We discussed in the section on the approximation intervals how a careful choice of the interval bounds can sometimes help in achieving convergence. Another trick that sometimes helps to stabilize the function value iteration is *damping*. Here we fit the new Chebychev coefficients as above. However, into the next iteration we pass on a convex combination of the new coefficients and the coefficients calculated in the previous iteration. In particular, if the function iteration is alternating damping can help us to converge to a better solution, or faster, or to converge at all. Finally, we can sometimes *perturb* a parameter that relates to the instability of the algorithm and only slowly increase or decrease the parameter to the desired value. For example, if some parameter value relates to the amount of gridpoints where (or the distance by how much) we jump out of the approximation intervals, we can start setting it to a value that ameliorates the approximation problem. Once we solved the problem, we can take the resulting value functions as an initial guess to resolve the problem with a value closer to the true parameter value, until we might be able to solve the problem we are actually interested in and which was unstable for an arbitrary guess of the value function.

The compecon toolbox. Miranda and Fackler (2002) ?? provide convenient toolboxes in Matlab for solving dynamic programming problems. In particular, they provide algorithms (commands) that create the fit the coefficients and evaluate the function approximation when using Chebychev polynomials or splines. You can download the compecon toolbox from Paul Fackler's homepage. The command line argument

²⁸Stachurski (?) has a nice discussion on "fitted value function iteration" where he analyzes the approximation step also in the theoretical fixpoint argument.

```
fspace=fundefn('basetype',n,a,b,order)
```

`basetype` \in {'cheb','spli'}, creates a structure that we called 'fspace' containing the basis underlying the function approximation. The parameter n determines the order of the basis functions. In a multidimensional application it is a vector determining the order in each dimension. The approximation interval is $[a, b]$, where both endpoints are vectors in the multidimensional case. The parameter 'order' is an optional argument that, in the case of splines, determines the order of the interpolating spline (default is 3, and 'lin' instead of 'spli' produces a linear spline interpolation). The command line

```
c=funfitxy(fspace,x,y)
```

finds the basis coefficients using the approximation structure 'fspace'.²⁹ Here, x is a list of the nodes (rows contains the coordinates of one point, columns a sequence of nodes) at which we evaluate the function, and y the corresponding values.

The command line

```
x=funnode(fspace)
```

returns a vector (onedimensional case) or array structure (multidimensional case) suggesting you the optimal interpolations nodes for the approximation structure `fspace`. These are Chebychev nodes for the Chebychev basis and equidistant nodes for splines. You obtain the basis in matrix notation evaluated at the point or vector of points x using the command line

```
B=funbas(fspace,x) ,
```

from which you find the approximate function value at x as

```
y=B*c .
```

Alternatively you can use the command

```
y=funeval(c,fspace,x)
```

to evaluate the function at the point(s) x . In particular, `funeval` also takes parameters that permit the efficient evaluation of derivatives.

²⁹The command `funfitf(fspace,f)` finds the coefficients directly if f is implemented as a function.

6.7 Other functional equations

We presented function approximation as a means of solving dynamic programming equations, but the same methods are useful in many other applications. Consider the problem of finding the function $f(x)$ that solves the functional equation

$$h(x, f(x)) = 0.$$

We can approximate $f(x)$ using $\hat{f}(x) = \phi(x)c$, with $\phi(x)$ an n dimensional row vector of basis functions. Given the interpolation nodes $x_j, j = 1, 2, \dots, n$, we can choose c to satisfy $h(x_i, \phi(x_i)c) = 0$ at each node. This method of solving functional equations is known as the collocation method. Equivalently, we can find f_i to satisfy $h(x_i, f_i) = 0$ and then choose c to satisfy $\mathbf{f} = \mathbf{\Phi}c$, where the i 'th element of \mathbf{f} is f_i and the i 'th row of $\mathbf{\Phi}$ is $\phi(x_i)$.

Functional approximation is also useful for solving two point boundary value (TPBV) problems, a system of differential equations arising in continuous time optimal control problems. We discuss this problem in detail in Chapter xx, and here merely indicate one approach to a numerical solution. For problems with one state variable and one control variable, the necessary conditions give rise to a TPBV problem consisting of a pair of differential equations with a split boundary condition: we have one boundary condition at the initial time, $t = 0$, and a second at the terminal time, $t = T$.

For example, consider the pair of differential equations

$$\frac{dx}{dt} = h(x, y, t) \quad \text{and} \quad \frac{dy}{dt} = k(x, y, t)$$

with boundary conditions

$$x(0) = x_0 \quad \text{and} \quad x(T) = x_T, \quad \text{given.}$$

The functions h and k and the boundary value x_T are obtained using the necessary conditions to the optimal control problem, as explained in Chapter xx, and the boundary condition x_0 is data. Here we take the differential equations and boundary conditions as given, and discuss the solution.

The solution consists of functions $x(t) = H(t)$ and $y(t) = K(t)$. The goal is to approximate the functions H and K . As above, we choose a set of n basis functions, $\phi(t)$ and approximate H using $\phi(t)c$ and K using $\phi(t)q$, where c and q are n dimensional column vectors of interpolation coefficients.

Substituting these approximants into the original differential equations, we have

$$\frac{d\phi(t)}{dt}c = h(\phi(t)c, \phi(t)q, t) \quad \text{and} \quad \frac{d\phi(t)}{dt}q = k(\phi(t)c, \phi(t)q, t). \quad (6.14)$$

For exposition, we choose $n - 1$ interpolation nodes. These range from 0 to T ; in terms of the notation above, $a = 0$ and $b = T$. We also have the boundary conditions

$$\phi(0)c = x_0 \quad \text{and} \quad \phi(T)c = x_T. \quad (6.15)$$

Requiring that the pair of equations in 6.14 hold at the $n - 1$ interpolation nodes results in $2(n - 1)$ equations; these, plus the two equations in ?? comprise a system of $2n$ equations, which can be used to find the $2n$ interpolation coefficients, the elements of c and q .

6.8 Finite Markov chains

So far, we considered a problem in which both the state and control spaces are continuous. Here we discuss an alternative in which both of these sets consist of a finite number of elements. The methods here are particularly useful in stochastic problems, so we move straight to that setting. We denote the finite set of states of the model by $S = \{x_1, \dots, x_n\}$. The control (or policy) variable can take on the values $a \in \{a_1, \dots, a_m\} \equiv A$. The policy function maps the state of the system into a control: $\sigma : S \rightarrow A$. We can represent the policy function by an n -vector σ with elements $\sigma_i = \sigma(x_i)$.

Range management provides an example of a renewable resource application of this setup. Suppose, as a discrete approximation, that rangeland quality (measured by biomass of available grass) can take a finite number of values. Rangeland quality is the state variable. Quality changes stochastically, and depends in part on stocking rate (number of animals per unit of land). The rancher can improve the quality of the land by applying a treatment (e.g. fertilizer or herbicide), a 0/1 variable. The control variables are stocking rate and the treatment decision. If there are $\frac{m}{2}$ possible stocking rates, then a_j , $j = 1, 2, \dots, m$, identifies an ordered pair; the first element is the stocking rate and the second element is a 0/1 variable identifying whether the rancher uses the treatment. Generalization to higher dimensional state variables, or

additional controls, is straightforward, but of course increases the size of the numerical problem.

We assume that the state variable follows a stationary Markov process: the probability distribution of the next period state depends only on the current state and the current action, not on the history of actions or calendar time. The Markov assumption is quite flexible. We can allow for the possibility that the probability of transition from state x_i to x_j depends on past values of the state and/or the control, by increasing the dimension of the state variable. For example, if the transition probability depends on both current and lagged values of the state variable, we define a new state variable, the ordered pair, $\tilde{x} \in S \times S$, where the first element equals the current value of the (original) state variable and the second element equals its lagged value. If x can take n possible values, \tilde{x} takes n^2 possible values. We express the stochastic transitions, replacing the equation of motion, using a *stochastic kernel*.

Definition 2 A function $p : S \times S \rightarrow [0, 1]$ is a stochastic kernel if

1. $p(x, y) \geq 0$ for all $(x, y) \in S \times S$, and
2. $\sum_{y \in S} p(x, y) = 1$ for all $x \in S$.

For any given state $x \in S$, the stochastic kernel $p(x, \cdot)$ defines a probability distribution over the state space. We use the stochastic kernel to specify the probability distribution of period's state, given the current state (or a distribution thereof).

In our dynamic programming application, the transition probabilities $p(x, \cdot)$ generally depend on the policy σ . Our model will usually characterize the probability that the next period state is x_j , if the current state is x_i and we choose policy a_k . Given a policy function σ , we then obtain the probability $p_\sigma(x_i, x_j)$ as the probability that the next period state is x_j , if the current state is x_i and we choose policy $\sigma(x_i)$. Then, for any policy function σ we have a stochastic kernel p_σ . We can represent these stochastic kernels in matrix form as

$$\mathbf{p}_\sigma = \begin{pmatrix} p_\sigma(x_1, x_1) & \dots & p_\sigma(x_1, x_n) \\ \vdots & \ddots & \vdots \\ p_\sigma(x_n, x_1) & \dots & p_\sigma(x_n, x_n) \end{pmatrix}.$$

In our case of a finite state space, a stochastic kernel corresponds to the conditional probability of the next period's state being y , given that the current state is x . Let X_t denote the random variable characterizing the system's state in period t . Given the initial condition $X_0 = \bar{x}$ and the stochastic kernel p_σ , we generate a Markov chain $(X_t)_{t \geq 0}$ of states by setting $X_0 = \bar{x}$ and drawing

$$X_{t+1} \sim p_\sigma(X_t, \cdot) \quad (6.16)$$

for all $t \geq 0$. We dedicate the remainder of this subsection to finding the optimal policy σ^* .

As before, $U(x, a)$ is the single period payoff, and β the discount factor. The transition equation $X_{t+1} \sim p_\sigma(X_t, \cdot)$ replaces the deterministic or stochastic difference equation from Section 6.2. The objective is to maximize

$$E_0 \sum_{\tau=0}^{\infty} \beta^\tau U(x_\tau, a_\tau),$$

subject to the transition probabilities of the states and the initial condition. We assume that the time horizon is infinite, $T = \infty$, and that the problem is autonomous. In this case, the value of being in state x_0 , denoted $V(x_0)$, does not depend on calendar time. For an arbitrary initial condition, x , the dynamic programming equation is

$$V(x) = \max_a \{U(x, a) + \beta E_{x'} V(x')\}, \quad (6.17)$$

where x' is the value of the state variable in the next period. The probability of transition from a particular value of x to a particular value of x' depends on the choice of the current control a (policy function in equation 6.16). If we knew the function V , the fact that there are a finite number of possible values of the control variable makes it straightforward to solve the problem in equation (6.17). The key is to find the function $V(x)$. There are three methods of solving this problem: value function iteration, policy function iteration, and linear programming.

The *value function iteration* approach begins with an initial guess \mathbf{V}^0 , an n dimensional vector whose i 'th element equals the initial guess of the value of being in state x_i . Given the guess \mathbf{V}^s , we update the guess by solving

$$V^{s+1}(x) = \max_a \{U(x, a) + \beta E_{x'} V^s(x')\},$$

for all $x \in S$. We stop the iterations when the sup norm $\|\mathbf{V}^{s+1} - \mathbf{V}^s\|$ is less than a tolerance chosen by the modeler, or when the optimal decision at any state does not change from one iteration to the next. This algorithm is similar to that used in Section 6.2, except that here we store the estimated values of the value function at every possible state, rather than approximating the value function using basis functions. A variety of methods speeds convergence.

The *policy function iteration* begins with a guess of the policy function. Denote the guess of the policy function at iteration s as the n dimensional vector σ^s , whose i 'th element gives the guess, at iteration s , of the optimal action when the state is x_i . Let \mathbf{U}_{σ^s} denote the n dimensional column vector of single period payoffs under policy σ^s ; the i 'th element being $U(x_i, \sigma(x_i))$. The i 'th row vector in \mathbf{p}_{σ^s} captures the probability distribution over next period's states, given the current system state is x_i . For a given policy function σ , the value of the program equals

$$\begin{aligned} \mathbf{V}^s &= \mathbf{U}_{\sigma^s} + \beta \mathbf{p}_{\sigma^s} \mathbf{V}^s \\ \Rightarrow \mathbf{V}^s &= [\mathbf{I} - \beta \mathbf{p}_{\sigma^s}]^{-1} \mathbf{U}_{\sigma^s}, \end{aligned} \quad (6.18)$$

where the i 'th element of \mathbf{V}^j is the value of the program when the current state is x_i and the planner uses the policy function σ^s . Given this guess of \mathbf{V}^j , we update our policy function setting

$$\sigma^{s+1}(x) = \arg \max_a \{U(x, a) + \beta E_{x'} V^s(x')\}$$

for all $x \in S$. This algorithm, unlike value function iteration, converges to the optimal decision rule and value function in a finite number of iterations. It requires solving the linear system (6.18) in each iteration.

The third approach uses *linear programming*. Denote by \mathbf{V} the column vector of values of the program, with i 'th element equal to the value of the program when the current state variable is x_i . Denote the policy function using control a_j in all states by $\sigma = a_j$, and \mathbf{U}_{a_j} as the vector of values of the single period payoff, given that the planner uses the control a_j . The i 'th element of \mathbf{U}_{a_j} is $U(x_i, a_j)$. By optimality, it must be the case that

$$\mathbf{V} \geq \mathbf{U}_{a_j} + \beta \mathbf{p}_{a_j} \mathbf{V}.$$

The right side of this inequality equals the vector of values of the payoff (whose elements correspond to the different values of the state variable), given that the planner uses action a_j at every state. Of course, this action need not be optimal at any state; hence, the inequality. We have m systems of this form, one corresponding to each possible action. Stack the m vectors \mathbf{U}_{σ_j} for all $a_j \in A$ to create the mn dimensional vector \mathbf{U} and stack the m matrices \mathbf{p}_{a_j} to create the $mn \times n$ dimensional matrix \mathbf{P} . Also denote $\mathbf{1}$ as the n dimensional column vector consisting of 1's, and \mathbf{I} as the n dimensional identity matrix. The system of inequalities can be written as

$$[\mathbf{1} \otimes \mathbf{I}] \mathbf{V} \geq \mathbf{U} + \beta \mathbf{P} \mathbf{V} \quad \Rightarrow \quad [\mathbf{1} \otimes \mathbf{I} - \beta \mathbf{P}] \mathbf{V} \geq \mathbf{U}. \quad (6.19)$$

The value of the program is the solution to the linear programming problem

$$\min \mathbf{1}' \mathbf{V} \text{ subject to inequality (6.19).}$$

6.9 Comments on Markov chains

Finite Markov chains provide a simple means of formulating a dynamic problem, while permitting a general description of stochastics: the probability distribution of the next-period state can depend in any manner on the current state and control. The main cost of this method is that it might be necessary to use a large number of values of the state variable to describe the problem with an acceptable degree of accuracy. If we need significantly fewer nodes to approximate the continuous functions, then the methods discussed in Section 6.2 probably provide a more efficient modeling strategy. However, those models typically use a simple representation of stochastics, obtained by making the equation of motion a function of a random variable. This model restricts the manner in which the current state and control affect the probability distribution of the state in the next period.

The analysis of the optimally controlled state also differs under the two approaches. Here, we restrict attention to the autonomous case. With continuous action and state space, the approximation of the autonomous control rule is $\phi(x) \mathbf{d}$, where $\phi(x)$ is the row vector of basis functions and \mathbf{d} is the column vector of basis coefficients used to approximate the control rule. The state evolves according to

$$x_{\tau+1} = g(x_\tau, \phi(x_\tau) \mathbf{d}, \varepsilon_\tau)$$

where ε_τ is the random variable. By taking many draws of the random variable, we can use Monte Carlo methods to study the stochastic evolution of the state, and to approximate, for example, the steady state distribution and the probability that the state variable crosses some threshold.

With a finite Markov chain, we can calculate those kinds of probabilities exactly, rather than by means of Monte Carlo experiments. Denote the optimal policy function by σ^* . The ij 'th element \mathbf{p}_{σ^*} equals the probability that the state transitions from x_i to x_j when the planner uses the optimal action at state x_i . Let the n -dimensional row vector q_t represent a probability distribution over S in period t . If we know the value of the state at time t , then q_t is a unit vector, consisting of 1 in the position of the known value of the state, and 0's elsewhere. The probability that the state transitions from i to j in t periods is $\mathbf{p}_{\sigma^*}^t$; see problem 1. The equation of motion for the probability distribution q_t satisfies the linear difference equation

$$q_{t+1} = q_t \mathbf{p}_{\sigma^*}. \quad (6.20)$$

For example, if we know, with probability 1, that $X_t = x_i$, then equation (6.20) returns the i 'th row of \mathbf{p}_{σ^*} as q_{t+1} , the probability distribution of the next period state variable.

An eigenvalue λ and a “right” eigenvector x solve $\mathbf{p}_{\sigma^*}x = \lambda x$. An eigenvalue λ and a “left” eigenvector x solve $x\mathbf{p}_{\sigma^*} = \lambda x$. The sets of right and left eigenvectors are not the same, but the sets of right and left eigenvalues are the same. The matrix \mathbf{p}_{σ^*} has a principal eigenvalue of 1: the largest eigenvalue λ solving $x\mathbf{p}_{\sigma^*} = \lambda x$ is $\lambda = 1$. Thus, there exists a vector q_∞ that solves $q_\infty \mathbf{p}_{\sigma^*} = q_\infty$. The vector q_∞ is a steady state distribution: if the distribution equals q_∞ in the current period, then it remains unchanged in each subsequent period. In general, there may be multiple steady states to the difference equation (6.20). However, if there is a unique steady state (the principal eigenvalue is of multiplicity 1), then the Markov chain is ergodic. Such a chain asymptotically approaches its unique steady state distribution from any initial distribution, including from a degenerate distribution, where we know the value of the state. In the rangeland example, we can calculate the long-run expected value of the rangeland. If it is optimal to perform the “treatment” only when the rangeland falls below a critical level, we can calculate the expected long-run frequency of the treatment.

If there are multiple steady state probability distributions, then the long-run

probability distribution may depend on the initial probability distribution. For example, if there are two steady state distributions, q_∞^A and q_∞^B , then there are three sets of values of x , say A, B, C . Degenerate probability distributions associated with initial conditions in set A approach the steady state distribution q_∞^A , and those with initial conditions in set B approach q_∞^B . Initial conditions in set C might transition to set A or to set B . We can calculate the probability that the state moves from an initial condition into a particular set, e.g., the probability that it moves from a point in set C in the initial period, to either the set A or B . For example, suppose that the pollution stock in a lake changes stochastically, depending on the current stock and the level of emissions. Moreover, the expected decay of the stock is high for small stocks, but low for large stocks. In this case, once the stock is larger than a threshold it might be prohibitively expensive or even infeasible to cause the stock to fall below this threshold (e.g. by reducing emissions). Chapter xx considers a deterministic version of this “shallow lake problem”. In the stochastic version, under the optimal emissions policy, there may be two steady state distributions, corresponding to high and low pollution stocks; denote these as q_∞^A and q_∞^B . When the process is modelled using a finite Markov chain, we can calculate these distributions and the associated sets A, B, C described above. The set A is the set of initial conditions from which the stock asymptotically approaches, with probability 1, the high pollution steady state distribution, and the set C is the set of initial conditions from which the stock approaches the high pollution steady state distribution with probability greater than 0 and less than 1. We can also calculate the mean passage time, defined as the expected amount of time it takes the stock to transit from a given initial condition, into a particular set. The nature of the problem determines the type of comparative dynamics experiments one is likely to undertake. In the shallow lake problem, we can, for example, see how a change in the cost of emissions control affects the probability distribution of the pollution stock.

6.10 An Application to the Integrated Assessment of Climate Change

The problem replicates a stylized version of William (Nordhaus 2008) DICE model as a recursive dynamic programming model. DICE is an open source integrated assessment model. Basically, it is a Ramsey growth model that is

enriched by emissions, pollution stock, and temperature related damages to world GDP. Some of the key exogenous parameters in the original model are time dependent. We omit the time of these parameters. In contrast to DICE, this problem allows you to model uncertainty in a coherent way. Uncertainty is persistent and affects decisions in every period. Uncertainty enters as an iid shock on climate sensitivity, which characterizes the temperature response to the radiative forcing caused by a doubling in atmospheric CO₂ concentrations with respect to the preindustrial level.³⁰ The climate sensitivity parameter is one of the key unknowns in modeling global warming because of several feedback processes involved when translating CO₂ increase into temperature change. We have an annual time step and face an infinite planning horizon. The model solves for the optimal expected trajectories of the climate economy controlling for emissions and investment.

6.10.1 Welfare

We use the intertemporally additive expected utility standard model. It evaluates scenarios by aggregating instantaneous welfare linearly over time and aggregating linearly over risk states by taking expected values. Thus, the social planner maximizes

$$U = \sum_t \beta^t L_t u\left(\frac{C_t}{L_t}\right) = \sum_t \beta^t L_t \frac{\left(\frac{C_t}{L_t}\right)^\rho}{\rho} = L_t^{1-\rho} \sum_t \beta^t \frac{C_t^\rho}{\rho}$$

where L_t is the population size and β is the discount factor stemming from the additional assumption of stationary preferences. The assumption of a power utility function returns to us a constant elasticity of intertemporal substitution. We can drop L_t in the welfare equation because it comes down to a multiplicative constant of welfare having no real effects. Anticipating that the climate-economy equations introduced in section 6.10.2 will depend on the two state variables capital K_t and CO₂ pollution stock M_t and will be controlled by the consumption-investment decision and the abatement rate μ_t we can write the according dynamic programming equation as

$$V(K_t, M_t) = \max_{C_t, \mu_t} \frac{C_t^\rho}{\rho} + \beta V(K_{t+1}, M_{t+1}) ,$$

³⁰Obviously the iid nature is not quite realistic but saves us additional state variables.

with control variables consumption c_t and emission control rate μ_t . Because we will assume a constant population, we can substitute per capita consumption c_t by aggregate global consumption C_t .³¹

6.10.2 The Climate Economy

The decision maker maximizes his value functions under the constraints of the following stylized model of a climate enriched economy. The model is largely a (very) reduced form of Nordhaus (2008) DICE-2007 model. All parameters are characterized and quantified in table ?? on page 354. The economy accumulates capital according to

$$K_{t+1} = (1 - \delta_K) K_t + Y_t - C_t ,$$

where δ_K denotes the depreciation rate, Y_t denotes net production (net of abatement costs and climate damage), and C_t denotes aggregate global consumption of produced commodities. Instead of trying to model the full carbon cycle, which would be very costly in terms of stock variables, we assume an exponential decay of CO₂ in the atmosphere at rate δ_M

$$M_{t+1} = M_t (1 - \delta_M) + E_t .$$

The variable E_t characterizes overall yearly CO₂ emissions. We use values for M_t and E_t characterizing CO₂ only. However, at the given level of abstraction a rescaled version of M_t could be thought of as representing greenhouse gases in CO₂ equivalents more generally. Emissions are composed of industrial emission (first term) and emissions from land use change and forestry B (which are assumed to be constant)

$$E_t = \sigma (1 - \mu_t) (AL)^{1-\kappa} K_t^\kappa + B$$

The constant σ specifies the emissions to GDP ratio and the control variable μ_t is the abatement rate. The constant AL represents effective labor (technology level and labor). The product $(AL)^{1-\kappa} K_t^\kappa$ is the global gross product (gross of abatement costs and climate damage). The net product is obtained

³¹The change results in an affine transformation of the value function that leaves the decision problem unchanged.

from the gross product as follows

$$Y_t = \frac{1 - \Lambda(\mu_t)}{1 + D(T_t)} (AL)^{1-\kappa} K_t^\kappa = \frac{1 - \Psi \mu_t^{a_2}}{(1 + b_1 T_t + b_2 T_t^{b_3})} (AL)^{1-\kappa} K_t^\kappa$$

where

$$\Lambda(\mu_t) = \Psi \mu_t^{a_2}$$

characterizes abatement costs as percent of GDP. The parameter Ψ is the cost coefficient and the cost is convex in the emission control rate μ_t . The equation

$$D(T_t) = b_1 T_t + b_2 T_t^{b_3}$$

characterizes the climate damage as percent of GDP depending on the temperature difference T_t of current with respect to temperatures in 1900.

In the model, temperature is an immediate response to the radiative forcing caused by the stock of CO_2 in the atmosphere

$$T_t = s_t \frac{\ln \frac{M_t}{M_{preind}}}{\ln 2} + \frac{EF}{\eta} . \quad (6.21)$$

where s_t denotes climate sensitivity, i.e. the temperature response to a doubling of CO_2 in the atmosphere with respect to preindustrial concentrations. The second term in equation (6.21) represents external forcing that is caused by other greenhouse gases. The model uses a climate sensitivity of $s_t = s \approx 3$. Finally the system is bound by the following constraints on the control variables consumption

$$0 \leq C_t \leq [1 - \Lambda(\mu_t)] Y_t^*$$

and abatement

$$0 \leq \mu_t \leq 1 .$$

The constraint on consumption uses gross output less climate damages $Y_t^* = \frac{(AL)^{1-\kappa} K_t^\kappa}{1 + D(T_t)}$. Rewriting the constraints in terms of abatement expenditure Λ_t

rather than the abatement rate $\mu_t = \left(\frac{\Lambda_t}{\Psi}\right)^{\frac{1}{\alpha_2}}$ makes the ‘right hand side’ constraints linear in the controls

$$\begin{aligned} C_t + \Lambda_t Y_t^* &\leq Y_t^* \\ \Lambda_t &\leq \Psi . \end{aligned}$$

As linear constraints are preferred by numerical solvers, we actually use Λ_t rather than μ_t as the control variable in the numerical implementation of this model.

6.10.3 The Code

The model is an implementation of the abstract procedures we discussed in the class when introducing some of the mechanics of solving numerical problems. For evaluating the Chebyshev polynomials we use the package by ?. The same is true for calculating the optimal nodes where to approximate the value function and for fitting the new coefficients. You can download the package from³²

<http://www4.ncsu.edu/~pfackler/compecon/toolbox.html>

The code has the option to use ‘damping’ as discussed above: rather than employing the new coefficients c^{new} for the subsequent iteration we use a convex combination of the previous and the new coefficients $c = (1-d)c^{new} + dc^{prev}$. For $d = 0$ we are back to the standard procedure. If you use the standard maximization routine in Matlab *fmincon*, you will probably not see an effect of damping. However, if you use other feeware of commercial solvers some of which are much faster than *fmincon*, you might observe different convergence properties (e.g. when using *knitro*). The code also enables you not to ‘force’ the new fit of the Chebychev function to exactly run through the values generated at the nodes on the right hand side of the dynamic programming equations. We can evaluate the function at more nodes, but ‘drop’ some of them when fitting the new coefficients. Thus, we then have more evaluation nodes than coefficients and do a least square fit.³³ The algorithm that you can download from the course website approximately

³²Of course you are happily invited to add the Matlab code for the according procedures to the problem set rather than using the package...good exercise.

³³? *funfit* routine used for acquiring the new coefficients automatically does a least square fit when it receives more values than it has coefficients.

works as follows:

File *model-KM* (main file)

- “run param-KM-DICE-initial” loads the parameters. These approximately correspond to the initial parameters of DICE in the initial year
- you set: topdir= ‘C: path to where you saved the model folder’ (change!)
- you can set “loadswitch” if you have initial values you want to use from previous runs. For this purpose you have to go to the directory with the initial values and set loadswitch to 4 rather than 0
- defines the variables relevant for the function approximation, i.e. number of nodes, approximation intervals, damping, number of nodes to be dropped
- creates the function space (fspace) used for the approximation and the nodes used for evaluation (s)
- calls the function iteration procedure
- plots results

File: *function-iterate-KM* (does the actual work)

- Saves old values
- Iterates over all nodes doing the following
 - solve the maximization problem on the r.h.s. of the DPE
 - the outbounds stuff is all about calculating information that comes up on the screen (and is also stored in .txt files) whether we jumped out of the approximation interval in $t + 1$ - it is only for informational purposes
- Fits the Chebyshev coefficients for the newly calculated values at the different nodes
- Calculates by how much coefficients and function values have changed from the old iteration and puts it out on the screen (together with the information how many nodes jumped out of the approximation interval).

In the end you will have a folder created at the location that you specified. Apart from the coefficient matrix and some other information it contains plots of the value function, the control rules, and the time paths.

6.10.4 The Questions

1. Download and install (?) toolbox. Test that it is properly installed. Download the model from the course website. Set your Matlab path to include to folder where you have saved the model code.
2. Familiarize with the code and run the model as is. Which intervals did you use for the value function iteration? How many nodes did the predefined model run put on each dimension? Did the model run use damping?
3. Go to the results folder and check out the plots. Do the optimal pollution stock and the optimal capital stock increase or fall over time? In which range does the capital stock move over the 300 depicted years? Do you think the interval used for the value function approximation contains the steady state level of capital? Does the value function increase or fall in the capital and the pollution stock? Does the social cost of carbon increase or fall in the capital and the pollution stock? Note that the social cost of carbon is measured in units of capital.
4. Increase the number of nodes to 10 on both dimensions. What do you find?
5. Leave the number of nodes at 10 and use the intervals [80, 220] and [670, 920]. What do you find?
6. Set the intervals to [100, 180] and [670, 920]. Set the number of nodes in both dimensions to 3. Comment out the current parameterization in line 36 of model-KM-cert and use instead the parameterization called param-KM-DICE-initial-doubleA. It is the same as the previous param-KM-DICE-initial except that it doubles the effectivity of labor, which happens in the full version of DICE after about 80 years. Run the model.

Is the CO_2 stock increasing or decreasing over time? Relate the observation to your previous result with half the A value.

7. Increase the number of nodes in each dimension to 8. Run the model. What do you find? Increase the number of evaluation nodes to 9 but drop one node when fitting coefficients (i.e. set “nodetocoeff = [1 1]”). Moreover, set damping to “damp=0.3”. Run the model. What do you find?
8. Set damping and nodedropping back to zero. Change loadswitch to “loadswitch=4” and go into the directory carrying the name of your run with 3 nodes and double labor productivity. Chose 8 nodes on every dimension. Run the model. What do you find?
What can you learn from this finding? (That question is not about the graphs, which should look like in the earlier run...)

The main drivers of the full DICE model are the exogenous changes in productivity (A) and abatement costs (Ψ) as well as an exogenous decline in the carbon intensity of production (σ in the text and *emint* in the code). These exogenous changes explain a large part of the differences that you see in the plots that compare abatement rate, social cost of carbon, CO_2 stock and temperature to those of the original DICE model. I included a parameterization that halves the abatement cost as it would happen after about 80 years in DICE and a parameterization that puts both A and Ψ to the levels after approximately 80 years in case you want to play with it. Another shortcoming of the simplified two state model is that it does not capture the delay in temperature increase.

6.11 Related literature

Section 6.2 is based primarily on Miranda and Fackler (2002), especially chapter 6. That book is particularly valuable for practitioners, because it is integrated with a Matlab-based “toolbox” containing routines that implement the methods discussed in the book, and many solved problems. Stachurski (2009) ?? offers an excellent combination of a thorough treatment of the underlying theory and a discussion including code in Python of the underlying numerical implementation. Judd (199?) is also a valuable reference, containing additional methods and going deeper into some technical issues. Section

6.8 is based primarily on Bertsekas (1976) and Bertsekas (200?). Parzen (1962 – check for later editions) provides a good introduction to finite Markov chains. The rangeland example is from Karp and Pope (1984).

6.12 Problems

1. Suppose that $n = 2$, i.e. the state takes two possible values. Let r equal the probability of transition from state 1 to state 1 (i.e. remaining in state 1) and let v equal the probability of transition from state 2 to state 2. Write the state transition matrix for this example, \mathbf{p} , and show that \mathbf{p}^2 equals the state transition matrix whose elements give the probability of transition from state i to state j in two periods. Using an inductive proof, show that \mathbf{p}^t equals the transition matrix whose elements give the probability of transition from state i to j in t periods.
2. Approximate the 2-dimensional function $f(x, y) = (1 + x)^0.1 \ln(1 + y)$ on the interval $[-1, 1]$. Do not use a toolbox, but use the formulas discussed in this chapter. Use a tensor basis of Chebychev functions and Chebychev nodes. Approximate the function first using only 3 and then using 10+ basis function in each dimension (whatever number larger or equal to 10 you like). Plot the original function and the two approximations. Send me code and plots by email. Mark the resulting Chebychev coefficients in your output.
3. Go over the problem described in Section 6.10 and answer the questions in Section 6.10.4.

7 The Maximum Principle and Pollution Control

The section introduces a wide-spread approach to intertemporal optimization in continuous time. In economics it runs under the names “Maximum Principle” and “optimal control theory”. A related approach in physics dates back quite a bit longer and runs under “Hamilton’s canonical equations”. The method is particularly convenient for optimization under certainty. It translates the intertemporal optimization problem into a static optimization problem and a set of ordinary differential equations. We will analyze these equations using the phase diagram and local approximation techniques introduced in the previous section. We discuss the example of a stock pollution problem.

7.1 Intertemporal Optimization: Motivation

Many problems in the field of environmental and resource economics as well as macroeconomics are of the form

$$\max \int_0^T U(x_t, y_t, t) dt \quad (7.1)$$

$$\begin{aligned} \text{s.t. } i) \quad & \dot{x}_t = f(x_t, y_t, t) && \text{(equation of motion)} \\ ii) \quad & x_0 \text{ given} && \text{(initial condition)} \\ iii) \quad & x_T = \bar{x} && \text{(terminal condition)} \\ iii') \quad & x_T \text{ free} && \text{(alternative terminal condition)} . \end{aligned}$$

Here, x is the state variable and y is the control. Frequently, the utility function will be of the form

$$U(x_t, y_t, t) = u(x_t, y_t) \exp(-\rho t)$$

indicating discounting with the constant rate of pure time preference ρ . Alternatively, we can take $u(x_t, y_t)$ to be a monetary payoff and ρ (or then rather writing it r) as the market interest rate. In the example of a greenhouse gas stock pollution problem we can think of the state variable x_t as representing the stock of CO_2 in the atmosphere and the control y_t as representing emissions. The stock would cause disutility, while the emissions

would be tied to welfare increasing consumption processes. In this example a reasonable first approximation³⁴ to the equation of motion is

$$\dot{x}_t = f(x_t, y_t, t) = y_t - \delta x_t . \quad (7.2)$$

Given that by now we should be very familiar with the discrete time equation of motion for the stock pollution problem, we want to derive equation (7.2) from its discrete time analog. Hereto we replace the unit time step by a time step Δt in the discrete equation of motion:³⁵

$$\begin{aligned} x_{t+\Delta t} &= y_t \Delta t + (1 - \delta \Delta t) x_t & (7.3) \\ \Rightarrow x_{t+\Delta t} - x_t &= y_t \Delta t - \delta \Delta t x_t \\ \Rightarrow \frac{x_{t+\Delta t} - x_t}{\Delta t} &= y_t - \delta x_t \\ \Rightarrow \dot{x}_t &= y_t - \delta x_t . \end{aligned}$$

In these intertemporal optimization problems we also have to specify a terminal condition. Example *iii*) above fixes the terminal stock to some given quantity \bar{x} , which in our example could correspond to a desired long-run CO_2 concentration target. Alternatively, example *iii'*) states that at time T the stock is free and, thus, the terminal state becomes part of the optimization problem. We will encounter more terminal conditions further below. While for the moment you might want to think of x and y as simple numbers, the following reasoning permits these variables to be vectors of real numbers.

Recall how you would solve a static constraint optimization problem using the Lagrange function (or method of Lagrange multipliers). For example, for the problem

$$\max u(x, y) \text{ s.t. } g(x, y) = 0$$

you would maximize the Lagrange function

$$\max L(x, y, \lambda) = u(x, y) + \lambda g(x, y) .$$

³⁴Atmospheric carbon does not actually decay but gets absorbed over time in different sinks. The time scale of the decay depends on the various sinks and is non-exponential.

³⁵In this step you actually have to think carefully about where a “1” turns into a Δt . You cannot simply read that off from the equation of motion, but you have to think about the underlying dynamics.

You could also use the Lagrangian to solve the discrete time version of our stock pollution problem

$$\begin{aligned} \sum_{t=0}^T U(x_t, y_t, \tau) \text{ s.t. } & x_1 = y_0 + (1 - \delta)x_0 \\ & x_1 = y_1 + (1 - \delta)x_1 \\ & \vdots \\ & x_{t+1} = y_t + (1 - \delta)x_t \end{aligned}$$

for all $t \in \{0, \dots, T-1\}$ and $x_T = \bar{x}_t$, x_0 given. You would have to maximize the Lagrangian³⁶

$$\begin{aligned} \max L(x_1, y_1, \lambda_1, \dots, x_T, y_T, \lambda_T) &= \sum_{t=0}^T U(x_t, y_t, t) \\ &+ \sum_{t=1}^T \lambda_t [x_t - (y_{t-1} + (1 - \delta)x_{t-1})]. \end{aligned} \quad (7.4)$$

The Lagrange multipliers are the shadow values of the pollution stock. The value of λ_t tells you how much (present value) welfare you would gain if you had an exogenous increase of the pollution stock x_t by 1 unit in period t . It should obviously be negative for pollution.

We discuss the discrete time case for two reasons. First, it gives an intuition why the derivation of the maximum principle in continuous time starts out with a somewhat surprising but clever idea. Second, simply taking the continuous time limit of the Lagrangian in equation (7.4) above and employing some dirty (not quite correct) math, will yield a good heuristics for arriving at the necessary conditions for the intertemporal optimization problem laid out in the maximum principle. We have already derived the continuous time limit of the constraints (equation 7.3). At the same time the sums in equation (7.4) turn into integrals, leaving us with the continuous time limit of the

³⁶Note that under the terminal condition $x_T = \bar{x}$ we cannot optimize the Lagrangian over x_T , but have instead the equation $x_T = \bar{x}$. However, if x_T is free we can optimize of x_T . In particular, if the utility would be independent of the stock, it would give us the condition $\lambda_T = 0$.

Lagrange function

$$\begin{aligned} \max L((x_t)_{t \in [0, T]}, (y_t)_{t \in [0, T]}, (\lambda_t)_{t \in [0, T]}) \\ = \int_0^T U(x_t, y_t, t) + \lambda_t [\dot{x}_t - (y_t - \delta x_t)] dt . \end{aligned} \quad (7.5)$$

This equation will be central to deriving the maximum principle in the next section.

7.2 The Maximum Principle: Derivation

We want to maximize the objective function

$$\max \int_0^T U(x_t, y_t, t) dt \quad \text{s.t.} \quad \dot{x}_t = f(x_t, y_t, t)$$

given initial and terminal state. We *assume* that U and f are continuously differentiable. As long as we make sure that the constraint

$$\dot{x}_t = f(x_t, y_t, t) \Leftrightarrow f(x_t, y_t, t) - \dot{x}_t = 0$$

is satisfied we can just as well maximize the objective function

$$\int_0^T U(x_t, y_t, t) + \lambda_t [f(x_t, y_t, t) - \dot{x}_t] dt , \quad (7.6)$$

for an arbitrary function λ_t . Given we intuitively derived this expression from the discrete time Lagrange function the reformulation hopefully does not strike you as odd anymore. Think of the second term in the integral as a zero in evening gown. While in the moment λ_t is an arbitrary function, we later choose it in a convenient way that will also restore its interpretation as the shadow value of capital. For the moment, we only require the function λ to be (piecewise) differentiable, so that we can use integration by parts³⁷

³⁷Integration by parts is the analog to the product rule of differentiation:

$$\frac{d}{dt} uv = \dot{u}v + u\dot{v} \Rightarrow \int_a^b \frac{d}{dt} uv dt = \int_a^b \dot{u}v dt + \int_a^b u\dot{v} dt \Rightarrow uv|_a^b - \int_a^b \dot{u}v dt = \int_a^b u\dot{v} dt .$$

to transform equation (7.6) into

$$\int_0^T U(x_t, y_t, t) + \lambda_t f(x_t, y_t, t) + \dot{\lambda}_t x_t dt - \lambda_t x_t|_0^T .$$

We define the first two terms of the integrand as the so called Hamiltonian

$$H(x_t, y_t, \lambda_t, t) = U(x_t, y_t, t) + \lambda_t f(x_t, y_t, t) ,$$

delivering the objective function

$$\int_0^T H(x_t, y_t, \lambda_t, t) + \dot{\lambda}_t x_t dt - \lambda_T x_T + \lambda_0 x_0 . \quad (7.7)$$

Remark: Here is an intuitive preview of what we do more rigorously below.

A mathematician would bite your head off if you do this, but it might be useful to see what we are getting at if you have never seen the calculus of variation before. Let us maximize the objective function $L \equiv (7.7)$ with respect to the same arguments that you would maximize a Lagrange function in equation (7.4) or (7.5). Starting with the controls y_t at every point in time we obtain

$$\frac{\partial L}{\partial y_t} = 0 \quad \Rightarrow \quad \frac{\partial H}{\partial y_t} = 0 .$$

Optimizing with respect to the state variables x_t yields

$$\frac{\partial L}{\partial x_t} = 0 \quad \Rightarrow \quad \frac{\partial H}{\partial x_t} + \dot{\lambda}_t = 0 \quad \Rightarrow \quad \frac{\partial H}{\partial x_t} = -\dot{\lambda}_t . \quad (7.8)$$

The derived equation turn out to be the correct necessary conditions for a maximum of equation (7.7). However, why would a mathematician bite your (or our) head off? First, the Hamiltonian is a density in equation (7.7) and, as such, the Hamiltonian at a given point in time has a Lebesgue measure of zero. Thus, we have to invoke some sort of continuity argument, which we do by employing the calculus of variation. Second, we cannot vary the states independently of the controls at every point in time, so it is not obvious that we really have the freedom to arrive at equation (7.8).

Let $y^* : [0, T] \rightarrow \mathbb{R}$ be the optimal control. We *assume* that y^* is (piecewise) continuous. A change of the optimal control path cannot increase the value of the integral in equation (7.7). In order to employ this simple insight for our calculus we define a set of possible deviations from the optimal control by the set of continuous functions $\mathcal{H} = \{h \in C^0[0, T]\}$.³⁸ Then

$$y = y^* + ah$$

for $h \in \mathcal{H}$ and $a \in [-1, 1]$ defines an alternative control path.³⁹ Let $\tilde{x}(a, h)$ be the solution to the equation of motion $\dot{x}_t = f(x_t, y_t, t)$ for the given initial condition x_0 under the control $y = y^* + ah$. Given our assumption that control is piecewise continuous and the function $f(x_t, y_t, t)$ is continuously differentiable, the solution $\tilde{x}(a, h)$ is (piecewise) continuously differentiable in a at every point in time. In the case that we require a fixed terminal state $x_T = \bar{x}$ we have to restrict the set of feasible deviations to the subset $\mathcal{H}^{fts} = \{h \in \mathcal{H} : \tilde{x}_T(a, h) = \bar{x} \forall a\}$, that is we can only permit deviations that yield the required terminal state.⁴⁰ For any feasible deviation path h we define

$$J_h(a) = \int_0^T H(\tilde{x}_t(a, h), y_t^* + ah_t, \lambda_t, t) + \dot{\lambda}_t \tilde{x}_t(a, h) dt \\ + \lambda_T \tilde{x}_T(a, h) + \lambda_0 x_0 .$$

Now we are in a position to analyze small deviations from the optimal control path without worrying about a Lebesgue zero measure or the dependence between the stock and the control. All is taken care of if we take the derivative of $J_h(a)$ with respect to a for a feasible h . To grasp the meaning of the optimality condition $J'_h(a) = 0$ it might be helpful to expand $J_h(a)$ around $a = 0$. Abbreviating terms of higher order in a by the Landau symbol $o(a)$

³⁸In case we allow for (only) piecewise continuous controls we take piecewise continuous deviation paths here.

³⁹You might realize that the parameter a does not extend the class of deviations that we take into consideration. However, the parameter a will be used to making the welfare changes caused by these deviations better tractable.

⁴⁰By requiring that the terminal condition is met for all a our requirement is slightly stronger, but the deviations are still sufficiently general to yield the sought for necessary conditions.

we find

$$\begin{aligned}
J_h(a) &= J_h(0) + J'_h(0) a + o(a) \\
&= \int_0^T H(\tilde{x}_t(0, h), y_t^*, \lambda_t, t) + \dot{\lambda}_t \tilde{x}_t(0, h) dt + \lambda_T \tilde{x}_T(0, h) + \lambda_0 x_0 \\
&\quad + \int_0^T \frac{\partial H(\tilde{x}_t(0, h), y_t^*, \lambda_t, t)}{\partial x_t} \frac{\partial \tilde{x}_t(a, h)}{\partial a} \Big|_{a=0} a \\
&\quad + \frac{\partial H(\tilde{x}_t(0, h), y_t^*, \lambda_t, t)}{\partial y_t} h_t a + \dot{\lambda}_t \frac{\partial \tilde{x}_t(a, h)}{\partial a} \Big|_{a=0} a dt \\
&\quad + \lambda_T \frac{\partial \tilde{x}_T(a, h)}{\partial a} \Big|_{a=0} a \\
&\quad + o(a) ,
\end{aligned}$$

By assumption, $a = 0$ maximizes $J_h(\cdot)$. This $J_h(0)$ term corresponds to the second line in the above equations. The third to fifth line therefore have to be zero for all h_t . Otherwise, a small plus or minus a deviation would increase $J_h(\cdot)$. Now comes a neat trick. Up to the current step λ_t was an arbitrary function. Now we choose λ_t so that it satisfies the differential equation⁴¹

$$\frac{\partial H(x_t^*, y_t^*, \lambda_t, t)}{\partial x_t} = -\dot{\lambda}_t , \tag{7.9}$$

where $x_t^* = \tilde{x}_t(0, h)$ is the optimal level of the stock in t . For such a choice of λ the two terms under the integral (in line three and four) that depend on the deviation of the stock cancel each other for all deviation paths h (and for all a). Having chosen λ in this way the necessary condition for an optimum requires that

$$\int_0^T \frac{\partial H(x_t^*, y_t^*, \lambda_t, t)}{\partial y_t} h_t dt + \lambda_T \frac{\partial \tilde{x}_T(a, h)}{\partial a} \Big|_{a=0} = 0 .$$

⁴¹Note that it is sufficient to make the integrand zero at all but a finite number of points. Thus, in case we permit for only piecewise continuous controls and the derivative of the Hamiltonian can jump, we are still good in terms of satisfying the requirement that the integral vanishes, even if the equation below is not satisfied at the times where the control jumps.

The equality has to hold for all deviations $h \in \mathcal{H}$. By assumption $\frac{\partial H(x_t^*, y_t^*, \lambda_t, t)}{\partial y_t}$ is continuous and h_t an arbitrary continuous function, then

$$\frac{\partial H(x_t^*, y_t^*, \lambda_t, t)}{\partial y_t} = 0 \quad (7.10)$$

has to hold for all $t \in [0, T]$.⁴² Moreover, in the case that our terminal state is fixed we know that $\tilde{x}_T(a, h) = \bar{x}$ cannot vary and $\frac{\partial \tilde{x}_T(a, h)}{\partial a} \Big|_{a=0} = 0$. Therefore, the second term in equation (7.10) is zero independently of λ_T . However, if the terminal state is free, then also there will exist some deviation h such that $\frac{\partial \tilde{x}_T(a, h)}{\partial a} \neq 0$, and we obtain another necessary condition stating that

$$\lambda_T = 0 .$$

The intuition for this finding is straight forward. If the stock still had a positive shadow value, we should have depleted it even more. If the stock has a negative value (e.g. in the case of emissions), we should have emitted even more (because we would not care about the future beyond T). Collecting conditions we found that together equations (7.1 i), (7.9), and (7.10) form a set of necessary conditions for a maximum of the optimization problem (7.1). Note that also the equation of motion (7.1 i) can be written as a condition on the Hamiltonian by requiring

$$\frac{\partial H(x_t^*, y_t^*, \lambda_t, t)}{\partial \lambda_t} = \dot{x}_t .$$

In addition, we have either the boundary condition $x_T = \bar{x}$ or the transversality condition $\lambda_T = 0$ if x_T is free. The subsequent section summarizes the results including a broader spectrum of terminal conditions than in our derivation.

7.3 The Maximum Principle: Formal statement

We formulate the maximum principle for general dynamic optimization problem of the form

⁴²A simple way to see this fact in the case of the free terminal state is by choosing $h_t = \frac{\partial H(x_t^*, y_t^*, \lambda_t, t)}{\partial y_t}$ so that the integrand is everywhere (weakly) positive. For a fixed terminal state our deviation path has to satisfy the boundary conditions, so that the derivation of the statement is slightly more intricate.

Problem (DMP):

$$\begin{aligned} & \max \int_0^T U(x_t, y_t, t) \\ \text{s.t. } & \begin{array}{ll} i) & \dot{x}_t = f(x_t, y_t, t) \quad (\text{equation of motion}) \\ ii) & x_0 \text{ given} \quad (\text{initial condition}) \\ \left. \begin{array}{ll} iii.a) & x_T = \bar{x} \quad \text{or} \\ iii.b) & x_T \geq 0 \quad \text{or} \\ iii.c) & x_T \text{ free} \end{array} \right\} & (\text{terminal conditions}) \\ \left. \begin{array}{ll} iv.a) & T \text{ fix} \quad \text{or} \\ iv.b) & T \text{ free} \end{array} \right\} \end{array} \end{aligned}$$

Case *iii*) describes whether there are restrictions with respect to the terminal state of the system. Case *iv*) describes whether the terminal time is fixed or free (in *iv.b* the planning horizon T has to be chosen optimally).

Assumption 1: The functions U and f are continuously differentiable.

Definition (Hamiltonian): Let $H(x_t, y_t, \lambda_t, t) = U(x_t, y_t, t) + \lambda_t f(x_t, y_t, t)$

Proposition 2: Let $(x_t^*, y_t^*)_{t \in [0, T]}$ solve problem (DMP) with $(y_t^*)_{t \in [0, T]}$ piecewise continuous. Then there exists a continuous function $(\lambda_t)_{t \in [0, T]}$ such that for all $t \in [0, T]$:

- y_t^* maximizes $H(x_t^*, y_t^*, \lambda_t, t)$ for $y \in Y \subset \mathbb{R}$.
- Except at points of discontinuity of y_t^* : $\frac{\partial H(x_t^*, y_t^*, \lambda_t, t)}{\partial x_t} = -\dot{\lambda}_t$.
- The following transversality conditions are satisfied:

case *iii.a*) no condition or

case *iii.b*) $\lambda_T \geq 0$ and $\lambda_T x_T = 0$ or

case *iii.c*) $\lambda_T = 0$ and

case *iv.a*) no condition or

case *iv.b*) $H(x_T^*, y_T^*, \lambda_T, T) = 0$.

λ_t is called the costate variable or shadow value.

Intuitive Formulation: Assume H is strictly concave⁴³ and differentiable in y_t . The necessary conditions for an optimal solution of problem (DMP) are:

$$\frac{\partial H(x_t^*, y_t^*, \lambda_t, t)}{\partial y_t} = 0,$$

$$\frac{\partial H(x_t^*, y_t^*, \lambda_t, t)}{\partial x_t} = -\dot{\lambda}_t,$$

$$\frac{\partial H(x_t^*, y_t^*, \lambda_t, t)}{\partial \lambda_t} = \dot{x}_t$$

plus transversality condition(s).

Remark: x_t and y_t can be vectors. The formulation of problem (DMP) and proposition 2 stays unaltered except for

- $y_t \in Y \subset \mathbb{R}$ becomes $y_t \in Y \subset \mathbb{R}^n$ and
- $\frac{\partial H(x_t^*, y_t^*, \lambda_t, t)}{\partial x_t} = -\dot{\lambda}_t$ becomes $\frac{\partial H(x_t^*, y_t^*, \lambda_t, t)}{\partial x_{it}} = -\dot{\lambda}_{it}$ for all $i \in \{1, \dots, n\}$.

The maximum principle gives us a set of necessary conditions for an optimal solution to our intertemporal optimization problem. In the following we state two sufficient conditions ensuring that the candidate solution is indeed the optimum. We assume that the time horizon is fixed. Mangasarian sufficiency condition (fixed terminal time):

Proposition 3: Given a fixed terminal time (case iv.a), the necessary conditions in proposition 2 are sufficient for an optimal solution to problem (DMP) if $H(x, y, \lambda_t, t)$ is concave in (x, y) for all $t \in [0, T]$, the set of feasible controls is convex, and

$$\frac{\partial H(x_t^*, y_t^*, \lambda_t, t)}{\partial y_t} (y_t^* - \tilde{y}_t) \geq 0$$

⁴³If H is linear maximization is not characterized by the differential condition.

for all feasible control paths \tilde{y} .

If $H(x, y, \lambda_t, t)$ is strictly concave in (x, y) for all $t \in [0, T]$, then $(x_t^*, y_t^*)_{t \in [0, T]}$ is the unique solution to problem (DMP).

A less demanding sufficiency condition only evaluates concavity along the optimized path. Arrow sufficiency condition (fixed terminal time):

Proposition 4: Given a fixed terminal time (case iv.a), the necessary conditions in proposition 2 are sufficient for an optimal solution to problem (DMP) if

$$\hat{H}(x_t, \lambda_t, t) = \max_y H(x, y, \lambda_t, t)$$

exists and is concave in x_t for all t (evaluated for the optimal shadow value path).

If $H(x, y, \lambda_t, t)$ is strictly concave for all x_t and $t \in [0, T]$, then x_t is the unique solution to problem (DMP), but the control does not have to be unique.

7.4 Economic Interpretation

Interpretation of λ , H and the necessary conditions:

λ : Shadow value of the stock.

H : A period's contributions to the overall value/welfare. It is a combination of the current utility (or profit) and the current change in stock value.

$$H(x_t, y_t, \lambda_t, t) = \underbrace{U(x_t, y_t, t)}_{\text{current utility}} + \underbrace{\lambda_t f(x_t, y_t, t)}_{\text{change in stock value}}$$

The change in stock value $\lambda_t f(x_t, y_t, t)$ captures the shadow price valued change in stock ($\lambda_t \dot{x}_t$) caused by the growth (or production) function $f(x_t, y_t, t)$.

$\frac{\partial H}{\partial y_t} = 0$: Cannot increase the current contribution to overall value by increasing or decreasing the control. Or breaking it up:

$$\Leftrightarrow \frac{\partial U}{\partial y_t} = -\lambda_t \frac{\partial f}{\partial y_t}$$

The optimal control y_t has to balance an increase in current welfare (profits) implied by a change in y_t with the resulting decrease in stock value.

$\frac{\partial H}{\partial x_t} = -\dot{\lambda}_t$: The standard interpretation of equation (7.11) is that the shadow price of capital has to depreciate at the rate at which capital contributes to the change in overall value (represented by the Hamiltonian). Recall that the equation is an equilibrium condition. If capital value would decay at a faster rate than its contribution to overall welfare, then you are likely to have over-accumulated capital.

This is probably the hardest first order condition to interpret, so we will try some alternative ways to get at the intuition. If you are happy with the statement above, move on to the next condition. Note that can rewrite the condition as

$$\Rightarrow -\dot{\lambda}_t = \frac{\partial H}{\partial x_t} = \frac{\partial U}{\partial x_t} + \lambda_t \frac{\partial f}{\partial x_t} \quad (7.11)$$

$$\Rightarrow \frac{\partial U}{\partial x_t} + \lambda_t \frac{\partial f}{\partial x_t} + \dot{\lambda}_t = 0 \quad (7.12)$$

The shadow value for capital is caused by future capital being valuable either to derive utility or to maintain a certain stock level. Thus, the value of an additional unit of capital has to be higher at t_0 than at a later point of time t_1 if the unit is productive in the meanwhile ($\frac{\partial H}{\partial x_t} > 0 \forall t \in [t_0, t_1]$) either by contributing to utility, or by further increasing capital through stock production. But then, as the shadow value is higher at t_0 than at t_1 it obviously had to fall.

Another way to think about the condition is as follows. The Hamiltonian expression gives the implied value change caused by a change in the quantity of capital. However, if we think about optimally managing the stock level, we have to take into account that a stock has a value of its own. If the value of capital is decreasing over time, then, along

an optimal path the additional unit of capital must yield a higher immediate return. On the other hand, if capital gains value over time it compensates for a lower return to an additional unit of capital via the Hamiltonian. The following split of the effects might help this intuition:

- First, assume that $\dot{\lambda}_t = 0$. Then, the stock level is optimal if a decrease in stock growth from an additional unit of x_t (weighted by its value) is offset by an increase in current utility.
- Second, assume that U is independent of x_t and that an additional unit of x_t increases stock growth ($\frac{\partial f}{\partial x_t} > 0$) and, thus, future stock value. Then, we would generally like to further increase the stock, unless the value of the stock actually declines over time (e.g. because we approach the end of a finite planning horizon where the stock becomes useless).
- Third, putting all effects together, equation (7.12) spells out that we have to balance the three different reasons why we would like to raise (or lower) the stock. First, there is immediate utility payoff. Second, it causes an increase in the production of new stock (which is valuable). Third, the value of a unit of capital increases. In an optimum, whatever positive contribution there is has to be balanced by negative contributions of the same magnitude.

$\lambda_T = 0$: With a free terminal state, whatever stock remains should be of no value.

$H_T = 0$: If T is optimal, the contribution at T to overall value must be zero (i.e. the sum of current and future profits realized in T must be zero). If it would be positive, we should keep going at least for a little bit, if it is negative we should have stopped earlier.

7.5 Current Value Hamiltonian

Most problems including the stock pollutant example exhibit utility of the form

$$U(x_t, y_t, t) = u(x_t, y_t) \exp[-\delta t] .$$

The same form is achieved if $u(x_t, y_t)$ denotes monetary benefits that are discounted at a constant interest rate δ . In these cases we can simplify the analysis by defining the current value Hamiltonian

$$\begin{aligned} H^c(x_t, y_t, \mu_t, t) &= u(x_t, y_t) + \mu_t f(x_t, y_t, t) \\ &= \exp[\delta t] H(x_t, y_t, \lambda_t, t) \end{aligned}$$

$$\text{where } \mu_t = \lambda_t \exp[\delta t] .$$

Then the necessary conditions for an optimum become

$$\begin{aligned} \frac{\partial H(x_t^*, y_t^*, \lambda_t, t)}{\partial y_t} = 0 &\Rightarrow \frac{\partial H^c(x_t^*, y_t^*, \mu_t, t)}{\partial y_t} = 0 \\ \frac{\partial H(x_t^*, y_t^*, \lambda_t, t)}{\partial \lambda_t} = \dot{x}_t &\Rightarrow \frac{\partial H^c(x_t^*, y_t^*, \mu_t, t)}{\partial \lambda_t} = \dot{x}_t \\ \frac{\partial H(x_t^*, y_t^*, \lambda_t, t)}{\partial x_t} = -\dot{\lambda}_t &= -[\dot{\mu}_t \exp(-\delta t) - \delta \mu_t \exp(-\delta t)] \\ \Rightarrow \frac{\partial H^c(x_t^*, y_t^*, \mu_t, t) \exp[-\delta t]}{\partial x_t} &= (\delta \mu_t - \dot{\mu}_t) \exp[-\delta t] \\ \Rightarrow \frac{\partial H^c(x_t^*, y_t^*, \lambda_t, t)}{\partial x_t} &= \delta \mu_t - \dot{\mu}_t \end{aligned}$$

For a finite time horizon, the transversality conditions stay the same.

7.6 Infinite Time Horizon

For many problems it is not obvious when to crop the planning horizon. Then, we generally assume an infinite time horizon and discount future welfare. Replacing T by ∞ in the maximization problem we have to be aware of some subtleties.

- The objective function can potentially become infinite. Then our problem formulation no longer yields a complete order of the feasible paths. In order to employ the tools presented here, we have to assume (and check) that the maximum (or supremum) in problem (DMP) exists and is finite.

- The transversality condition $\lim_{t \rightarrow \infty} \lambda_t = 0$ is not necessary in the case of a free terminal state.
- Under rather moderate assumptions the transversality condition⁴⁴

$$\lim_{t \rightarrow \infty} H(x_t^*, y_t^*, \lambda_t, t) = 0 .$$

is necessary in the infinite horizon setting. Here H is the present value Hamiltonian. The intuition for this transversality condition is that with an infinite time horizon, there is no fixed terminal time.

For stating a sufficiency condition, let us assume a strictly positive discount rate and the following terminal condition

$$\lim_{t \rightarrow \infty} b_t x_t \geq 0 \text{ for some } b : \mathbb{R}_+ \rightarrow \mathbb{R}_+ \text{ with existing } \lim_{t \rightarrow \infty} b_t < \infty .$$

For $b = 1$ the terminal condition simply states that there is a minimal stock level. We use the current value formulation. Then, the following proposition holds.

Proposition 5: The necessary conditions in proposition 2 [not including the transversality conditions] are sufficient for an optimal solution to problem (DMP) if

$$\hat{H}(x_t, \lambda_t, t) = \max_{y_t} H(x_t, y_t, \lambda_t, t)$$

is concave in x_t for all t (evaluated for the optimal shadow value path) and

$$\begin{aligned} \lim_{t \rightarrow \infty} \exp(-\rho t) H(x_t^*, y_t^*, \mu_t, t) &= 0 \text{ and} \\ \lim_{t \rightarrow \infty} \exp(-\rho t) \mu_t x_t &\geq 0 \end{aligned}$$

are satisfied for all feasible paths (y_t, x_t) .

If, moreover, $H(x, y, \lambda_t, t)$ is strictly concave for all x_t then x_t is the unique solution to problem (DMP).

⁴⁴These conditions use the value function which we introduce in the next section. They basically state that the long-run (current value) value function is differentiable and does not explicitly depend on time. See e.g. Acemoglu theorem 7.12, page 251.

Quite frequently it is simply assumed that we can replace the transversality condition by the assumption that we converge to the steady state, i.e. that things “settle down in the long run”.

7.7 Stock Pollutants: The Euler Equation

Recall the stock pollution problem where we wanted to maximize

$$\max \int_0^{\infty} U(E_t, S_t) \exp(-\delta t) dt$$

subject to the constraint

$$\dot{S}_t = E_t - \beta S_t ,$$

and given an initial stock S_0 . Note that we switched over to an infinite time horizon. We assume that $S_t \geq 0$ and, thus, that $\lim_{t \rightarrow \infty} S_t \geq 0$. We maximize over the path $(E_t)_{t \in (0, \infty)}$ and assume $U_E > 0$, $U_S < 0$, $U_{EE} < 0$ (decreasing marginal utility from consumption that produces emissions), and $U_{SS} < 0$ (implying convex damages $-U_{SS} > 0$). In order to make use of the maximum principle we define the (current value) Hamiltonian

$$H^c(E_t, S_t, \mu_t) = U(E_t, S_t) + \mu_t(E_t - \beta S_t)$$

and find the following necessary conditions:

$$\begin{aligned} \frac{\partial H_c}{\partial E_t} &= U_E + \mu_t \stackrel{!}{=} 0 \\ \Rightarrow \mu_t &= -U_E \end{aligned} \tag{7.13}$$

$$\Rightarrow \dot{\mu}_t = -U_{EE} \dot{E}_t - U_{ES} \dot{S}_t , \tag{7.14}$$

$$\begin{aligned} \frac{\partial H_c}{\partial S_t} &= U_S - \mu_t \beta \stackrel{!}{=} \mu_t \delta - \dot{\mu}_t \\ \Rightarrow \dot{\mu}_t &= \mu_t(\delta + \beta) - U_S . \end{aligned} \tag{7.15}$$

Plugging equation (7.13) and (7.14) into equation (7.15) yields

$$-U_{EE} \dot{E}_t - U_{ES} \dot{S}_t = -U_E(\delta + \beta) - U_S . \tag{7.16}$$

This equation is known as the Euler equation⁴⁵ for our problem. You can always obtain the Euler equation (i.e. also for other optimal control problems) by the method used above:

1. Obtain an equation for the shadow value λ_t (or μ_t) from the optimization with respect to the control.
2. Take the time derivative of this equation to also obtain an equation for the time change of the shadow value $\dot{\lambda}_t$ (or $\dot{\mu}_t$).
3. Plug these equations for λ_t and $\dot{\lambda}_t$ (or μ_t and $\dot{\mu}_t$) into the condition on the derivative of the Hamiltonian with respect to the stock (which differs depending on whether you deal with the present or the current value Hamiltonian).

Recall that in addition to the equations above, we have two more necessary conditions for an optimum. First, the equation of motion, which can also be obtained from requiring

$$\frac{\partial H^c}{\partial \mu_t} = E_t - \beta S_t \stackrel{!}{=} \dot{S}_t . \quad (7.17)$$

Second, a transversality condition has to hold. However, for the moment we will replace it by the assumption that we converge into a steady state in the long run.

In the *steady state* our state and control variables are constant (by definition) implying $\dot{E} = \dot{S} = 0$. Then the Euler equation simplifies to the form

$$U_E = \frac{-U_S}{\delta + \beta} . \quad (7.18)$$

This equation looks quite similar to a condition we would expect in a static model: The marginal benefits from emission U_E should be equal to the

⁴⁵While to economists it is mostly known as the Euler equation, in science it is more frequently referred to as the Euler-Lagrange or simply Lagrange equation. The two mathematicians derived it already in the 1750s. More precisely, the Euler equation is a general first order condition for a dynamic optimization problem (written in a slightly different form but similar to problem DMP). This general condition leads to equation (7.16) when using the particular structure of the objective function and the equation of motion of our stock pollution problem.

marginal damages from emission $-U_S$. However, in equation (7.18) marginal benefits from current emission should rather be equal to $\frac{-U_S}{\delta+\beta}$. The denominator comes in because emission benefits are instantaneous while damages are cumulative in time as we are dealing with a stock pollutant. Assume that we (exogenously⁴⁶) add one unit of emissions to the stock at time $t = 0$, which creates a damage D in that (instantaneous) period. Some part of that emission unit will decay until the next period. The rest will still cause damage in later periods. The fraction of the additional unit remaining in the atmosphere is $\exp[-\beta t]$ because our equation of motion assumes exponential decay⁴⁷. However, as we discount the future we don't care quite as much for the damage caused in the future so that the damage of the remaining fraction of the emission unit at time t only yields the present value damage $D \exp[-\beta t] \exp[-\delta t]$. Integrating the damage caused by that additional emission unit at time $t = 0$ over an infinite time horizon yields

$$\int_0^{\infty} D \exp[-(\beta + \delta)t] dt = \frac{1}{-(\beta + \delta)} D \exp[-(\beta + \delta)t] \Big|_0^{\infty} = \frac{D}{\beta + \delta},$$

explaining the steady state form of the Euler equation (7.18).

In general, the Euler equation yield the following condition on the marginal value of an emission unit

$$U_E = \frac{-U_S + U_{EE}\dot{E}_t + U_{ES}\dot{S}_t}{\delta + \beta}. \quad (7.19)$$

In addition to the damage $\frac{-U_S}{\beta+\delta}$ there are two more values that our current benefit from emitting another unit must make up for in order for it to be optimal. First, there is the term $U_{EE}\dot{E}_t$. We assumed $U_{EE} < 0$ implying decreasing marginal benefits from emitting (and using the emission flow for consumption purposes). Assume that emissions are increasing over time ($\dot{E}_t > 0$). Then, the decreasing marginal benefits from emissions imply that a unit emission tomorrow is less valuable than today. Hence, as opposed to the steady state, we rather emit a little more today, or, in value terms:

⁴⁶We undertake the thought-experiment of adding one marginal unit to the steady state stock. In this thought-experiment we neglect that it is infeasible in our model to add any finite amount instantaneously to the stock, or that changing the stock while at the steady state level would actually throw us out of the steady state.

⁴⁷In the steady state we have $0 = \dot{S}_t = -\beta S_t + E_t \Rightarrow \dot{S}_t + \Delta \dot{S}_t = -\beta[S_t + \Delta S_t] + E_t \Rightarrow \Delta \dot{S}_t = -\beta \Delta S_t \Rightarrow \Delta S_t = \exp[-\beta t]$.

The marginal value we derive from another emitted unit today does not have to be as high as in a steady state ($U_{EE}\dot{E}_t$ is negative). Second, there is the term $U_{ES}\dot{S}_t$. An assumption on this mixed derivative is not as straight forward as it was for U_{EE} . Let us assume $U_{ES} > 0$. The interpretation of this assumption is that at a high emission stock, our welfare is even more sensitive to creating emissions for consumption purposes. A story could be that because it is so hot we appreciate AC even more. Assume that also the emission stock is increasing over time ($\dot{S}_t > 0$). Then the value of a marginal unit of emissions increases in the future. In consequence, as compared to the steady state, the value of current emissions has to be higher in order to make it optimal to emit the unit already today as opposed to the future.⁴⁸

7.8 Stock Pollutants: Phase Space Analysis

Now we would like to analyze the optimal dynamics for our model of stock pollution. In general, a full analytic solution cannot be achieved. Therefore, we introduce a graphical analysis frequently applied in economic models of optimal control. Aiming at a helpful diagram in the $S_t - E_t$ plain, we first search for those curves where either the stock, or the optimal emissions do not change over time. To simplify our analysis we assume in the following that $U_{SE} = 0$. Observe that this assumption yields ‘another’ helpful simplification. The condition also implies that the first derivatives only depend on one of the two variables and we can write them as $U_E = U_E(E_t)$ and $U_S = U_S(S_t)$.⁴⁹ Moreover, we continue assuming $U_E > 0, U_S < 0, U_{EE} < 0$ and $U_{SS} < 0$. The following is a useful step by step procedure to obtain the desired diagram used for the so called phase space analysis. A phase space is simply a space pinning down a set of variables that characterizes the dynamics of the system.

1. Draw a coordinate system with S_t on the axis of abscissa (x-axis) and E_t on the ordinate (y-axis).

⁴⁸Similarly, under the assumptions $U_{ES} < 0$ and $\dot{S}_t > 0$ the value of future emissions decreases and current emissions become efficient already at a lower value. A story underpinning $U_{ES} < 0$ would be that once the stock and the implied damages are high enough we no longer can enjoy our former consumption habits (creating emissions) as much as we used to.

⁴⁹From $U_{SE} = 0$ we know that the derivative of U_E with respect to S is zero, implying that U_E is independent of S . Similarly U_S has to be independent of E .

2. Where is $\dot{S}_t = 0$?

From the equation of motion (7.17) we have $\dot{S}_t = 0 \Leftrightarrow E_t = \beta S_t$.
Draw this curve (here a straight line) into the diagram.

3. What happens to the right of the $\dot{S}_t = 0$ line?

Again by equation (7.17) we have $\dot{S}_t = E_t - \beta S_t < 0$

Indicate that S_t is falling in this region by means of corresponding arrows.

Remark: A more formal and more mechanical way to obtain this result is as follows. Define $L(E_t, S_t) = E_t - \beta S_t$ so that by equation (7.17) we have $\dot{S}_t = L(E_t, S_t)$. Then take the derivative $\frac{\partial L(E_t, S_t)}{\partial S_t}$ and evaluate it on the $\dot{S}_t = 0$ -line:

$$\frac{\partial L(E_t, S_t)}{\partial S_t} = -\beta \quad \Rightarrow \quad \left. \frac{\partial L(E_t, S_t)}{\partial S_t} \right|_{\dot{S}_t=0} = -\beta < 0$$

We know that as we move right \dot{S} increases and becomes positive.

4. What happens to the left of the $\dot{S}_t = 0$ line?

By equation (7.17) we have $\dot{S}_t = E_t - \beta S_t > 0$.

Indicate that S_t is increasing in this region by corresponding arrows.

5. Where is $\dot{E}_t = 0$?

From the Euler equation (7.19) we have

$$U_E(E_t) = \frac{-U_S(S_t) + 0 + 0}{\delta + \beta} . \quad (7.20)$$

Implicit differentiation⁵⁰, taking E_t as a function of S_t , yields

$$\begin{aligned} U_{EE}(E_t) \frac{dE_t}{dS_t} &= \frac{-U_{SS}(S_t)}{\delta + \beta} \\ \Rightarrow \frac{dE_t}{dS_t} &= \frac{-U_{SS}}{(\delta + \beta)U_{EE}} < 0. \end{aligned}$$

Draw a falling line representing this curve into the diagram.

6. What happens to the right of the $\dot{E}_t = 0$ line?
By the Euler equation in form (7.16) we have

$$\dot{E}_t = \frac{U_E(E)(\delta + \beta) + U_S(S)}{U_{EE}(E)} > 0,$$

where we used that going right from the \dot{E}_t line the stock S increases and therefore U_S decreases ($U_{SS} < 0$) while the other terms depending only on E stay unaltered. As $U_{EE} < 0$ the right hand side increases and \dot{E} becomes positive.

Indicate that E_t is increasing in this region by corresponding arrows.

Remark: Here the more formal and mechanical way pays off higher than in the above case where we confronted a linear equation. Again, define $G(E_t, S_t) = \frac{U_E(E_t)(\delta + \beta) + U_S(S_t)}{U_{EE}(E_t)}$ so that by equation (7.16) we have $\dot{E}_t = G(E_t, S_t)$. Then take the derivative $\frac{\partial G(E_t, S_t)}{\partial S_t}$ and evaluate it on the $\dot{E}_t = 0$ -line:

$$\begin{aligned} \frac{\partial G(E_t, S_t)}{\partial S_t} &= \frac{U_{SS}(S_t)}{U_{EE}(E_t)} \\ \Rightarrow \frac{\partial G(E_t, S_t)}{\partial S_t} \Big|_{\dot{E}_t=0} &= \frac{U_{SS}(S_t)}{U_{EE}(E_t)} \Big|_{\dot{E}_t=0} > 0 \end{aligned} \quad (7.21)$$

We know that as we move right \dot{E} increases and becomes positive.

⁵⁰There are several ways to do this. One way is to take a total derivative of both sides of equation (7.20) and solve for $\frac{dE_t}{dS_t}$. Another way is to define $E_t = g(S_t)$ and replace E_t by the function $g(S_t)$ in equation (7.20). Then derive both sides of the equation by S_t and solve for $g'(S_t)$. Finally, you can arrange equation (7.20) to the form $h(E_t, S_t) \equiv (\delta + \beta)U_E(E_t) - U_S(S_t) = 0$ and imply the implicit function theorem to obtain the same result (the non-degeneracy condition is simply $U_{EE} < 0$).

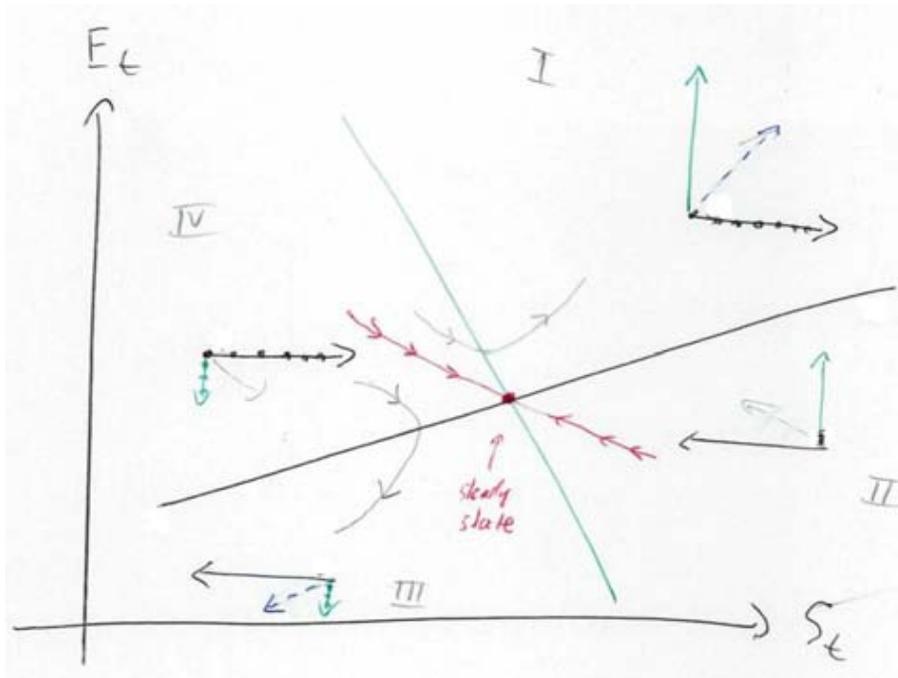


Figure 6 Phase diagram for the stock pollution problem

7. What happens to the left of the $\dot{E}_t = 0$ line?

By equation (7.16) we have

$$\dot{E}_t = \frac{U_E(E)(\delta + \beta) + U_S(S)}{U_{EE}(E)} < 0$$

by an analogous reasoning as in the previous step.

Indicate that E_t is falling in this region by corresponding arrows.

As you probably realized, the arrows flip whenever switching the side of the corresponding curve. Now label the quadrants starting with the top quadrant and proceeding clockwise by *I*, *II*, *III*, and *IV*. From your arrows you observe the following dynamics

- Where the $\dot{E} = 0$ -line and the $\dot{S} = 0$ -line cross the system is in the steady state.
- In quadrants *I* and *III* we are moving away from the steady state.

- In quadrants *II* and *IV* there is a particular path leading into the steady state. Each of these paths is called a separatrix and a stable manifold.
- Given an initial stock S_0 we can pick the control E_0 in exactly one way to get onto the stable manifold. The arrows indicate the movement along which the equation of motion and the Euler equation are satisfied. If we keep picking the control E_t so that we stay on the separatrix we move the system into the steady state. Moreover, we know that along the separatrix the equation of motion and the Euler equation are satisfied.⁵¹

As we converge to the steady state the current value quantities converge to a constant. Thus, it is straight forward that also the necessary transversality condition for the Hamiltonian is met. Moreover, also the condition $\lim_{t \rightarrow \infty} \exp[-\delta t] S_t \mu_t$ would be met even though it is not always necessary.

Let us have another look at the sufficiency condition for our trajectory being the optimum. The concavity of the Hamiltonian in control and state boils down to concavity of the utility function because of the linearity of the equation of motion. Using Sylvester's criterion⁵² for the negative definiteness of the corresponding Hessian leads to the conditions $U_{EE} < 0$ and $U_{EE}U_{SS} - U_{ES}^2 = U_{EE}U_{SS} > 0$. These conditions are satisfied. Thus we are left to show that

$$\begin{aligned} \lim_{t \rightarrow \infty} \exp(-\rho t) \mu_t S_t &= \lim_{t \rightarrow \infty} \exp(-\rho t) [-U_E(E_t)] S_t \geq 0 \\ \Leftrightarrow \lim_{t \rightarrow \infty} \exp(-\rho t) U_E(E_t) S_t &\leq 0 \end{aligned} \quad (7.22)$$

holds for all feasible paths. For this purpose we need another assumption. For example, we can assume that the control has an upper bound \bar{E} and that $U_E(0)$ is finite. Then the long-run limit of the pollution stock has the upper

⁵¹Of course, for our qualitative derivation of the curves and arrows, we only know that somewhere in the according quadrant there must be such a separatrix on which the arrows indicating the movement corresponding to the equation of motion and the Euler equation indeed lead into the steady state.

⁵²Negative definiteness means that both Eigenvalues are strictly negative. Sylvester's criterion verifies negative definiteness by checking that the first leading minor is negative and subsequent leading minors alternate signs. The leading minors of a matrix are the determinants of the upper left $k \times k$ submatrices for $k = 1, 2, \dots$

limit $\lim_{t \rightarrow \infty} S_t \leq \bar{S} = \frac{\bar{E}}{\beta}$. Moreover, the function $U_E(E_t)$ is bounded and equation (7.22) holds with equality for all feasible paths.

7.9 Stock Pollutants: Local Dynamics in the Neighborhood of the Steady State

We maintain the assumptions of the preceding section. Thus, the Euler equation (7.16) and the equation of motion (7.17) are

$$\begin{aligned} -U_{EE}(E)\dot{E}_t &= -(\delta + \beta)U_E(E) - U_S(S) \\ \Rightarrow \dot{E}_t &= \frac{(\delta + \beta)U_E(E) + U_S(S)}{U_{EE}(E)} \end{aligned} \quad (7.23)$$

and

$$\dot{S}_t = E_t - \beta S_t . \quad (7.24)$$

In the following we approximate these functions linearly in order to find the approximate behavior of our system close to the steady state. The approximate system of equations will give us enough information to find out whether the steady state is stable and some other qualitative features of the system. We denote the steady state values of the stock and the control by S^* and E^* . In the steady state the above equations imply

$$(\delta + \beta)U_E(E^*) = -U_S(S^*) \quad \text{and} \quad E^* = \beta S^* . \quad (7.25)$$

In order to approximate equations (7.23) and equation (7.24) we define

$$\begin{aligned} G(E_t, S) &= \frac{(\delta + \beta)U_E(E) + U_S(S)}{U_{EE}(E)} \quad \text{and} \\ L(E_t, S_t) &= E_t - \beta S_t . \end{aligned}$$

The partial derivatives of these functions indicate how \dot{E} and \dot{S} change as we move in the $E_t - S_t$ plane. We find

$$\frac{dG(E, S)}{dE} = \frac{(\delta + \beta)U_{EE}(E)U_{EE}(E) - [(\delta + \beta)U_E(E) + U_S(S)]U_{EEE}(E)}{U_{EE}^2(E)} .$$

Evaluated at the steady state where equations (7.25) are satisfied the squared bracket vanishes leaving us with

$$\left. \frac{dG(E, S)}{dE} \right|_{stead} = \delta + \beta .$$

Similarly we obtain for the first order change in S_t as derived in the remark of the previous section (equation 7.21)

$$\left. \frac{\partial G(E, S)}{\partial S} \right|_{stead} = \frac{U_{SS}(S^*)}{U_{EE}(E^*)}$$

These two equations together give us our linear approximation of the Euler equation (7.23). Denote deviations from the steady state by $\Delta E_t = E_t - E^*$ and $\Delta S_t = S_t - S^*$. Then the approximate Euler equation can be written as

$$\Delta \dot{E}_t = (\delta + \beta) \Delta E_t + \frac{U_{SS}(S^*)}{U_{EE}(E^*)} \Delta S_t , \quad (7.26)$$

where $\Delta \dot{E}_t = \dot{E}_t - \dot{E}^* = \dot{E}_t$.

Similarly we can approximate the equation of motion for the pollution stock employing that

$$\frac{dL(E_t, S_t)}{dE_t} = 1 \quad \text{and} \quad \frac{dL(E_t, S_t)}{dS_t} = -\beta ,$$

which holds everywhere in the $E_t - S_t$ -plane and, thus, in the steady state. We therefore approximate the equation of motion (7.24) for the pollution stock by

$$\Delta \dot{S}_t = \Delta E_t - \beta \Delta S_t . \quad (7.27)$$

Observe that, because the original equation of motion was linear, the linear ‘approximation’ is actually exact (and defining L and taking partial derivatives was not necessary). Together equations (7.26) and (7.27) can be written as the matrix equation

$$\begin{pmatrix} \Delta \dot{E}_t \\ \Delta \dot{S}_t \end{pmatrix} = \begin{pmatrix} \delta + \beta & \frac{U_{SS}(S^*)}{U_{EE}(E^*)} \\ 1 & -\beta \end{pmatrix} \begin{pmatrix} \Delta E_t \\ \Delta S_t \end{pmatrix} . \quad (7.28)$$

Denoting the trace of a matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

by $\text{tr}(A) = a + d$, the determinant by $\det(A) = ad - bc$, and the discriminant by $\Delta(A) = \text{tr}(A)^2 - 4\det(A) = (a + d)^2 - 4(ad - bc)$, recall from linear algebra that the Eigenvalues (or characteristic roots) of A are given by

$$\begin{aligned} \xi_{1,2} &= \frac{1}{2} \left(\text{tr}(A) \pm \sqrt{\Delta(A)} \right) \\ &= \frac{1}{2} \left(a + d \pm \sqrt{(a + d)^2 - 4(ad - bc)} \right) . \end{aligned}$$

Note that the sign of the discriminant $\Delta(A)$ decides whether the Eigenvalues are real or complex. For the matrix characterizing the dynamics of our system of equations (7.28) these quantities translate into

$$\begin{aligned} \text{tr}(A) &= \delta + \beta - \beta = \delta &> 0 \\ \det(A) &= -\beta(\delta + \beta) - \frac{U_{SS}(S^*)}{U_{EE}(E^*)} &< 0 \end{aligned} \quad (7.29)$$

$$\begin{aligned} \Delta(A) &= \delta^2 - 4\det(A) &> \delta^2 \\ &= \delta^2 + 4 \left(\beta(\delta + \beta) + \frac{U_{SS}(S^*)}{U_{EE}(E^*)} \right) \\ \xi_1 &= \frac{1}{2} \left(\delta + \sqrt{\Delta(A)} \right) &> 0 \\ \xi_2 &= \frac{1}{2} \left(\delta - \sqrt{\Delta(A)} \right) &< 0 . \end{aligned} \quad (7.30)$$

Our system of equations determining the dynamics around the steady state has two real Eigenvalues, one positive and one negative. The next subsection 7.10 discusses in detail what these Eigenvalues tell us about the dynamics in the neighborhood of the steady state and why. Here I summarize some of the information most important to us. The fact that we have one positive and one negative Eigenvalue tells us that there exist precisely two separatrices (connecting in the steady state) that lead us into the steady state on a trajectory that satisfies the necessary conditions. Off these separatrices we would move away from the steady state if we were to satisfy the necessary

conditions for an optimum. For these reasons such a steady state is called saddle point stable.

If both Eigenvalues would be real and negative, we would converge from the neighborhood into the steady state from anywhere in the phase space while satisfying the necessary conditions. The steady state would be stable. If both Eigenvalues were real and positive, there would be no separatrix that would carry us from the neighborhood into the steady state while satisfying the equations of motions derived from the necessary conditions for an optimum. The steady state would be unstable. If we had complex Eigenvalues we would circle into or out of the steady state, rather than following a more or less straight line. The real part of the complex Eigenvalue determines whether we circle into the steady state (for a negative real part) or whether we circle out (for a positive real part).

The solutions characterizing the approximate dynamics in the neighborhood of the steady state are of the form

$$\begin{pmatrix} \Delta E_t \\ \Delta S_t \end{pmatrix} = a_1 \vec{v}_1 \exp[\overbrace{\xi_1}^+ t] + a_2 \vec{v}_2 \exp[\overbrace{\xi_2}^- t] \quad (7.31)$$

with $a_1, a_2 \in \mathbb{R}$ and \vec{v}_1 being the Eigenvector corresponding to the Eigenvalue ξ_1 and \vec{v}_2 being the Eigenvector corresponding to the Eigenvalue ξ_2 . From equation (7.31) you can observe why there is precisely one path leading into the steady state. Take the solution where $a_1 = 0, a_2 > 0$. Then the deviations from the steady state ΔE_t and ΔS_t decay to zero as time goes to infinity. However, for all solutions where $a_2 > 0$ we diverge away from the steady state exponentially (with the local approximation quickly losing validity). Thus, the direction of the separatrix that you explored in your phase diagram is, in the neighborhood of the steady state, given by (plus-minus) the Eigenvector \vec{v}_2 , which is the Eigenvector to the negative characteristic root that you can calculate from equation (7.30). Denoting the matrix in equation (7.28) by A we know that the Eigenvectors satisfy

$$A\vec{v}_i = \xi_i \vec{v}_i \quad \Rightarrow \quad (A - \xi_i \mathbf{I})\vec{v}_i = 0 ,$$

with \mathbf{I} denoting the unit matrix. Moreover, by the definition of an Eigenvalue we know that the rows of $(A - \xi_i \mathbf{I})$ are linearly dependent so that we can use e.g. the lower row to calculate the elements of the Eigenvector which we

7.10 Dynamics in the Neighborhood of a Steady State: General Remarks 201

denote by ΔE^i and ΔS^i :

$$\Delta E^i - (\beta + \xi_i)\Delta S^i = 0 \quad \Rightarrow \quad \Delta E^i = (\beta + \xi_i)\Delta S^i .$$

For the negative Eigenvalue ξ_2 the latter equation gives us the slope of the separatrix in the neighborhood of the steady state:

$$\begin{aligned} \frac{\Delta E^{i=2}}{\Delta S^{i=2}} &= \beta + \frac{1}{2} \left(\delta - \sqrt{\delta^2 + 4 \left(\beta(\delta + \beta) + \frac{U_{SS}(S^*)}{U_{EE}(E^*)} \right)} \right) \\ &= \beta + \frac{1}{2} \left(\delta - \sqrt{\delta^2 + 2(2\beta)\delta + (2\beta)^2 + 4 \frac{U_{SS}(S^*)}{U_{EE}(E^*)}} \right) \\ &= \beta + \frac{1}{2} \left(\delta - \sqrt{(\delta + 2\beta)^2 + 4 \frac{U_{SS}(S^*)}{U_{EE}(E^*)}} \right) . \end{aligned}$$

7.10 Dynamics in the Neighborhood of a Steady State: General Remarks

If we like to learn about the local dynamics in the neighborhood of a steady state we linearize our first order differential system. Locally, the generally nonlinear system will behave approximately like the linearized one and a system of linear differential equations is easy to solve. This step is also a starting point for numerically calculating the optimal control path of a general dynamic optimization problem. This section derives and generalizes the results already stated in the preceding section and explains why the Eigenvalues provide the information already mentioned.

In general our dynamic system will be of the form

$$\begin{aligned} \dot{z}_1 &= g_1(z_1, \dots, z_n) \\ \dot{z}_2 &= g_2(z_1, \dots, z_n) \\ &\vdots = \vdots \\ \dot{z}_n &= g_n(z_1, \dots, z_n) \end{aligned}$$

We linearize the system around the steady state. Letting z_1^* denote steady

state values and $\zeta_i = z_i - z_i^*$ (zeta) we obtain the form

$$\begin{pmatrix} \dot{\zeta}_1 \\ \dot{\zeta}_2 \\ \vdots \\ \dot{\zeta}_n \end{pmatrix} = \begin{pmatrix} \frac{\partial g_1(z_1^*, \dots, z_n^*)}{\partial z_1} & \frac{\partial g_1(z_1^*, \dots, z_n^*)}{\partial z_2} & \cdots & \frac{\partial g_1(z_1^*, \dots, z_n^*)}{\partial z_n} \\ \frac{\partial g_2(z_1^*, \dots, z_n^*)}{\partial z_1} & \frac{\partial g_2(z_1^*, \dots, z_n^*)}{\partial z_2} & \cdots & \frac{\partial g_2(z_1^*, \dots, z_n^*)}{\partial z_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_n(z_1^*, \dots, z_n^*)}{\partial z_1} & \frac{\partial g_n(z_1^*, \dots, z_n^*)}{\partial z_2} & \cdots & \frac{\partial g_n(z_1^*, \dots, z_n^*)}{\partial z_n} \end{pmatrix} \begin{pmatrix} \zeta_1 \\ \zeta_2 \\ \vdots \\ \zeta_n \end{pmatrix} \quad (7.32)$$

We denoting the Jacobian, i.e. the ‘derivative matrix’, by A . Then, this system of linear first order differential equation can be written as

$$\dot{\zeta} = A\zeta. \quad (7.33)$$

The trial solution $\zeta = v \exp(\lambda t)$ with an arbitrary n -vector v implies

$$\begin{aligned} \lambda v \exp(\lambda t) &= A v \exp(\lambda t), \\ \Leftrightarrow \lambda v &= A v, \end{aligned}$$

and, thus, that $\zeta = v \exp \lambda t$ indeed solves the differential equation for pairs of eigenvalues λ (characteristic roots) and eigenvectors v (characteristic vectors).

Assumption 2: The matrix A is (complex) diagonalizable.

Then the n eigenvectors v_i corresponding to the n eigenvalues λ_n are linearly independent and the general solutions of the linearized differential system is of the form

$$\zeta(t) = \alpha_1 v_1 \exp(\lambda_1 t) + \alpha_2 v_2 \exp(\lambda_2 t) + \dots + \alpha_n v_n \exp(\lambda_n t) \quad (7.34)$$

with $\alpha_1, \alpha_2, \dots, \alpha_n \in \mathbb{C}$.

Another way to reach equation (7.34) is longer but slightly more intuitive. Observe that due to assumption 2 there exists an invertible matrix S such that

$$SAS^{-1} = D$$

where D is a diagonal matrix of the form

$$D = \begin{pmatrix} \lambda_1 & 0 & \dots & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ 0 & 0 & \lambda_3 & 0 & \dots & 0 \\ \vdots & & & \ddots & & \vdots \\ 0 & \dots & \dots & \lambda_{n-1} & 0 \\ 0 & \dots & \dots & 0 & \lambda_n \end{pmatrix}$$

The system

$$\dot{\xi} = D\xi \tag{7.35}$$

has the obvious, linearly independent solutions

$$\mathbf{e}_i \exp(\lambda_i t) \text{ for } i = 1, \dots, n,$$

where \mathbf{e}_i denotes the i th unit vector. But equation (7.35) is equivalent to

$$\begin{aligned} S^{-1}\dot{\xi} &= S^{-1}DS^{-1}\xi \\ \Leftrightarrow S^{-1}\dot{\xi} &= AS^{-1}\xi \end{aligned}$$

so that $S^{-1}\xi$ is a solution to system (7.33) iff ξ is a solution to system (7.35). Moreover, recall from linear algebra that the column vectors of the matrix S^{-1} are the eigenvectors of A . Thus, once more we have that the eigenvectors $v_i \exp(\lambda_i t) = S^{-1}\mathbf{e}_i \exp(\lambda_i t)$ for $i = 1, \dots, n$ span the solution space of system (7.33). Again we obtain the general solution (7.34).

For real eigenvalues (and real coefficients) the interpretation of equation (7.34) is straight forward. If we happen to move along an eigenvector v_i corresponding to a negative eigenvalue λ_i (i.e. all $\alpha_{j \neq i} = 0$ in equation 7.34) we converge into the steady state as $\exp(\lambda_i t)$ falls to zero. If we happen to move along an eigenvector v_j corresponding to a positive eigenvalue λ_j we diverge away from the steady state as $\exp(\lambda_j t)$ grows exponentially (note that the local approximation quickly loses its validity).

Complex eigenvalues (and eigenvectors) have no immediate economic interpretation as economic variables are not moving in a complex plane. However, complex roots of the characteristic polynomial have an immediate implication for the dynamics in the real space. They describe a spiral movement. If

the real part of the eigenvalue is positive we spiral out of the steady state, and if the real part of the eigenvalue is negative we spiral into the steady state. The following reasoning explains this conjecture. First, it can be shown that complex roots always come in conjugate pairs. Let $\lambda = a + ib$ be a complex root of the characteristic polynomial, then so is $\bar{\lambda} = a - ib$. Second, it can be shown that if v is a (generally complex) eigenvector of A with eigenvalue λ then its complex conjugate \bar{v} is an eigenvector to the eigenvalue $\bar{\lambda}$. Then, in particular, the system (7.32) has solutions of the form

$$\zeta(t) = \alpha_1 \underbrace{v \exp(\lambda t)}_{\equiv \nu(t)} + \alpha_2 \underbrace{\bar{v} \exp(\bar{\lambda} t)}_{\equiv \bar{\nu}(t)}$$

with arbitrary coefficients $\alpha_1, \alpha_2 \in \mathbb{C}$. In particular, we can choose $\alpha_1 = \alpha_2 = \frac{1}{2}$ or also $\alpha_1 = -\alpha_2 = \frac{1}{2i}$. These give us the solutions

$$\begin{aligned} \zeta^*(t) &\equiv \frac{\nu(t) + \bar{\nu}(t)}{2} = \operatorname{Re}(v \exp(\lambda t)) \quad \text{and} \\ \zeta^\dagger(t) &\equiv \frac{\nu(t) - \bar{\nu}(t)}{2i} = \operatorname{Im}(v \exp(\lambda t)) \quad . \end{aligned}$$

Both of these solutions are real. If we write the complex Eigenvalue λ as $\lambda = (a + ib)t$ we can use Euler's formula

$$\exp(ix) = \cos(x) + i \sin(x)$$

to rewrite the solutions as

$$\begin{aligned} \zeta^*(t) &= \operatorname{Re}([\operatorname{Re}(v) + i \operatorname{Im}(v)] \exp(at) \exp(ibt)) \\ &= \operatorname{Re}([\operatorname{Re}(v) + i \operatorname{Im}(v)] \exp(at) [\cos(bt) + i \sin(bt)]) \\ &= \exp(at) \operatorname{Re}[\operatorname{Re}(v) \cos(bt) + i \operatorname{Re}(v) \sin(bt) + i \operatorname{Im}(v) \cos(bt) \\ &\quad + i^2 \operatorname{Im}(v) \sin(bt)] \\ &= \exp(at) [\operatorname{Re}(v) \cos(bt) - \operatorname{Im}(v) \sin(bt)] \end{aligned}$$

and similarly

$$\zeta^\dagger(t) = \exp(at) [\operatorname{Re}(v) \sin(bt) - \operatorname{Im}(v) \cos(bt)] \quad .$$

This form of writing the solution shows clearly that the real part of the Eigenvalue, a , informs us whether we converge to the steady state ($a < 0$ implies that the deviation $\zeta(t)$ from the steady state decays to zero) or whether we diverge ($a > 0$). At the same time we observe that imaginary part of the complex Eigenvalue, b , characterizes the oscillation frequency.

7.11 Stock Pollutants: Comparative Statics and Dynamics

This subsection discusses how changes in the exogenous variables affect the steady state and the optimal trajectories. The steady state is characterized by equation (7.23) for $\dot{E}_t = 0$ and equation (7.24) for $\dot{S}_t = 0$. We analyze the effects of first order changes in the parameters by taking the total derivative of this system of equations with respect to ρ , β , and the endogenous variables E and S , which are constants in the steady state. We have already calculated the changes of the equations in E and S when we derived the local approximation of our dynamic system around the steady state. The result is summarized in equation (7.28), which gives us the left hand side of

$$\underbrace{\begin{pmatrix} \delta + \beta & \frac{U_{SS}(S^*)}{U_{EE}(E^*)} \\ 1 & -\beta \end{pmatrix}}_A \begin{pmatrix} dE \\ dS \end{pmatrix} = \begin{pmatrix} -\frac{U_E(E^*)}{U_{EE}(E^*)} \\ S^* \end{pmatrix} d\beta + \begin{pmatrix} -\frac{U_E(E^*)}{U_{EE}(E^*)} \\ 0 \end{pmatrix} d\rho .$$

From equation (7.29) we know that $\det A$ is negative. Using Cramer's rule (or inverting the matrix) we find e.g.

$$\frac{dE}{d\beta} = \frac{\det A_{1,\beta}}{\det A} > 0$$

where

$$\det A_{1,\beta} = \det \begin{pmatrix} -\frac{U_E(E^*)}{U_{EE}(E^*)} & \frac{U_{SS}(S^*)}{U_{EE}(E^*)} \\ S^* & -\beta \end{pmatrix} = \beta \frac{U_E(E^*)}{U_{EE}(E^*)} - S_t \frac{U_{SS}(S^*)}{U_{EE}(E^*)} < 0$$

is the determinant of the matrix obtained from replacing the first column in A by the vector characterizing the changes in β . Similarly we find

$$\frac{dE}{d\rho} = \frac{\det A_{1,\rho}}{\det A} > 0 \quad \text{and}$$

$$\frac{dS}{d\rho} = \frac{\det A_{2,\rho}}{\det A} > 0$$

where

$$\det A_{1,\rho} = \det \begin{pmatrix} -\frac{U_E(E^*)}{U_{EE}(E^*)} & \frac{U_{SS}(S^*)}{U_{EE}(E^*)} \\ 0 & -\beta \end{pmatrix} = \beta \frac{U_E(E^*)}{U_{EE}(E^*)} < 0 \quad \text{and}$$

$$\det A_{2,\rho} = \det \begin{pmatrix} \delta + \beta & -\frac{U_E(E^*)}{U_{EE}(E^*)} \\ 1 & 0 \end{pmatrix} = \frac{U_E(E^*)}{U_{EE}(E^*)} < 0 .$$

Both, emission flow and emission stock fall unambiguously for a lower rate of pure time preference. Moreover, we saw above that the emission flow decreases whenever the pollution stock decays at a lower rate. In determining the pollution stock's reaction to a change in the decay rate the reduced emission flow acts in the opposite directions of the lower decay rate. Indeed, both, a net increase or a net decrease in the pollution stock under a decrease in the decay rate are possible. This is observed from

$$\frac{dS}{d\beta} = \frac{\det A_{2,\beta}}{\det A}$$

where the sign of

$$\det A_{1,\beta} = \det \begin{pmatrix} \delta + \beta & -\frac{U_E(E^*)}{U_{EE}(E^*)} \\ 1 & S^* \end{pmatrix} = (\delta + \beta) S^* + \frac{U_E(E^*)}{U_{EE}(E^*)}$$

is undetermined at the current level of generality. We can transform the determinant employing the steady state equation $E^* = \beta S^*$ and denoting the elasticity of marginal utility with respect to emissions (and thus consumption producing the emissions) by $\eta_{U,E}$ to the form

$$\det A_{1,\beta} = \left(1 + \frac{\delta}{\beta}\right) E^* + \frac{E^*}{\frac{dU_E(E^*)}{dE} \frac{E^*}{U_E(E^*)}} = E^* \left(1 + \frac{\delta}{\beta} - \frac{1}{\eta_{U,E}}\right) .$$

We find that the pollution stock decreases with a lower decay rate whenever

$$\det A_{1,\beta} < 0 \quad \Leftrightarrow \quad 1 + \frac{\delta}{\beta} < \frac{1}{\eta_{U,E}} \quad \Leftrightarrow \quad \left(1 + \frac{\delta}{\beta}\right) \eta_{U,E} < 1 .$$

Under the assumption that $\eta_{U,E}$ is constant,⁵³ we have the following interpretation. A low rate of time preference makes it ‘more likely’ that the pollution stock is reduced. So does a high responsiveness of marginal utility to changes in emissions. A high decay rate also increases the ‘likelihood’ that the pollution stock is reduced with a reduction of the decay rate.

So far we have only analyzed the effects of parameter changes on the steady state. In the following we examine a change of the control along the optimal trajectory under a change in the discount rate. We already know that the steady state value of emissions E^{*new} and of the pollution stock S^{*new} are higher under an increase in the discount rate to $\rho^{new} > \rho$. We would like to show that, for any given pollution stock S , the optimal control $E^{new}(S)$ is higher than the optimal control $E(S)$ under the lower discount rate. The argument runs by contradiction. We already know that in the steady state $E^{*new} > E^*$ and $S^{*new} > S^*$. A look at Figure 6 should convince you that the new steady state lies above the optimal trajectory under the old parameterization. Assume there existed a pollution stock S° such that for the optimal controls it would hold $E^{new}(S^\circ) < E(S^\circ)$. Then the stable manifolds of the old and the new scenario would have to intersect at some point (S^\dagger, E^\dagger) .

1. Assume we are on the left of the new steady state. At the intersection point (S^\dagger, E^\dagger) the optimal trajectory under the old parameterization would have to fall steeper than the optimal trajectory under the new parameterization. Employing equations (7.23) and (7.24) we derive the implication

$$\begin{aligned} \left. \frac{dE}{dS} \right|_{S^\dagger, E^\dagger, \rho^{new}} &= \left. \frac{\dot{E}_t}{\dot{S}_t} \right|_{S^\dagger, E^\dagger, \rho^{new}} = \frac{(\delta^{new} + \beta)U_E(E^\dagger) + U_S(S^\dagger)}{U_{EE}(E^\dagger) (E^\dagger - \beta S^\dagger)} \\ &> \frac{(\delta + \beta)U_E(E^\dagger) + U_S(S^\dagger)}{U_{EE}(E^\dagger) (E^\dagger - \beta S^\dagger)} = \left. \frac{\dot{E}_t}{\dot{S}_t} \right|_{S^\dagger, E^\dagger, \rho} = \left. \frac{dE}{dS} \right|_{S^\dagger, E^\dagger, \rho} . \end{aligned}$$

⁵³Otherwise, $\eta_{U,E}$ is itself a function of β and ρ spoiling the subsequent reasoning.

Another look at Figure 6 should convince you that left of the steady state $\dot{S}_t = E^\dagger - \beta S^\dagger$ is positive so that the above equation implies $\rho^{new} < \rho$, which delivers the contradiction.

2. Assume we are on the right of the new steady state. At the intersection point (S^\dagger, E^\dagger) the optimal trajectory under the new parameterization would have to fall steeper than the optimal trajectory under the old parameterization. Employing equations (7.23) and (7.24) once more we derive the implication

$$\begin{aligned} \left. \frac{dE}{dS} \right|_{S^\dagger, E^\dagger, \rho^{new}} &= \left. \frac{\dot{E}_t}{\dot{S}_t} \right|_{S^\dagger, E^\dagger, \rho^{new}} = \frac{(\delta^{new} + \beta)U_E(E^\dagger) + U_S(S^\dagger)}{U_{EE}(E^\dagger)(E^\dagger - \beta S^\dagger)} \\ &< \frac{(\delta + \beta)U_E(E^\dagger) + U_S(S^\dagger)}{U_{EE}(E^\dagger)(E^\dagger - \beta S^\dagger)} = \left. \frac{\dot{E}_t}{\dot{S}_t} \right|_{S^\dagger, E^\dagger, \rho} = \left. \frac{dE}{dS} \right|_{S^\dagger, E^\dagger, \rho}. \end{aligned}$$

Another look at Figure 6 should convince you that right of the steady state $\dot{S}_t = E^\dagger - \beta S^\dagger$ is negative so that the above equation implies $\rho^{new} < \rho$, which delivers the contradiction.

Hence, an increase in the discount rate increases the optimal emission level at all levels of the pollution stock.

Nonrenewable resources

This chapter is an introduction to the economics of nonrenewable resource extraction. We consider a resource nonrenewable if its renewal time can be considered negligible compared to its typical consumption flow or within human planning horizons. Common examples are geological resources such as oil, coal, gas, and ores, but also old growth natural forests or top soil, which need centuries to recover. The potential for irreversible loss also qualifies biodiversity as a nonrenewable resource. Our analysis focuses on those nonrenewables whose main benefit arises from their consumption. The first chapter on nonrenewable resource use focuses on the optimal management in a social planner setting, paying particular attention to the time path of extraction, the time to full exploitation, and the trajectory of resource prices (or shadow values). The second chapter analyzes extraction in a competitive market and under a monopolistic ownership of the resource and it compares the resulting extraction paths to those of the social optimum. This subsequent chapter also analyzes resource taxation and the impact of stock dependent extraction costs and technological progress on the evolution of resource prices.

8 Nonrenewable Resources: Socially Optimal Extraction

We start by analyzing the optimal use of a nonrenewable resource in a general setting. Because the stock of a nonrenewable resource can only decline, we will eventually run out of the resources. A steady state might not exist. As a result, the terminal conditions play a particularly important role when analyzing the optimal extraction of nonrenewable resources.

8.1 Problem formulation

We let R_t denote the stock of the resource left for future use, and we let q_t denote the resource flow. In this nonrenewable resource context, q_t is commonly referred to as the extraction rate, even if it is merely the resource flow and not truly a rate in the sense of the rate of change of the stock ($-\frac{\dot{R}}{R}$). The

present chapter focuses on some core insights governing optimal extraction patterns. For this purpose, we abstract from an explicit production sector. Instead, we assume the existence of a direct utility function $U(q_t)$ characterizing society's use value of the resource. This utility function captures the stylized facts about how much our society depends on at least a minimal availability of the resource, and how quickly our use value falls with an increase of the resource flow. We note that this utility function represents the net benefit of the resource flow. Possible extraction costs⁵⁴ or flow-related externalities are contained in $U(q_t)$. The social planner's optimal control problem is

$$\begin{aligned} \max_{q_t} \int_0^T U(q_t) e^{-\rho t} dt & \quad (8.1) \\ \text{s.t. } i) \quad \dot{R}_t = -q_t & \\ ii) \quad R_T \geq 0 & \quad (\text{terminal condition}) \\ iii) \quad R_0 = \text{given} & \quad (\text{initial resource stock}) \\ iv) \quad T \text{ free.} & \end{aligned}$$

The equation of motion for nonrenewables is simple, $\dot{R}_t = -q_t$; what is extracted reduces the stock of the resource left for tomorrow. We cannot go negative on a natural resource. Thus, the stock R_t has to be nonnegative at all times. In most cases, it suffices to impose the non-negativity constraint on the last period, where it plays a crucial role for evaluating the terminal condition. In the present section, we will skip over a potentially binding nonnegativity constraint before the end of the time horizon; we address this possibility later in Section 8.7. The complete problem statement furthermore contains the initial stock of the resource R_0 and the time horizon. Here, we assume that the time horizon is endogenous (free). The choice of the utility function will ultimately determine whether we want to fully exploit the resource in finite time. Alternatively, we could stipulate a particular time horizon after which the social planner has no more interest in consuming or preserving the resource, which generally results in a different solution.

⁵⁴The functional form can only capture constant and resource flow dependent extraction costs. See Section (9.2) for an example where we explicitly split utility into costs and benefits to compare it to a market setting. We will discuss stock dependent and time dependent extraction costs (exogenous technological progress) later in Section 9.5.

The current value Hamiltonian associated with the optimal control problem (8.1) is

$$H^C(q_t, \mu_t) = U(q_t) + \mu_t(-q_t).$$

Assuming an interior solution, the resulting necessary conditions for an optimal extraction program are

$$\frac{\partial H^C}{\partial q_t} = U'(q_t) - \mu_t \stackrel{!}{=} 0 \quad \Rightarrow \quad U'(q_t) = \mu_t \quad (8.2)$$

$$\frac{\partial H^C}{\partial R_t} = 0 \stackrel{!}{=} \rho\mu_t - \dot{\mu}_t \quad \Rightarrow \quad \mu_t = \mu_0 e^{\rho t} \quad (8.3)$$

$$\frac{\partial H^C}{\partial \mu_t} = -q_t \stackrel{!}{=} \dot{R}_t.$$

Equation (8.2) states that the marginal utility of extraction has to equal the shadow value of the resource in the ground (measured in utils, the units of the objective). Equation (8.3) states that the resource's shadow value should grow at the discount rate ρ .⁵⁵ Alternatively, we find that the present value shadow price $\lambda_t = e^{-\rho t} \mu_t = \mu_0$ is constant. As usual, the third necessary condition returns the equation of motion.

8.2 Euler equation

We obtain the Euler equation taking a total time differential of equation (8.2), finding

$$\frac{d}{dt} U'(q_t) = \dot{\mu}_t \quad \Rightarrow \quad \dot{\mu}_t = U''(q_t) \dot{q}_t, \quad (8.4)$$

and then plugging the result for $\dot{\mu}_t$ and equation (8.2) for μ_t into the differential version of equation (8.3) delivering the Euler equation

$$U''(q_t) \dot{q}_t = \rho U'(q_t) \quad \Leftrightarrow \quad \dot{q}_t = \rho \frac{U'(q_t)}{U''(q_t)}. \quad (8.5)$$

⁵⁵We have formulated a social welfare objective for resource use, measuring the objective in welfare units or 'utils'. If the welfare is a representative agent's utility, then ρ is the agent's pure rate of time preference. If we measure welfare in real or monetary units, the discount rate would generally correspond to the consumption discount rate.

We now rewrite this equation using the Arrow-Pratt measure of relative risk aversion, $\text{RRA}(q) = -\frac{U''(q)}{U'(q)}q$. In our context, it is a measure for the decision maker's desire to smooth consumption over time, and the inverse of the elasticity of intertemporal substitution. Then the actual extraction

rate \hat{q}_t is

$$\hat{q}_t \equiv \frac{\dot{q}_t}{q_t} = \frac{-\rho}{-\frac{U''(q_t)}{U'(q_t)}q_t} = -\frac{\rho}{\text{RRA}(q_t)}. \quad (8.6)$$

This rate of change of extraction is negative, implying a declining resource use over time. Optimal resource use declines more quickly for a more impatient society ρ . It falls more slowly, if society has a greater desire to smooth consumption over time (or generations). This Euler equation tells us that high patience as well as a high sense of intergenerational justice both tend to “stabilize” extraction.

The right side version of equation (8.3) suggests that deriving the (usual form of the) Euler equation might not be necessary if our only interest was to solve the problem. Neither the objective function nor the resource's equation of motion depend explicitly on the level of the current resource stock. As a result, the second necessary condition turns into a particularly simple differential equation that is easily integrated. Inserting equation (8.2) straight into the right side version of equation (8.3) delivers

$$U'(q_t) = \mu_0 e^{\rho t} \quad \Leftrightarrow \quad e^{-\rho t} U'(q_t) = \mu_0. \quad (8.7)$$

We learn that, under optimal extraction, marginal utility increases inversely to the discount factor, or, that the discounted marginal utility of a unit of the resource is constant over time. Marginal utility is generally strictly falling in consumption so that we can invert $U'(q_t)$. Denoting the inverse by $U'^{-1}(q_t)$ delivers

$$q_t = U'^{-1}(\mu_0 e^{\rho t}). \quad (8.8)$$

8.3 How to stop?

Neither the Euler equation (8.6) nor equation (8.8) in section 8.2 solve the resource extraction problem by themselves. The Euler equation merely gives

us the speed at which extraction should fall, but not the starting level. Similarly, equation (8.8) still depends on the initial shadow value of the resource, which remains unknown. Alternatively, we can replace the initial shadow value with the initial extraction flow $q_0 = U'^{-1}(\mu_0)$, which is also yet to be determined.

Our necessary conditions involve two ordinary first order differential equation. Their solution has two degrees of freedom. The initial resource stock pins down one of these degrees of freedom. The second degree of freedom manifests in the unknown initial level of the resource flow or, equivalently, in the unknown initial shadow value of the resource. For further progress, we have to make use of the terminal condition. The terminal condition $R_T \geq 0$ corresponds to case *iii.b*) in our formulation of the Maximum Principle. Proposition 2 in Section 7.3 states that this terminal condition imposes the transversality conditions

$$\begin{aligned} \lambda_T &\geq 0 \quad \text{and} \quad \lambda_T R_T = 0 \\ \Leftrightarrow \mu_T &\geq 0 \quad \text{and} \quad e^{-\rho T} \mu_T R_T = 0. \end{aligned} \quad (8.9)$$

Proposition 2 is formulated in present value and we transformed the present value shadow prices $\lambda_t = e^{-\rho t} \mu_t$ into current value shadow prices μ_t . Intuitively, the right side condition states that in the final period, either the resource has to be depleted or it must be of no value. If the resource stock has been depleted, its value must be (weakly) positive.

By equation (8.3), we know that our optimal control problem has a constant present value shadow price $\lambda_t = e^{-\rho t} \mu_t = \mu_0$. Thus, the transversality condition can only be met with $R_T = 0$.⁵⁶ This finding implies that an optimal extraction program exploits all of the resource between the present and the terminal period T . For a given T we obtain the condition

$$\int_0^T q_t dt = \int_0^T U'^{-1}(\mu_0 e^{\rho t}) dt \stackrel{!}{=} R_0, \quad (8.10)$$

which will help us to find the initial extraction rate and the initial shadow value (related by equation 8.2). We used equation (8.8) for the extraction rate.

⁵⁶In the infinite time horizon, the corresponding transversality condition $\lim_{T \rightarrow \infty} R_T = 0$ is not generally a necessary condition for an optimum. Yet, the cases where it fails are somewhat pathological considering economic application.

For a fix time horizon, we would now have collected all the information required to solve the problem, assuming that the necessary conditions uniquely identify the solution. However, our problem features a free (endogenous) final period T and we have to identify this terminal time before we can apply equation (8.10) to pin down the solution. Again, Proposition 2 comes to our rescue. The free terminal time corresponds to case *iv.b*) and the proposition requires the final period's Hamiltonian to be zero

$$\begin{aligned} H(R_T, q_T, \lambda_T, T) = 0 &\Leftrightarrow e^{-\rho T} U(q_T) - \lambda_T q_T = 0 \\ &\Leftrightarrow e^{-\rho T} (U(q_T) - \mu_T q_T) = 0. \end{aligned} \quad (8.11)$$

We recall that the Hamiltonian characterizes period T 's value gain along our optimal extraction program, accounting for both use value and shadow value of the remaining resource stock. If it is optimal to stop, then the value gain from going on for another period has to be zero. The Hamiltonian in Proposition 2 is in present value, and the round bracket's in equation (8.11) contain its present value version. As long as the extraction stops in finite time, we can exchange the present value and the current value Hamiltonian in the terminal condition. However, in case resource use stretches over an infinite time horizon, only the present value Hamiltonian has to approach zero (see Section 7.6). Then, we might not find that $U(q_T) = \mu_T q_T \Leftrightarrow H^c(R_T, q_T, \mu_T, T) = 0$ (in the limit), but condition (8.11) will be met because $\lim_{T \rightarrow \infty} e^{-\rho T} \rightarrow 0$ (and the term in brackets not diverging faster).

Given our finding that the present shadow value $\lambda_t = \mu_0$ is constant in the nonrenewable resource extraction problem, we can write equation (8.11) as⁵⁷

$$q_T = \frac{e^{-\rho T} U(q_T)}{\mu_0}. \quad (8.12)$$

In the case of an infinite extraction horizon, the equation requires (and is satisfied if)

$$\lim_{T \rightarrow \infty} q_T = 0. \quad (8.13)$$

⁵⁷In an economically interesting setting we have $\mu_0 > 0$. Otherwise, equation (8.2) implies $0 = \mu_0 = U'(0)$; already the first unit of extraction would not yield positive marginal utility. As a result the shadow price is always zero and extraction would not take place. The reasonable assumption that $U'(0) > 0$ rules out such a degeneracy.

assuming a well-behaved utility function whose value is bounded given the scarcity of the resource.

Equation (8.11) shows that a necessary condition for optimally exploiting the resource stock in *finite time* is a vanishing *current* value Hamiltonian, i.e., $U(q_T) = \mu_T q_T$ has to have a solution. Thus, assuming resource exploitation in finite time, equations (8.11) and (8.7), $U'(q_t) = \mu_t$, imply that one of the following conditions has to hold⁵⁸

$$1. \quad \mu_T \text{ finite, } q_T = 0 \text{ and } U(0) = 0 \quad \text{or} \quad (8.15)$$

$$2. \quad \mu_T \text{ finite, } q_T > 0 \text{ and } U'(q_T) = \frac{U(q_T)}{q_T}. \quad (8.16)$$

Assuming an increasing and strictly concave utility function, average utility cannot equal marginal utility if $U(0) \geq 0$. In this case, it has to hold that $q_T = 0$ for finite time termination.⁵⁹ If $U(0) < 0$, condition (8.16) can be satisfied. The planner has an incentive to stop resource extraction at a strictly positive flow level before his objective function becomes negative. In fact, for $U(0) < 0$ stopping resource extraction in finite time at a level $q_T > 0$ is the only solution because $U(0) \neq 0$ and equation (8.12) cannot be satisfied as extraction approaches zero or any nonnegative value.

Summarizing our findings, the Euler equation (8.6) shows that the optimally controlled resource flow falls over time. It falls faster the more impatient we are, and it falls slower the more we like to smooth consumption over time. In the nonrenewable resource problem, equation (8.8) shows that we can

⁵⁸We can rule out a third potential case

$$3. \quad q_T = 0, \lim_{t \rightarrow T} \mu_T q_T \rightarrow c \in \mathbb{R}_{++}, U(0) = c \text{ and } \lim_{q \rightarrow 0} U'(q) = \infty. \quad (8.14)$$

Because of the finite extraction time we evaluate the shadow value function on the compact interval $[0, T]$ so that μ_T has to be finite as the Maximum Principle stipulates that the shadow price path is continuous.

⁵⁹We will return to the issue what happens if $U(0) > 0$ in section 8.7. In short, a finite time termination of the planning horizon cannot be optimal. However, the resource might nevertheless be optimally extracted in finite time as the result of the internal boundary condition that we skipped over in the present section. Despite the fully exploited resource, the planner wants to continue the program because $U(0)$ still delivers positive utility and adds to his or her welfare objective. This case is somewhat degenerate and the actual extraction path will not differ from a path where an additive constant normalizes utility to satisfy $U(0) = 0$.

find an integrated version of the Euler equation for general welfare functions, which is as close as we can hope to get in terms of a general closed-form expression. The Euler equation only captures the time change of the extraction flow, but it does not pin down the initial level of extraction. The terminal condition $R_T \geq 0$ implies that we have to use up all of the resource over the optimal time horizon (equation 8.10). At the present level of generality, full exhaustion of the resource can either happen in finite time, or resource use might be stretched over an infinite time horizon. As long as a continuously differentiable utility function is positive, strictly increasing, and concave, resource use will continuously approach zero. If that happens in finite time, we have $q_T = 0$, otherwise resource use will approach zero in the long-run $\lim_{T \rightarrow \infty} q_T = 0$. If utility becomes negative for a low extraction rate, the planner will optimally terminate resource use in finite time at a strictly positive level of the flow.

8.4 Power utility – the eternal stretch

We now analyze the optimal resource extraction problem (8.1) further under the assumption that utility takes the form $U(q_t) = \frac{q_t^{1-\eta}}{1-\eta}$ with $0 < \eta < 1$. The utility function exhibits the constant elasticity of intertemporal substitution η^{-1} , or the coefficient of relative risk aversion $\eta = \text{RRA}(q_t) = -\frac{U''(q_t)}{U'(q_t)}q_t$. We call it a power utility function because it differs by an additive constant from the common form usually referred to as CIES or CRRA (see Section 8.8). In typical intertemporal optimization problems, an affine transformation⁶⁰ of the utility function does not change the solution. However, this reasoning changes when the time horizon is endogenous, as in our case. Then, the zero level of utility becomes meaningful (see Sections 8.6-8.8).

Then Hamiltonian associated with problem (8.1) under power utility is

$$H^c(q_t, \mu_t) = \frac{q_t^{1-\eta}}{1-\eta} - \mu_t q_t$$

⁶⁰An affine transformation multiplies the utility function with a constant and adds another constant. The multiplicative constant has to be the same over time, the additive constant can be time-dependent.

and the first two necessary conditions for an interior solution are

$$\begin{aligned}\frac{\partial H^c}{\partial q_t} &= q_t^{-\eta} - \mu_t \stackrel{!}{=} 0 & \Rightarrow q_t &= \mu_t^{-\frac{1}{\eta}} \\ \frac{\partial H^c}{\partial R_t} &= 0 \stackrel{!}{=} \rho\mu_t - \dot{\mu}_t & \Rightarrow \mu_t &= \mu_0 e^{\rho t}\end{aligned}$$

delivering

$$q_t = \mu_0^{-\frac{1}{\eta}} e^{-\frac{\rho t}{\eta}}. \quad (8.17)$$

In particular, we learn that $q_0 = \mu_0^{-\frac{1}{\eta}}$, illustrating how finding the initial shadow value links to finding the initial level of the extraction rate. Alternatively, we obtain equation (8.17) by integrating the Euler equation (8.6), recognizing that $\text{RRA}(q_t) = \eta$ is constant. As a consequence, the interpretation of the determinants ρ and η governing the extraction flow coincides with our discussion of equation (8.6).

Now we have to find the initial level of the extraction rate (or shadow value). We have already shown that the resource has to be fully extracted by the terminal time (equation 8.10). We still have to determine the terminal time to use equation (8.10) to pin down the initial extraction level. We start by checking the conditions for a possible exploitation in finite time. Equation (8.15) requires that

$$q_T = 0 \text{ and } U(0) = \frac{0^{1-\eta}}{1-\eta} \stackrel{!}{=} 0.$$

The second condition is satisfied (given $\eta < 1$), but the first condition cannot be satisfied because of the exponential decline of the extraction rate derived in equation (8.17). If extraction was zero in the final period it has to be zero in all periods, which we can intuitively rule out as a reasonable solution when extracting a valuable resource. Formally, it would violate $R_T = 0$. The problem can still have a solution with finite time exploitation if condition (8.16) is met

$$q_T > 0 \text{ and } U'(q_T) = q_T^{-\eta} \stackrel{!}{=} \frac{q_t^{1-\eta}}{1-\eta} = \frac{U(q_T)}{q_T} \Rightarrow q_T > 0 \text{ and } 1 - \eta = 1.$$

We excluded the corresponding case $\eta = 0$ early on in our derivation. For $\eta = 0$ “maximizing” the (then linear) current value Hamiltonian delivers

$\mu_t = 1$ as the necessary condition for an interior solution. But we have to keep in mind boundary solutions. For a positive discount rate $\rho > 0$, $\mu_t = 1$ is incompatible with the condition $\mu_t = \mu_0 e^{\rho t}$ on any interval $[0, T]$ of finite length. Intuitively, if $\eta = 0$ consumption is perfectly substitutable across periods. If the decision maker has a positive discount rate, there is no reason to postpone any extraction. She will extract everything in the first instant. Indeed, the solution to this problem is degenerate with immediate extraction.⁶¹

We are left to test whether the necessary conditions can be satisfied with an infinite extraction horizon. By equation (8.17), the extraction rate converges to zero and satisfies the terminal condition (8.13), $\lim_{T \rightarrow \infty} q_T = 0$. Then $\lim_{T \rightarrow \infty} R_T = 0$ delivers equation (8.10) which translates into

$$\int_0^{\infty} q_t dt = \int_0^{\infty} q_0 e^{-\frac{\rho t}{\eta}} dt = -\frac{\eta}{\rho} q_0 e^{-\frac{\rho t}{\eta}} \Big|_0^{\infty} = \frac{\eta}{\rho} q_0 \stackrel{!}{=} R_0,$$

where we used equation (8.17) and $q_0 = \mu_0^{-\frac{1}{\eta}}$. We find

$$\begin{aligned} \Rightarrow q_0 &= \frac{\rho}{\eta} R_0 & \text{and} & \mu_0 = \left(\frac{\eta}{\rho R_0} \right)^\eta \\ \Rightarrow q_t &= \frac{\rho}{\eta} R_0 e^{-\frac{\rho t}{\eta}} & \text{and} & \mu_t = \left(\frac{\eta}{\rho R_0} \right)^\eta e^{\rho t}, \end{aligned}$$

which solves our extraction problem. The optimal initial resource use increases resource availability R_0 and the discount rate ρ . It falls with the desire to smooth consumption over time η . Correspondingly, the speed at which our optimal resource use declines increases with our discount rate and falls with our desire to smooth consumption over time. The decision maker will not fully exploit the resource in finite time. As in the general formulation, the current value shadow price of the resource μ_t increases exponentially at the discount rate, leaving the present value shadow price λ_t constant.

⁶¹With initial and terminal time $T = 0$ our formulation of the maximum principle is not really equipped to solve the problem. But we observe that with $q_0 > 0$ and $\frac{\partial H^c}{\partial q_0} \geq 0$ as well as $\frac{\partial H^c}{\partial q_t} \leq 0$ and $q_t = 0$ for $t > 0$ we maximize the linear Hamiltonian with an increasing shadow price starting at $\mu_0 = 1$. We also satisfy $U'(q_T) = 1 \stackrel{!}{=} \frac{q_T}{q_T} = \frac{U(q_T)}{q_T}$ as suggested above. The condition $R_T = 0$ cannot be satisfied in zero but only instantly after.

8.5 Constant absolute risk aversion – running out

In the case of constant absolute risk aversion (CARA) utility takes the form $U(q_t) = \frac{1-e^{-aq_t}}{a}$, where $a > 0$ is the measure of absolute Arrow-Pratt risk aversion $a = -\frac{U''(q_t)}{U'(q_t)}$.⁶² The Hamiltonian associated with problem (8.1) becomes

$$H^c(q_t, \mu_t) = \frac{1 - e^{-aq_t}}{a} - \mu_t q_t$$

and the first two necessary conditions for an interior solution are

$$\frac{\partial H^c}{\partial q_t} = e^{-aq_t} - \mu_t \stackrel{!}{=} 0 \quad \Rightarrow \quad \mu_t = e^{-aq_t} \quad (8.18)$$

$$\frac{\partial H^c}{\partial R_t} = 0 \stackrel{!}{=} \rho \mu_t - \dot{\mu}_t \quad \Rightarrow \quad \mu_t = \mu_0 e^{\rho t} \quad (8.19)$$

delivering $\mu_0 = e^{-aq_0} \Leftrightarrow q_0 = -\frac{\log(\mu_0)}{a}$ and

$$\begin{aligned} \Rightarrow e^{-aq_t} &= \mu_0 e^{\rho t} & \Rightarrow & -aq_t = \log(\mu_0) + \rho t \\ \Rightarrow q_t &= -\frac{\log(\mu_0) + \rho t}{a} = q_0 - \frac{\rho}{a}t. \end{aligned} \quad (8.20)$$

Extraction falls linearly over time. It falls faster for a higher discount rate and slower for a stronger desire to smooth consumption over time, here expressed by the absolute measure a . The fact that extraction falls linearly already suggests that the resource will be fully exploited in finite time. We note that we can obtain equation (8.20) directly from our general result, equation (8.8), finding that $q_t = U'^{-1}(\mu_0 \exp(\rho t)) = -\frac{1}{a} \log(\mu_0 \exp(\rho t)) = -\frac{\log(\mu_0)}{a} - \frac{1}{a}\rho t$.

Finite time exploitation requires that equation (8.15) or (8.16) hold. Starting with equation (8.15) we find that

$$q_T = 0 \text{ and } U(0) = \frac{1 - e^0}{a} \stackrel{!}{=} 0,$$

⁶²In the limit $a \rightarrow 0$ we find $\lim_{a \rightarrow 0} \frac{1-e^{-aq_t}}{a} = \lim_{a \rightarrow 0} \frac{1-e^{-aq_t}}{1} = q_t$, where we use L'Hospital's rule observing that both numerator and denominator approach zero in the original expression. Thus for $a \rightarrow 0$ the utility function corresponds to the linear case we briefly discussed as the case $\eta = 0$ in the section 8.4 and, in particular, footnote 61

which is satisfied. The finding $q_T = 0$ implies that by equation (8.20)

$$\begin{aligned} 0 = q_T = q_0 - \frac{\rho}{a}T &\quad \Rightarrow \quad q_0 = \frac{\rho}{a}T \\ \Rightarrow q_t = \frac{\rho}{a}(T - t). &\quad (8.21) \end{aligned}$$

This reformulation of equation (8.20) fleshes out that we run down the resource at a linearly falling extraction rate hitting zero at the terminal time. Note that T is still endogenous and will depend on ρ and a as well. In terms of the shadow value we note that $\mu_T = e^{-a q_T}$ implies that $\mu_T = 1$. We call the terminal price of the resource, when extraction comes to an end, the *choke price*. Here, the choke price is unity. This value is a shadow price measured in utils. If we had a more seriously calibrated model, we could transform it into consumption equivalents dividing by marginal utility.

We still have to find the terminal time T . Using the terminal condition we find by equation (8.10)

$$\begin{aligned} \int_0^T q_t dt &= \int_0^T \frac{\rho}{a}(T - t)dt = \frac{\rho}{a} \left[Tt - \frac{t^2}{2} \right]_0^T = \frac{\rho T^2}{2} \stackrel{!}{=} R_0 \\ \Rightarrow T &= \sqrt{\frac{2aR_0}{\rho}}. \end{aligned}$$

The terminal time increases in the initial stock of the resources. The more of a resource we have, the longer we want to stretch its exploitation time. The terminal time horizon also increases with society's desire to smooth consumption over time expressed by the absolute aversion coefficient a . Finally, the terminal time draws closer for a more impatient society with a higher rate of pure time preference ρ .

Substituting our solution for the terminal time into equation (8.21) we obtain the level of initial extraction and the fully specified solution for the extraction rate

$$q_t = \sqrt{2R_0} \frac{\rho}{a} - \frac{\rho}{a}t. \quad (8.22)$$

Four times a resource stock R_0 doubles initial extraction and the time to full extraction.⁶³ Such a change in the initial resource stock does not affect

⁶³Given the linear reduction in extraction, this result is geometrically intuitive. The area under the extraction curve is proportional to total resource use and this area is also proportional to both initial extraction and extraction time.

the rate at which optimal extraction declines. Four times the discount rate doubles the initial resource use and cuts the time to full exploitation into half. Four times the desire to smooth consumption over time, measured by a , cuts initial resource use into half and doubles the time to full exploitation.

8.6 Constant absolute risk aversion – a variation

In typical intertemporal optimization problems, an additive constant in the utility function does not change the solution. However, this reasoning changes when the time horizon is endogenous. Then the zero level of utility becomes meaningful. If we can add more periods with utility above zero we would like to go on, if additional periods result in negative utility we stop. In the case of constant absolute risk aversion (CARA), where the exploitation time is finite, we therefore obtain a different solution if we add a constant, say $U(q_t) = \frac{1-e^{-aq_t}}{a} + b$. The parameter $a > 0$ still measures absolute Arrow-Pratt risk aversion. This variation of Section 8.5's setting changes the Hamiltonian associated with problem (8.1) to

$$H^c(q_t, \mu_t) = \frac{1 - e^{-aq_t}}{a} + b - \mu_t q_t.$$

Because the constant b drops out in the derivatives of the Hamiltonian, the first order conditions (8.18) and (8.19) do not change, still delivering

$$q_t = -\frac{\log(\mu_0) + \rho t}{a} = q_0 - \frac{\rho}{a}t. \quad (8.23)$$

The extraction rate still falls linearly over time at the same speed determined by pure time preference and the desire to smooth consumption over time.

Finite time exploitation requires that one of the terminal conditions (8.15) or (8.16) holds. Starting with equation (8.15) we find that

$$q_T = 0 \text{ and } U(0) = \frac{1 - e^0}{a} + b \stackrel{!}{=} 0,$$

which is no longer satisfied unless $b = 0$. Equation (8.16) is satisfied if

$$\begin{aligned} q_T > 0 \text{ and } U'(q_T) &= e^{-aq_T} \stackrel{!}{=} \frac{\frac{1-e^{-aq_T}}{a} + b}{q_T} = \frac{U(q_T)}{q_T} \\ \Rightarrow q_T > 0 \text{ and } \left(\frac{1}{a} + b\right) e^{aq_T} - q_T &= \frac{1}{a}. \end{aligned} \quad (8.24)$$

This equation does not have a solution for all, but for at least some parameterizations of the problem. For example, let $a = e^{-1} \approx 0.37$ and $b = 2 - e \approx -0.71$. Then the condition (8.24) turns into

$$(e + 2 - e) e^{\frac{qT}{e}} - qT = e \quad \Rightarrow \quad 2 e^{\frac{qT}{e}} - e = qT \quad \Rightarrow \quad qT = e. \quad (8.25)$$

The finding $qT = e$ implies by equation (8.23) that

$$\begin{aligned} e = qT &= q_0 - \frac{\rho}{a}T & \Rightarrow & \quad q_0 = \frac{\rho}{a}T + e \\ \Rightarrow \quad q_t &= \frac{\rho}{a}(T - t) + e. & & \quad (8.26) \end{aligned}$$

The time horizon T is still endogenous and the reformulation of equation (8.20) reflects that, once again, extraction will fall linearly until the stopping time. In contrast to the earlier setting, at the stopping time extraction will be strictly positive (and then drop instantaneously to zero). The shadow value at the stopping time is $\mu_T = e^{-aqT} = e^{-ae}$, dropping the choke price below unity. Given that we have not changed marginal utility, this comparison to the previous setting is somewhat meaningful, as it would also hold after converting utils into consumption equivalents.

We still have to find the terminal time T . Using the terminal condition we find by equation (8.10)

$$\begin{aligned} \int_0^T q_t dt &= \int_0^T \left(\frac{\rho}{a}(T - t) + e \right) dt = \frac{\rho}{a} \frac{T^2}{2} + eT \stackrel{!}{=} R_0 \\ \Rightarrow \quad T &= \frac{a}{\rho} \left(\sqrt{e^2 + 2\frac{\rho}{a}R_0} - e \right). & & \quad (8.27) \end{aligned}$$

We picked the positive root because the terminal time has to be positive. Comparing to our earlier objective function where $b = 0$, we find that the negative b reduces the time until terminal extraction, which was $\sqrt{\frac{2aR_0}{\rho}} = \frac{a}{\rho} \sqrt{2\frac{\rho}{a}R_0}$ in Section 8.5. The intuition is clear, society no longer considers low rates of extraction worthwhile. Given the same optimal decline of extraction over time, it starts at a higher extraction rate and runs out of the resource earlier. To calculate the initial rate of extraction, we substitute the terminal

time into our solution for the extraction rate in equation (8.26) and find

$$\begin{aligned} q_t &= \frac{\rho}{a} \left(\frac{a}{\rho} \left(\sqrt{e^2 + 2\frac{\rho}{a}R_0} - e \right) - t \right) + e \\ &= \sqrt{e^2 + 2\frac{\rho}{a}R_0} - \frac{\rho}{a}t. \end{aligned} \quad (8.28)$$

Comparing the result to our original CARA setting, where equation (8.22) found $q_t = \sqrt{2R_0\frac{\rho}{a}} - \frac{\rho}{a}t$, we confirm a higher initial extraction rate. We note that initial extraction increases by less than extraction in the terminal period, which is now e rather than 0. With the present preferences, four times the initial resource stock R_0 increases initial extraction and terminal time by less than a factor 2. This reduction in sensitivity compared to the earlier setting also holds for changes in the rate of time preference and the desire to smooth consumption over time.

8.7 More on CARA utility & terminal conditions

In Section 8.6 we discuss a particular solution to the optimal extraction problem under our variation of CARA utility. The example solves the terminal condition (8.24)

$$\left(\frac{1}{a} + b \right) e^{aq_T} - q_T = \frac{1}{a} \quad \Leftrightarrow \quad \frac{1}{a} + b = \frac{1}{a}e^{-aq_T} + q_Te^{-aq_T}. \quad (8.29)$$

We now discuss this equation more generally. First, this equation has no solution for $b > 0$. We confirm this claim as follows. Using the formulation on the right, we observe that the left side of the equality increases strictly in b . By contrast, the right side of the equality falls in q_T

$$\frac{\partial}{\partial q_T} \left(\frac{1}{a}e^{-aq_T} + q_Te^{-aq_T} \right) = -e^{-aq_T} + e^{-aq_T} - aq_Te^{-aq_T} = -aq_Te^{-aq_T} \leq 0,$$

assuming a consumption smoothing preference $a > 0$. From section 8.5 we know that equality (8.29) is satisfied for $b = q_T = 0$. Thus, it cannot have a solution where $b > 0$ and q_T remains nonnegative.

Second, we establish that the terminal condition (8.29) has a solutions only for $-\frac{1}{a} < b \leq 0$. Then b reduces the left side of the equality (8.29) to some

positive value (lower than the solution with $b = 0 = q_T$). Increasing q_T , the right side of the equality falls strictly monotonically to zero, so that equation (8.29) will have a (unique) solution. For $b \leq \frac{1}{a}$ the left side of the equality turns negative (or zero), which cannot be met by the right side (with finite resource use). We conclude that (only) for $-\frac{1}{a} < b \leq 0$ it is optimal to terminate the time horizon T . We note that in the case where $b < -\frac{1}{a}$ utility is always strictly negative independent of the extraction level. The social planner would maximize the objective by not extracting at all, a degenerate solution that our Maximum Principle does not pick up.

What happens for $b > 0$? To understand this case, we have to revisit the first order condition paying more careful attention to the resource constraint. For an interior solution equation (8.18) gave us

$$\frac{\partial H^c}{\partial q_t} = e^{-aq_t} - \mu_t \stackrel{!}{=} 0 \quad \Rightarrow \quad \mu_t = e^{-aq_t}. \quad (8.30)$$

However, if the constraint is binding, then $q_t = 0$ and

$$\frac{\partial H^c}{\partial q_t} = e^{-aq_t} - \mu_t \leq 0 \quad \Rightarrow \quad \mu_t \geq e^{-aq_t}. \quad (8.31)$$

Combined with $\mu_t = \mu_0 e^{\rho t}$ (equation 8.19) and accordingly $\mu_0 = e^{-a q_0}$, we obtain the boundary solution $q_t = 0$ and

$$\Rightarrow e^{-aq_t} \leq \mu_0 e^{\rho t} \quad \Rightarrow \quad q_t \geq q_0 - \frac{\rho}{a}t,$$

which is satisfied once extraction hits zero ($q_t = 0 \geq q_0 - \frac{\rho}{a}t$). Thus, we have zero extraction after $t = T$, also as an interior solution to our optimal control problem. Under standard CARA preferences in Section 8.5, $U(0) = \frac{1-e^0}{a} = 0$. Once extraction is zero, it does not matter whether we keep counting utility beyond T . Even if we were to impose an infinite planning horizon, our solution under CARA utility would remain the same. Instead of stopping the problem at T , we would find that we hit the boundary condition and that $q_t = 0$ is the optimal extraction path for $t > T$. This invariance of the solution to a planning horizon beyond the time where optimal extraction hits zero holds under standard CARA utility because $U(0) \geq 0$ when we hit our constraint $q_t \geq 0$. Under our modification with $b > 0$ and, thus, $U(0) > 0$, the social planner continues to derive positive utility even when extraction is zero. Then, she will not terminate the welfare maximization program even

at the point where society runs out of the resource. However, what happens beyond the point where society runs out of the resource does not change. The fact that we keep counting utility is meaningless as the extraction path does not change. Optimal extraction in the CARA variation problem with $b > 0$ is governed by the same equation (8.22) as in the original CARA setting. The terminal time $\sqrt{\frac{2aR_0}{\rho}}$ is now the endogenous time where extraction hits zero, but no longer the formal stopping time.

Such an invariance no longer holds for a variation of CARA utility that results in $U(0) < 0$. In our example of Section 8.6, $b \approx -0.71$ turns utility negative at a zero extraction rate. For example, there can be a fix cost arising from an environmental externality related to mining or to keeping an oil-pipeline open that renders extraction of quantities below some threshold inefficient. Then, it becomes crucial for the social planner to terminate the extraction before the extraction rate hits zero, as we observed in equations (8.25) and (8.26).

8.8 CRRA utility

We now return the class of utility functions exhibiting constant intertemporal elasticity of substitution (CIES), or constant relative risk aversion (CRRA). Adding the constant $-\frac{1}{1-\eta}$ to Section 8.4's power utility delivers the more common form $U(q_t) = \frac{q_t^{1-\eta} - 1}{1-\eta}$, $\eta > 0$. Adding this constant ensures that the utility function converges to the logarithm for $\eta \rightarrow 1$. The CARA variation already taught us that such an additive constant is not entirely innocent if the planning horizon is finite and endogenous. Now we will see that the additive constant $-\frac{1}{1-\eta}$ can change a solution that stretches the resource to infinity, as in our power function formulation of Section 8.4, into a solution with full exploitation in finite time.

The Hamiltonian associated with problem (8.1) under CRRA utility is

$$H^c(q_t, \mu_t) = \frac{q_t^{1-\eta} - 1}{1-\eta} - \mu_t q_t$$

and the necessary conditions remain the same as in Section 8.4 because the additional constant drops out in the derivatives. Equation (8.17) still describes the Euler equation $q_t = q_0 e^{-\frac{\rho t}{\eta}}$. Yet, we now have a solution with full

exploitation in finite time as the terminal condition (8.16) changes to

$$\begin{aligned} U'(q_T) = q_T^{-\eta} &\stackrel{!}{=} \frac{q_t^{1-\eta} - 1}{1-\eta} = \frac{U(q_T)}{q_T} &\Leftrightarrow & q_T^{1-\eta} \stackrel{!}{=} \frac{q_t^{1-\eta} - 1}{1-\eta} \\ \Leftrightarrow \eta q_T^{1-\eta} = 1 &&\Leftrightarrow & q_T = \eta^{-\frac{1}{1-\eta}}. \end{aligned} \quad (8.32)$$

We note that this solution also holds in the limiting case of $\eta \rightarrow 1$, which gives rise to log-utility and $q_T = e$, the Euler number. The CRRA utility functions start out negative and eventually turn positive for all $\eta > 0$. Evaluating utility at the terminal extraction rate (equation 8.32), we find $U(q_T) = \frac{1}{\eta} > 0$. The social planner stops extraction when utility is still strictly positive. Combing the Euler equation with result (8.32) delivers

$$\eta^{-\frac{1}{1-\eta}} = q_T = q_0 e^{-\frac{\rho T}{\eta}} \quad \Rightarrow \quad q_0 = e^{\frac{\rho T}{\eta}} \eta^{-\frac{1}{1-\eta}}. \quad (8.33)$$

By equation (8.10) we find the terminal extraction time solving

$$\begin{aligned} \int_0^T q_t dt &= \int_0^T q_0 e^{-\frac{\rho t}{\eta}} dt = -\frac{\eta}{\rho} q_0 e^{-\frac{\rho t}{\eta}} \Big|_0^T \stackrel{!}{=} R_0 \\ \Rightarrow \frac{\eta}{\rho} q_0 (1 - e^{-\frac{\rho T}{\eta}}) &= \frac{\eta}{\rho} e^{\frac{\rho T}{\eta}} \eta^{-\frac{1}{1-\eta}} (1 - e^{-\frac{\rho T}{\eta}}) \stackrel{!}{=} R_0 \\ \Rightarrow e^{\frac{\rho T}{\eta}} &= 1 + \frac{\rho}{\eta} R_0 \eta^{\frac{1}{1-\eta}} \\ \Rightarrow T &= \frac{\eta}{\rho} \log \left(1 + \rho \eta^{\frac{\eta}{1-\eta}} R_0 \right). \end{aligned} \quad (8.34)$$

In the CRRA setting, the time of resource exhaustion grows logarithmically in the initial resource stock R_0 .

Having characterized the terminal extraction period, equation (8.33) yields

$$q_0 = \left(1 + \frac{\rho}{\eta} R_0 \eta^{\frac{1}{1-\eta}} \right) \eta^{-\frac{1}{1-\eta}} = \eta^{-\frac{1}{1-\eta}} + \frac{\rho}{\eta} R_0. \quad (8.35)$$

and the fully specified solution for the extraction rate

$$q_t = \left(\eta^{-\frac{1}{1-\eta}} + \frac{\rho}{\eta} R_0 \right) e^{-\frac{\rho t}{\eta}}. \quad (8.36)$$

Extraction starts $\eta^{-\frac{1}{1-\eta}}$ higher than in Section 8.4's power function example, and it ends in finite time precisely when extraction has reached this level

$\eta^{-\frac{1}{1-\eta}}$. Plotting this increase in initial (and terminal) extraction we find that it is decreasing and convex in η . The higher the desire to smooth consumption over time, the lower the initial and the terminal extraction rate. Obviously, this result requires that the time to full exploitation increases in the consumption smoothing parameter. The planner spreads and smoothes consumption over more generations.

9 Nonrenewable Resources: Market-Based Extraction

This chapter analyzes resource extraction in a market setting and compares the resulting intertemporal resource allocations to the social optimum. Our first setting is a representative resource extraction firm that acts as a price taker (competitive market economy). We find that the decentralized market solution yields a socially optimal resource allocation. Geological resources are often state owned and occurrence is sometimes limited to a few regions. Therefore, market power has played an important role in resource extraction, and market power is still being exercised today in several markets. We analyze the implications of market power in a monopolistic setting. We prove a famous saying that the “monopolist is the conservationist’s best friend” and exhausts the resource later than a competitive economy (if extraction happens in finite time). After comparing the different market structures, we extend our setting to account for stock-dependent resource extraction costs as well as exogenous technological progress. Finally, we evaluate a (gross) revenue and a profit tax regarding their abilities to appropriate (part of) a firm’s resource rent without distorting the market.

9.1 Extraction by a Competitive Firm

Constant marginal extraction costs. Instead of maximizing welfare, we now analyze a representative firm’s profit maximization problem. For the moment, we assume constant extraction costs c per unit of the resource extracted. Then, the firm’s dynamic profit maximization problem is

$$\max_{\{q_t\}_{t \in [0, T]}} \int_0^T (p_t - c)q_t e^{-rt} dt \quad \text{s.t.} \quad \dot{R}_t = -q_t, \quad R_0 = \bar{R}, \quad R_T \geq 0, \quad T \text{ free}$$

giving rise to the (current value) Hamiltonian

$$H^c(p_t, q_t, \mu_t) = (p_t - c)q_t + \mu_t(-q_t) = (p_t - \mu_t - c)q_t.$$

This Hamiltonian is linear in the control. As a result, the firm will either be indifferent to the amount it extracts (interior solution) or it will extract everything ($p_t > \mu_t - c$) or nothing ($p_t < \mu_t - c$), a boundary solution. In

most markets of interest, extraction is positive and we do not exhaust the resource in an instant. Thus, we will *assume* an interior solution. Section 9.2 will add a demand side story to the argument why the price mechanism in the market equilibrium will usually imply an interior solution, at least until the resource is fully exhausted over time. We obtain the necessary conditions for an interior solution as

$$\frac{\partial H^C}{\partial q_t} = p_t - q_t - c = 0 \quad \Rightarrow \mu_t = p_t - c \quad (9.1)$$

$$\frac{\partial H^C}{\partial R_t} = r\mu_t - \dot{\mu}_t = 0 \quad \Rightarrow \frac{\dot{\mu}_t}{\mu_t} = r \quad \Rightarrow \mu_t = \mu_0 e^{rt}, \quad \mu_0 \geq 0 \quad (9.2)$$

The resulting Euler is

$$\frac{\frac{d}{dt}(p_t - c)}{p_t - c} = \frac{\dot{p}_t}{p_t - c} = r. \quad (9.3)$$

We refer to the difference of price and extraction costs as the rent of the resource. This resource rent is also called a royalty. Equation (9.3) states that the resource rent rises at the rate of interest. This finding is referred to as “Hotelling’s rule” after the economist and mathematician Harold Hotelling (1895-1973) who published this rule in 1931. In the absence of extraction costs, it states that the price of the resource has to grow at the rate of interest. In the presence of extraction costs, it states that only the net revenue from extracting the resource has to grow at the rate of interest.

We can also conceive Hotelling’s rule as an arbitrage argument. Assume

$$\frac{\frac{d}{dt}(p_t - c)}{p_t - c} > r.$$

Then, the firm is better off not extracting today, but extracting tomorrow because the net revenue from the resource grows faster than the firm’s opportunity cost, borrowing money from the bank (if needed). In case of the reverse inequality, the firm would extract all of the resource today. By extracting everything today and depositing the savings on a bank account it makes larger profits compared to waiting and selling some of the resource tomorrow. Implicitly, also the arbitrage argument assumes an interior solution, i.e., an equilibrium where the resource flow is positive in both periods.

Increasing marginal extraction costs. More generally, we assume that marginal extraction costs increase in the level of the extraction flow. Extracting large amounts in a short period of time is more expensive than doing so over a longer period of time (e.g., more or more expensive machinery will be needed when extracting a lot of the resource in parallel than when emptying a reservoir more slowly over time). We denote absolute extraction costs by $C(q_t)$ and assume that $C'(q_t), C''(q_t) > 0$, where the second condition captures increasing marginal extraction costs. For comparison to the above scenario with constant marginal extraction costs we let $c(q_t) = \frac{C(q_t)}{q_t}$ denote the average cost of extraction. We distinguish two cases. In the first case, we assume that also average costs increase $c'(q_t) > 0$. In the second case, average costs can initially fall before they increase ($c'(0) < 0, c''(q_t) > 0$), which is typical in the case of fix costs. We note that these are “per period fix costs” and not “sunk fix costs” that are only required once at the beginning of the extraction program. Such one time costs would not alter the analysis unless they are so large that extraction should not take place at all. We can think of per period fix costs either as machinery and man power that has to run independently of the extraction level or as an approximation to repeated drilling to tap sequentially into a “continuum” of reservoirs. The firm’s maximization problem becomes

$$\max_{\{q_t\}_{t \in [0, T]}} \int_0^T (p_t - c(q_t)) q_t e^{-rt} dt \quad \text{s.t.} \quad \dot{R}_t = -q_t, \quad R_0 = \bar{R}, \quad R_T \geq 0, \quad T \text{ free.}$$

The Hamiltonian $H^c(p_t, q_t, \mu_t) = (p_t - \mu_t - c(q_t))q_t$ is no longer linear and the firm can set the extraction rate to satisfy the first order conditions for an interior optimum

$$\frac{\partial H^c}{\partial q_t} = p_t - \mu_t - c(q_t) - c'(q_t)q_t = 0 \quad \Rightarrow \quad \mu_t = p_t - c(q_t) - c'(q_t)q_t \quad (9.4)$$

or, using absolute extraction costs instead, $\mu_t = p_t - C'(q_t)$, where we used that $c(q_t) + c'(q_t)q_t = C'(q_t) > 0$. Together with equation (9.2) we now find the Euler equation

$$\frac{\frac{d}{dt}(p_t - C'(q_t))}{p_t - C'(q_t)} = \frac{\dot{p}_t - C''(q_t)\dot{q}_t}{p_t - C'(q_t)} = \frac{\dot{p}_t - (2c'(q_t) + c''(q_t)q_t)\dot{q}_t}{p_t - c(q_t) - c'(q_t)q_t} = r. \quad (9.5)$$

Even though the equation looks more complicated, the interpretation remains the same. The net revenue or resource rent grows at the rate of interest. Here,

the resource rent grows not only because of a possible price increase, but also because of a possible change in the extraction rate over time, which affects the extraction costs. If the extraction rate falls over time, future extraction will be cheaper. The resulting increase in net revenue over time partially offsets the required price increase that would have to prevail in the absence of extraction costs.

Terminal conditions. The firm will stop extraction in *finite time* if the current value Hamiltonian is zero

$$\begin{aligned} H^c(p_T, q_T, \mu_T) &= (p_t - \mu_t - c(q_t))q_t \stackrel{!}{=} 0 \\ \Leftrightarrow \mu_T &= p_T - c(q_T) \text{ or } q_T = 0. \end{aligned} \quad (9.6)$$

By equation (9.4), the first condition is equivalent to

$$\begin{aligned} p_T - c(q_T) - c'(q_T)q_T &= \mu_T = p_T - c(q_T) \\ \Leftrightarrow c'(q_T)q_T &= 0. \end{aligned}$$

In the case of increasing average costs, $c'(q_t) > 0$, it must hold that $q_T = 0$. However, in the case of fix costs, $c'(x)$ cuts zero from below and we can have a solution where $q_T > 0$ and the firm stops operating when the resource flow is at the minimum of the average costs.

In a market setting, we can explore the choke price a little further. Recall that the choke price is that price of the resource where exploitation stops. Let $p(q)$ denote inverse demand; we are interested in the demand during the final period, but for convenience we assume that demand is stationary. In the *absence of fix costs*, we have $c'(q_t) > 0$, $q_T = 0$ and a choke price $p(q_T) = p(0)$. Here, the choke price is simply the maximal willingness to pay for the first unit of the resource. The final period's resource rent is $p(0) - C'(0)$ or, by equation (9.4), this shadow value is $\mu_T = p(0) - c(0)$ because $q_T = 0$. We summarize that without fix costs (continuously increasing average costs) we have the terminal conditions

$$q_T = 0 \quad \text{and} \quad \mu_T = p(0) - c(0). \quad (9.7)$$

In the *presence of fix costs*, we stop extraction at the point $q^* > 0$ where average costs are minimized $c'(q^*) = 0$. At this point average costs equal

marginal extraction costs, $c(q^*) = C'(q^*)$.⁶⁴ Assuming that demand falls in the price, we find a lower choke price $p(q^*) < p(0)$ and a last period resource rent of $p(q^*) - C'(q^*) = p(q^*) - c(q^*)$.

Now let us assume that it is *not* optimal to exhaust the resource in *finite time*. From our necessary first order conditions, equations (9.2) and (9.4), we find

$$p_t - c(q_t) = \mu_0 e^{rt} + c'(q_t)q_t. \quad (9.8)$$

In the case of increasing average costs we know that $c'(q_T)q_T \geq 0$. Then, as $\mu_0 e^{rt}$ grows to infinity, so does $p_t - c(q_t)$. In the case of fix costs, we already know that it cannot be optimal to let the extraction rate approach zero smoothly. Yet, the absence of fix costs is necessary to even consider an infinite horizon for extraction.⁶⁵ Thus, a solution that stretches resource extraction over an infinite time horizon can only exist if the average net revenue from extracting a unit of the resource converges to infinity. In particular, there is no finite choke price. We should think of this case as an approximation to a world where the first unit of a resource is extremely valuable to sustain life or the economy. Of course, we cannot avoid that, eventually, the resource flow approaches zero.

Formally, we derive this rather intuitive insight as follows. If finite time extraction is not optimal, the present value Hamiltonian has to converge to zero and we have the terminal condition

$$\begin{aligned} \lim_{T \rightarrow \infty} e^{-rT} H^c(p_T, q_T, \mu_T) &= \lim_{T \rightarrow \infty} e^{-rT} (p_T - \mu_T - c(q_T))q_T = 0 \\ \Rightarrow \lim_{T \rightarrow \infty} e^{-rT} (p_T - c(q_T))q_T &= \lim_{T \rightarrow \infty} \mu_0 q_T. \end{aligned}$$

Thus, either $\lim_{T \rightarrow \infty} q_T = 0$ or

$$\lim_{T \rightarrow \infty} (p_T - c(q_T)) = \lim_{T \rightarrow \infty} \mu_0 e^{-rT} \Rightarrow \lim_{T \rightarrow \infty} c'(q_T)q_T = 0 \quad (9.9)$$

⁶⁴From $c'(q) = 0$ we obtain the condition that $C'(q)q - C(q) = 0 \Rightarrow C'(q) = \frac{C(q)}{q} = c(q)$. In the static micro theory of the firm, the supply curve would start at this point: below it, the firm makes a loss, noting that for such a static competitive firm it is $p = C'(q)$. By contrast, our firm sets the price equal to marginal costs plus resource rent and, thus, still has strictly positive profits. It stops producing when the marginal profit no longer returns the proper rent (which increases over time).

⁶⁵A non-renewable resource flow cannot stay above a strictly positive number over an infinite time horizon because it would require an infinite supply of the resource. Thus, in the case of fix costs, we do not expect to find an infinite extraction horizon to start with.

by equation (9.8). Let $q^* > 0$ denote again the minimum of the average cost curve that the firm would approach in the long run (if $\lim_{T \rightarrow \infty} q_T > 0$). However, assuming a continuous extraction stream, there exist $\epsilon < q^*$ and a time \bar{T} at which extraction is within an ϵ -envelope of q^* so that we have

$$\int_0^\infty q_t dt \geq \int_{\bar{T}}^\infty q_t dt \geq \int_{\bar{T}}^\infty (q^* - \epsilon) dt = \infty, \quad (9.10)$$

contradicting our assumption that the resource is scarce. Summarizing our result, we find that resource use being stretched over an infinite time horizon requires

$$\lim_{T \rightarrow \infty} q_T = 0 \quad \text{and} \quad \lim_{T \rightarrow \infty} p_t - c(q_t) \rightarrow \infty. \quad (9.11)$$

Profits. In any of these scenarios of finite or infinite time exhaustion, equation (9.8) implies that the firm's profits are strictly positive. Why? Because the amount of the resource is fix and it is scarce. Our firm is selling something that cannot be reproduced. How does our result fit into the typical competitive market setting where firms cannot make positive profits? We should think of the resource as a factor of production. Just like capital or labor the resource is paid a rent, which is precisely our resource rent or royalty. In the standard general equilibrium setting, the firm would not be allowed to own the resource, and it would pay the resource owner a rent. Then, the firm itself is back to zero profits and the resource owner gets the rent from owning a scarce commodity.

9.2 Competitive Market Equilibrium & Comparison to Social Optimum

Take 1: Money metric utility generating exogenous demand. Assume that utility is money-metric and $\rho^{social} = r^{market}$. In this partial analysis of the resource market, the representative agent's utility can be expressed as

$$U(q_t) = \int_0^{q_t} p(q) dq \quad \Rightarrow \quad U'(q_t) = p(q_t) = p_t. \quad (9.12)$$

9.2 Competitive Market Equilibrium & Comparison to Social Optimum 234

Keeping in mind that $U'(q_t) = p_t$, we compare the socially optimal and the market's necessary conditions

$$\begin{aligned} H_{social}^C &= U(q_t) - \mu_t q_t & \leftrightarrow & \quad H_{firm}^C = (p_t - \mu_t)q_t - C(q_t) \\ \frac{\partial H_{social}^C}{\partial q_t} &= U'(q_t) - \mu_t = 0 & \leftrightarrow & \quad \frac{\partial H_{firm}^C}{\partial q_t} = p_t - \mu_t - C'(q_t) = 0 \\ \frac{\partial H_{social}^C}{\partial R_t} &= 0 = \rho\mu_t - \dot{\mu}_t & \leftrightarrow & \quad \frac{\partial H_{firm}^C}{\partial R_t} = 0 = r\mu_t - \dot{\mu}_t, \end{aligned}$$

Without extraction costs, $U'(q_t) = p_t$ is enough to see that the necessary dynamic equations of socially optimal extraction and those of a competitive market coincide.⁶⁶ For example, in our CARA utility example from Section 8.5, we find that

$$U(q_t) = \frac{1 - e^{-aq_t}}{a} \quad \Rightarrow \quad U'(q_t) = e^{-aq_t} \equiv p(q_t)$$

so that inverse demand is $p(q) = e^{-aq}$, or actual demand is

$$\Rightarrow \quad q = -\frac{\log(p)}{a} = D(p).$$

Combining the demand equation with the firm's Euler equation (Hotelling's rule) in the absence of extraction costs delivers

$$\Rightarrow \quad q_t = -\frac{\log(p_t)}{a} = -\frac{\log(p_0 \exp(rt))}{a} = -\frac{\log(p_0)}{a} - \frac{r}{a} t,$$

which – up to the initial condition – is the same equation of motion for the extraction rate as in the social optimum.

Without extraction costs, the firm optimally exhausts the resource in *finite time* only if $q_T = 0$.⁶⁷ Then, the representative firm will stop production when $\mu_T = p(0) - c(0)$. Thus, in the absence of extraction costs, we find that

⁶⁶In the market case, the Hamiltonian is once again linear without extraction costs. We rely on the assumption of an interior solution, which now links directly to our finding that the demand side requires $U'(q_t) = p_t$, where usual demand functions make boundary solutions unlikely until the resource is fully exhausted.

⁶⁷The current value Hamiltonian is proportional to $p_t = \mu_t$ and q_t . Only the latter can be zero if

the firm's terminal condition in the competitive market is $\mu_T = p(0) = U'(0)$. Recalling our discussion of equations (8.15) and (8.16) in the social planner setting we know that also the social planner sets $q_T = 0$ in the last period because the corresponding utility function satisfies $U(0) = \int_0^0 p(q) dq = 0$. Thus, the social planner's terminal condition is, by equation (8.2), $\mu_T = U'(0) = p(0)$. We have shown that in the case of finite time exhaustion, the necessary conditions for socially optimal extraction coincide with the necessary conditions generated by a representative firm who faces a corresponding demand function. In the case where it is optimal to stretch resource use over the *infinite horizon*, equation (8.13) in the social planner setting and equation (9.11) for the competitive firm both impose the same terminal condition that $\lim_{T \rightarrow \infty} q_T = 0$.

With extraction costs, we have to integrate the extractions costs into the planner's utility function. Assume utility $U(q_t) = B(q_t) - C(q_t)$ is split between benefits $B(q_t)$ that derive from consuming the resource and (total) extraction costs $C(q_t)$. Only the first of the first order conditions directly depends on extraction costs and it becomes

$$\frac{\partial H_{social}^c}{\partial q_t} = B'(q_t) - C'(q_t) - \mu_t = 0 \quad \leftrightarrow \quad \frac{\partial H_{firm}^c}{\partial q_t} = p_t - \mu_t - C'(q_t) = 0.$$

Again, the dynamic equations in the social optimum coincide with those of the market equilibrium, recognizing that the consumer price only reflects the consumer's marginal benefits $B'(q_t)$.

Assumption 1: We assume the absence of per period fix costs, i.e., $C(0) = 0$. Then we have increasing average costs $c'(q) > 0$ and the extraction flows will continuously approach zero. For the firm, we derived the stopping condition $\mu_T = p(0) - c(0)$. Because $U(0) = B(0) - C(0) = 0$, the social planner's terminal condition, by equation (8.2), equals $\mu_T = U'(0) = B'(0) - C'(0) = p(0) - C'(0) = p(0) - c(0)$, which coincides with the stopping condition of the firm. In the case resource use is stretched to infinity, we find once again $\lim_{T \rightarrow \infty} q_T = 0$ as the coinciding terminal condition in both settings.

In our CARA utility example, we could split the overall welfare into a part that represents consumer benefits and a part that represents extrac-

tion costs⁶⁸

$$\begin{aligned}
 B(q_t) &= \frac{1 - (1 + \gamma)e^{-aq_t}}{a} & \Rightarrow & & B'(q_t) &= (1 + \gamma)e^{-aq_t} \equiv p(q_t) & \text{and} \\
 c(q_t) &= \frac{1 - \gamma e^{-aq_t}}{a} & \Rightarrow & & c'(q_t) &= \gamma e^{-aq_t},
 \end{aligned}$$

not changing the social optimum from Section 8.5 as we keep $U(q_t) = B(q_t) - c(q_t) = \frac{1 - e^{-aq_t}}{a}$ and, in particular, $U'(q_t) = e^{-aq_t}$. The resulting demand side relation between price and quantity is

$$\begin{aligned}
 \text{inverse demand} & & p_t &= (1 + \gamma)e^{-aq_t} \\
 \text{demand} & & q_t &= -\frac{\log\left(\frac{p_t}{1+\gamma}\right)}{a} = D_\gamma(p_t).
 \end{aligned}$$

Recall from equation (9.4) that $\mu_t = p_t - c'(q_t)$ and so the firm sets $p_t = \mu_t + c'(q_t) = \mu_0 \exp(rt) + \gamma e^{-aq_t}$ (integrated form of the Euler equation 9.5). Combining this equation with the consumer side we find

$$\begin{aligned}
 \underbrace{B'(q_t)}_{\text{consumer}} &= (1 + \gamma)e^{-aq_t} = p_t = \underbrace{\mu_t + c'(q_t)}_{\text{firm}} = \mu_0 \exp(rt) + \gamma e^{-aq_t} \\
 \Rightarrow e^{-aq_t} &= \mu_0 \exp(rt) & \Rightarrow & & q_t &= -\frac{\log(p_0)}{a} - \frac{r}{a} t.
 \end{aligned}$$

Again, it is the same equation of motion as in the social optimum. The (coinciding) terminal condition $\mu_T = p(0) - c(0)$ fixes the initial shadow value to $\mu_0 = (p(0) - c(0))e^{-rT}$. To find p_0 we have to integrate the resource use over time as in Section 8.5 or compare the above equation to that section's result, equation (8.22) stating $q_t = \sqrt{2R_0 \frac{r}{a}} - \frac{r}{a} t$. Replacing time preference with market interest, we find

$$p_0 = \exp\left(-a\sqrt{2R_0 \frac{r}{a}}\right) = \exp\left(-\sqrt{2R_0 ar}\right).$$

The initial resource price falls in the resource abundance and the interest rate. Why does the initial price fall for a higher interest rate? Our terminal condition sets the price in the last period. With a higher interest rate, the price increases faster over time and the resource use falls more quickly. Thus,

⁶⁸Given the CARA functional form, the resulting cost function has decreasing marginal costs. The decomposition merely serves as a convenient illustration, not as a good example for a cost function.

we start out with a higher extraction rate at a lower price level. The expression also seems to suggest that the initial price would fall in the consumption smoothing preference. However, the price p_t is measured in utils and as we vary the parameter a we change the utility function and the meaning of a “util”. Thus, we cannot meaningfully interpret the expression for a variation of a .⁶⁹

Assumption 2: We assume $C(0) > 0$. Then, the firm stops resource extraction at the point $q^* > 0$ where average costs are minimized $c'(q^*) = 0$. At this point average costs equal marginal extraction costs, $c(q^*) = C'(q^*)$ or

$$C'(q_T) = \frac{C(q_T)}{q_T}. \quad (9.13)$$

For the social planner, we find that $U(0) = B(0) - C(0) < 0$ and equation (8.16) requires that the social planner’s terminal condition is

$$\begin{aligned} U'(q_T) = \frac{U(q_T)}{q_T} &\Rightarrow B'(q_T) - C'(q_T) = \frac{B(q_T) - C(q_T)}{q_T} \\ &\Rightarrow C'(q_T) = \frac{C(q_T)}{q_T} - \left(\frac{B(q_T)}{q_T} - B'(q_T) \right). \end{aligned} \quad (9.14)$$

In general, the terminal condition (9.13) and (9.14) differ. For a benefit with $B(0) = 0$, $B'(q) > 0$, and $B''(q) < 0$, we have the condition $C'(q_T) < \frac{C(q_T)}{q_T}$

⁶⁹To permit a more meaningful investigation of the initial price, we have to convert it into present consumption equivalents. Note that $U'(q_0)$ characterizes the util-value of the marginal consumption unit in period 0. If we invert this relation it gives us the value of a util in consumption equivalents, here, in terms of units of the resource. We find

$$\tilde{p}_0 = \frac{p_0}{U'(q_0)} = \exp(-\sqrt{2R_0ar})a \left(1 - \exp(-a(\sqrt{2R_0\frac{r}{a}}))\right)^{-1} = a \left(\exp(\sqrt{2R_0ar}) - 1\right)^{-1}$$

A short calculations shows that

$$\frac{d}{da} \left(\exp(\sqrt{2R_0ar}) - 1 \right)^{-1} > 0 \quad \Leftrightarrow \quad \exp(x) \left[1 - \frac{x}{2} \right] - 1 > 0,$$

where $x = \sqrt{2R_0ar}$. Plotting the function, it starts out positive and eventually turns negative (around $x = 1.6$). As the intertemporal consumption smoothing parameter increases, initial extraction falls, and its marginal value increases. Given we divide by marginal value to translate utils into consumption equivalents, we find that eventually the value of a unit of the resource in consumption equivalents falls.

because $\frac{B(q_T)}{q_T} > B'(q_T)$. Then the social planner would terminate at a lower resource flow than the competitive firm.

Take 2: Intertemporally optimizing consumer. Now let us explicitly model the dynamic decision problem of a representative agent. For simplicity, we continue to assume that the consumer is a direct user of the resource, e.g., oil for driving his or her car. Moreover, we assume that the consumer has a quasi-linear utility function $V(x_t, q_t) = x_t + U(q_t)$ where x_t represents the monetary value of all other consumption goods. The assumption that utility is linear in x implies that we are not interested in how the agent's overall well-being affects the resource use.⁷⁰ We are interested in the resource market, a small part of the economy and we merely model how the consumer responds to price changes in the resource everything else equal. The consumer is equipped with initial wealth W_0 , which earns interest over time at rate r . She spends her wealth on the resource and the money-metric consumption bundle reflecting other goods, whose price we normalize to unity. The consumer's optimization problem is

$$\begin{aligned} \max_{q_t, x_t} \int_0^T (x_t + U(q_t)) e^{-\rho t} dt \\ \text{s.t. } \dot{W}_t = rW_t - x_t - p_t q_t & \quad (\text{equation of motion}) \\ W_T \geq 0 & \quad (\text{terminal condition}) \\ W_0 = \text{given} & \quad (\text{initial wealth}). \end{aligned}$$

We left open the terminal condition as it governs her overall expenditures rather than the resource in particular. It might be interesting to realize the similarity between the consumer's consumption expenditure problem ("wealth extraction") and the original resource extraction problem. The main difference is that the wealth re-grows as long as it is not completely exhausted. The consumer's (current value) Hamiltonian is

$$H^C(q_t, \mu_t) = x_t + U(q_t) + \mu_t(rW_t - x_t - p_t q_t).$$

⁷⁰Precisely, we are abstracting from wealth and overall consumption affecting the agent's marginal utility through either of two channels. First, the wealth and resulting overall consumption level impacts marginal utility derived from "other consumption" and affects, in particular, the equilibrium interest rate. Second, changes in "other consumption" can directly interact with the marginal utility derived from resource use. Our quasi-linear utility function allows us to focus on the direct first order partial equilibrium effects that result from saturation in or scarcity of the resource flow.

The resulting necessary conditions for an optimal expenditure program assuming an interior solution are

$$\frac{\partial H^C}{\partial x_t} = 1 - \mu_t \stackrel{!}{=} 0 \quad \Rightarrow \quad \mu_t = 1 \quad (9.15)$$

$$\frac{\partial H^C}{\partial q_t} = U'(q_t) - \mu_t p_t \stackrel{!}{=} 0 \quad \Rightarrow \quad U'(q_t) = p_t \mu_t \quad (9.16)$$

$$\frac{\partial H^C}{\partial W_t} = r\mu_t \stackrel{!}{=} \rho\mu_t - \dot{\mu}_t \quad \Rightarrow \quad \mu_t = \mu_0 e^{(\rho-r)t} \quad (9.17)$$

$$\frac{\partial H^C}{\partial \mu_t} = rW_t - x_t - p_t q_t \stackrel{!}{=} \dot{W}_t$$

Equations (9.15) and (9.17) imply

$$1 = \mu_t = \mu_0 e^{(\rho-r)t} \quad \Rightarrow \quad \rho = r. \quad (9.18)$$

Modeling the household’s dynamic choice problem explicitly implies that market interest r equals the representative agent’s rate of time preference ρ , a finding we simply assumed in our earlier approach. We note that this relation does not always have to hold. It holds here because our quasi-linear utility function implies constant marginal utility from “all other consumption”. Without that assumption, the market interest rate would match a combination of time preference and decreasing (or increasing) marginal utility over time as the representative agent’s consumption level changes.⁷¹

Equation (9.15) reduces equation (9.16) to

$$U'(q_t) = p_t, \quad (9.19)$$

the consumer’s “Euler equation”. Thanks to the major simplification from finding (9.15), this “Euler equation” is just the “static” condition (9.12) that price equals (or more generally is proportional to) marginal utility. Given findings (9.18) and (9.19), the remainder of the comparison between the market equilibrium and the social optimal runs as above.

⁷¹Given we did not explicitly model other goods in our social planner problem, we would have to extend not only the consumer but also the social planner model to show equivalence in a more comprehensive general equilibrium settings.

9.3 Monopolistic resource extraction

Some nonrenewable resources are or have been concentrated in the hand of a few countries that have exercised market power in the past. Maybe the most famous price influencer is the Organization of the Petroleum Exporting Countries (OPEC), founded in 1960 by Iran, Iraq, Kuwait, Saudi Arabia, and Venezuela and today comprising 13 member states. It came to fame during the oil embargo 1973-74 when the member states used their market power to influence several countries supporting Israel by reducing or cutting their oil supply. During the last decade, OPEC has lost some of its market power, but the big players still tried to coordinate a reduction in oil supply during the recent recessions. Closer to an actual monopoly these days is China's share in the extraction of rare earth minerals.⁷² We use the market form of a monopoly rather than an oligopoly or a cartel merely to demonstrate the role of market power in a simple setting. Analytic solutions for monopolistic extraction are harder to obtain than in the case of socially optimal extraction. In the absence of a convincing analytic examples, we embrace a general analysis. Thereby, we also generalize a famous finding stating that the monopolist is the conservationist's best friend.⁷³

A monopolist does not take the price as given. We assume that the monopolist knows the consumer's stationary (inverse) demand function $p(q)$ and sets the extraction rate to maximize profits. The monopolist's optimal control problem is

$$\max \int_0^T (p(q_t) - c(q_t))q_t e^{-rt} dt \quad \text{s.t.} \quad \dot{R} = -q_t, R_0 = \bar{R}, R_t \geq 0, T \text{ free.}$$

We assume a free terminal horizon because we are particularly interested in the question whether a monopolist will fully exhaust the resources earlier or later than a social planner or a price-taking firm. In the present setting, we assume that per unit costs are weakly increasing in extraction and that

⁷²These rare earth minerals are, however, more abundant than the name suggests and if the price increases sufficiently other reservoirs will be worthwhile tapping into.

⁷³Conrad & Clark's (1987) classic text book on resource economics and Hanley, Shogren & White (2001) assume a linear demand. While they can solve the functional form of the resource extraction, they cannot solve for the terminal extraction time that will be of particular interest in our setting. Perman et. al. (2012) use the logarithmic demand curve we encountered in the CARA utility case. They derive an analytic solution but only by using a somewhat crude approximation.

$c'(q_t) \geq 0$ is continuous. We also assume that the demand price is strictly falling in extraction and that $p'(q_t) < 0$ is continuous.

The (current value) Hamiltonian

$$H^c(q_t, \mu_t) = p(q_t)q_t - c(q_t)q_t - \mu_t q_t$$

delivers the first order conditions

$$\begin{aligned} \frac{\partial H^c(q_t, \mu_t)}{\partial q_t} &= p'(q_t)q_t + p(q_t) - c'(q_t)q_t - c(q_t) - \mu_t \stackrel{!}{=} 0 \\ &\Rightarrow p'(q_t)q_t + p(q_t) - c'(q_t)q_t - c(q_t) = \mu_t \end{aligned} \quad (9.20)$$

$$\frac{\partial H^c(q_t, \mu_t)}{\partial R_t} = 0 \stackrel{!}{=} r\mu_t - \dot{\mu}_t \quad \Rightarrow \quad \dot{\mu}_t = r\mu_t \quad (9.21)$$

From equation (9.21) we obtain as usual

$$\dot{\mu}_t = r\mu_t \quad \Rightarrow \quad \mu_t = \mu_0 \exp(rt)$$

which we insert into equation (9.20) to find

$$p_t = c'(q_t)q_t + c(q_t) + \mu_0 \exp(rt) + \underbrace{-p'(q_t)q_t}_{\text{mark up } (\geq 0)} . \quad (9.22)$$

The underbraced expression $-p'(q_t)q_t$ is positive because the demand price falls with quantity. It is this mark up that distinguishes the dynamic equations of the monopoly from those of the competitive firm (compare equation 9.4). and the socially optimal extraction.

Can we use the mark up in equation (9.22) to claim that the resource price is always higher under monopolistic extraction? No. First, the initial shadow value of the resource might differ in the case of the monopoly. Second, in case of a finite extraction time, it is unlikely that a price can always be higher under monopolistic extraction. Assume that the competitive price was always lower. Then consumers in the competitive market would consume more in every period and the resource would run out earlier. Then, the price should at least be higher in the competitive scenario when the resource is running out in the competitive setting but still abundant under the monopoly. We will return to the comparison between monopolistic and competitive extraction in the next section.

We now show that the monopolist's extraction path is *continuous and strictly falling* over time. The function

$$g(q_t) \equiv p(q_t) - c(q_t) + p'(q_t)q_t - c'(q_t)q_t$$

is continuous by our assumption that $p'(q_t)$ and $c'(q_t)$ are continuous (implying that their integrals $p(q_t)$ and $c(q_t)$ are continuous as well). Along the optimal path, q_t maximizes the Hamiltonian at every point in time. Therefore, the second order condition for an optimum implies

$$\frac{\partial^2 H^c(q_t, \mu_t)}{\partial q_t^2} = g'(q_t) \stackrel{!}{\leq} 0. \quad (9.23)$$

Rearranging equation (9.22) we find that

$$g(q_t) = \mu_0 \exp(rt) \quad \Rightarrow \quad g'(q_t)\dot{q}_t = \mu_0 r \exp(rt). \quad (9.24)$$

Assuming $\mu_0 \neq 0$,⁷⁴ we therefore have $g'(q_t) \neq 0$ and

$$\dot{q}_t = \mu_0 r \frac{\exp(rt)}{g'(q_t)} < 0.$$

Extraction falls strictly over time. Moreover, the function $g(q_t)$ is invertible because $g'(q_t) < 0$. By the left side of equation (9.24), we can write the solution to the extraction problem as

$$q_t = g^{-1}(\mu_0 \exp(rt)),$$

establishing that the extraction path is continuous.

Finally, we analyze when the monopolist's terminal condition assuming exhaustion of the resource in finite time. Then, the current value Hamiltonian has to equal zero in the last period

$$\begin{aligned} H^c(q_T, \mu_T) &= p_T(q_T)q_T - c(q_T)q_T - \mu_T q_T \stackrel{!}{=} 0 \\ \Rightarrow p_T(q_T)q_T - c(q_T)q_T &= \mu_T q_T \\ \Rightarrow p_T(q_T) - c(q_T) &= \mu_T \text{ or } q_T = 0. \end{aligned}$$

⁷⁴If μ_0 was zero, then $\mu_t = 0$ at all times and the resource would be of no (shadow-) value in all period. Moreover, by equation (9.24) (l.h.s.) we had $g(q_t) = 0$ for all t and, in particular, $g(q_T) = 0$. But we already know that $q_T = 0$ implying that $g(0) = p(0) - c(0) = 0$ would have to hold ($g(q)$ and $p'(q)$ are continuous and, thus, finite on any compact interval including 0). Even for the first marginal extraction unit the demand value would not exceed the extraction costs, and extraction should not happen.

and by inserting equation (9.20) for μ_T we find that either $q_T = 0$ or $q_T \neq 0$ and

$$\begin{aligned} p_T(q_T) - c(q_T) &\stackrel{!}{=} p'(q_T)q_T + p(q_T) - c(q_T) - c'(q_T)q_T \\ \Leftrightarrow p'(q_T) &= c'(q_T). \end{aligned}$$

But we have assumed increasing average costs $c'(q) \geq 0$ and falling demand $p'(q) < 0$. As a result, the equality cannot be satisfied. We conclude that resource use has to continuously approach $q_T = 0$ if full exploitation happens in finite time. Our reasoning corresponds directly to the one in the social planner setting: given the monopolist's objective function, equation (8.16) cannot be satisfied and, thus, equation (8.15) has to hold. Here, profits replace utility and are zero in the last period when $q_T = 0$. Because $q_T = 0$, equation (9.20) implies that the last period shadow value is $\mu_T = p(0) - c(0)$. Thus, there will be no mark-up in the last period. Summarizing these results from the terminal condition we know that for finite time exhaustion

$$q_T = 0 \quad \text{and} \quad \mu_T = p(0) - c(0), \quad (9.25)$$

which coincide with the corresponding case for the competitive firm, i.e., where $c'(q) \geq 0$ implies $q_T = 0$.

9.4 Comparison of monopolistic extraction to competitive and socially optimal extraction

We formalize the famous statement that the monopolist is the conservationists best friend, showing that, if resources are extracted in finite time, the monopolist will spread resource extraction over a longer time horizon. We continue to assume increasing average costs $c'(q) > 0$. The basic idea is simple, the monopolist charge a mark-up on the resource, and a higher price reduces consumption. However, we already know that any change of the Euler equation also changes the rent. As a result the comparison of the two price and consumption paths is a bit more sophisticated. Our proof proceeds in three steps. First, we combine the first order conditions and the transversality condition to express the resource's shadow value as a function of the last period shadow value. This step is convenient because we found that the choke price coincides across the monopolistic, the socially optimal, and the competitive market setting. Second, we express the competitive firm's

shadow value as a function of the monopolist's terminal time T^M . These two steps help us to make the two scenarios comparable by shifting price and extraction paths to anchoring them with respect to terminal rather than initial conditions, using the coinciding choke price as the common anchor point. The top left panel of Figure 7 shows the price paths for the competitive firm and the monopoly as they might look in calendar time, if our suspicion is right that the monopoly charges a higher price and, as a result, the competitive setting exhausts the resource faster.⁷⁵ The lower left panel shows how steps one and two shift the competitive price path to the right in a way that both exhaust the resource at the same point; the path of the competitive firm is no longer in calendar time but shifted. Once we have shifted the price paths accordingly, we can show that the competitive price lies below the monopolist's price because of the mark-up. Finally, using our finding that either market form will use up all of the available resource, we show that it takes longer along the monopolist's extraction path to use up all of the resource.

Let C label "competitive". Given our assumptions, we know that the competitive setting decentralizes the social optimum. Let M label "monopolistic", and let $i \in \{C, M\}$. In both the competitive market equilibrium (equation 9.7) and the monopolistic setting (equation 9.25) we found that $\mu_{T^i} = p(0) - c(0)$. Therefore

$$\begin{aligned}\mu_T &\equiv p(0) - c(0) = \mu_{T^i}^i = \mu_0^i \exp(rT^i) \\ \Rightarrow \mu_t^i &= \mu_T \exp(r(t - T^i)),\end{aligned}$$

where we emphasize that μ_T does not carry an index because it is the same for both settings. We define $\delta = T^M - T^C$ as the difference in time to full exhaustion between the monopolistic and the competitive setting. We shift the competitive firm's shadow value path by δ to enable comparability of the extraction rates

$$\begin{aligned}\mu_t^C &= \mu_T \exp(r(t - T^C)) = \mu_T \exp(r(t + \delta - T^M)) \\ \Rightarrow \mu_{t-\delta}^C &= \mu_T \exp(r(t - T^M)) = \mu_t^M.\end{aligned}$$

⁷⁵Note that the competitive price has to overtake the monopoly's price at some point in order to reach the choke price. While it is indeed charging more at this point in time, this point in time corresponds to different reserves left in the ground. In the competitive setting, the resource will be much closer to exhaustion than under the monopolistic setting at the point where the two curves intersect.

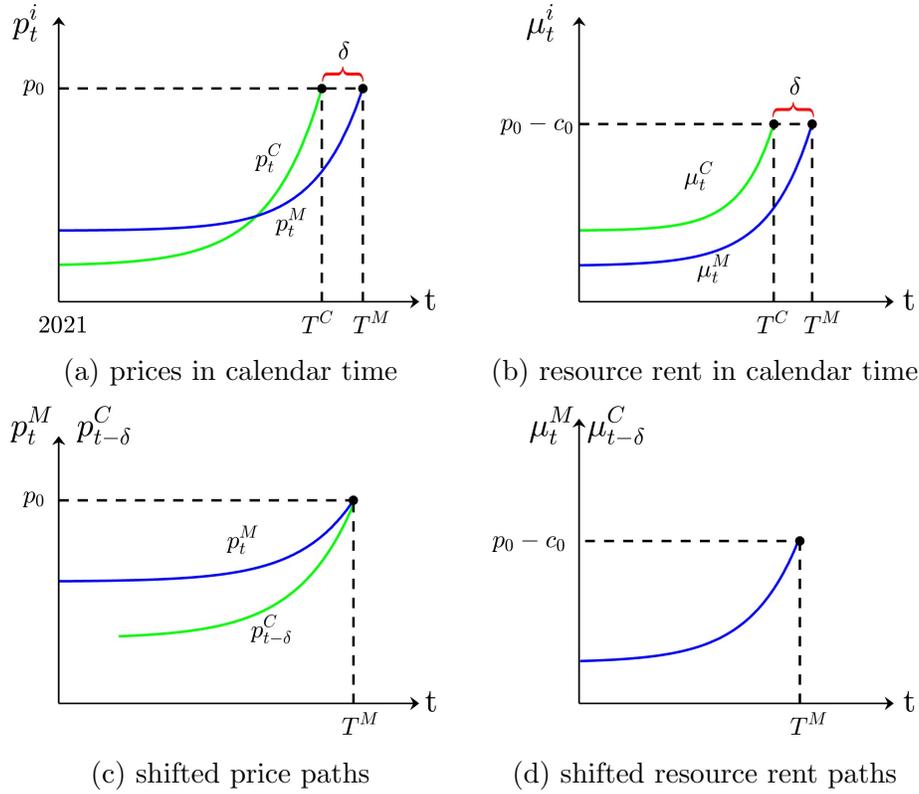


Figure 7 The monopolist is the conservationist’s best friend: intuition for the proof.

Using aggregate cost $C(q_t)$, the price in the monopolistic setting is given by equation (9.22) as

$$p_t = C'(q_t) + \mu_t^M + \underbrace{-p'(q_t)q_t}_{\equiv M(q_t)} \geq 0$$

and in the competitive setting by equation (9.4) as

$$p_t = \mu_t + C'(q_t) \quad \Leftrightarrow \quad p_{t-\delta} = \mu_{t-\delta} + C'(q_{t-\delta}).$$

Together, these last three equations imply

$$\begin{aligned}
& p(q_{t-\delta}^C) - C'(q_{t-\delta}^C) = \mu_{t-\delta}^C = \mu_t^M = p(q_t^M) - C'(q_t^M) - M(q_t^M) \\
\Rightarrow & p(q_{t-\delta}^C) - C'(q_{t-\delta}^C) = p(q_t^M) - C'(q_t^M) - M(q_t^M) \\
\Rightarrow & p(q_{t-\delta}^C) - C'(q_{t-\delta}^C) \leq p(q_t^M) - C'(q_t^M) \\
\Rightarrow & q_{t-\delta}^C \geq q_t^M, \tag{9.26}
\end{aligned}$$

where the last step uses that the function $h(q) = p(q) - C'(q)$ is strictly decreasing because $p'(q) < 0$ and $C''(q) > 0$. The comparison states that, if we were to shift the extraction paths in such a way that the terminal time coincides, then extraction in the competitive setting would be higher. In the final period both settings extract nothing and the inequality is weak. However, in most periods the monopolist sets a positive mark up $M(q_t) > 0$ and the inequality will be strict. We can now compare the aggregate resource use in the two settings over suitable parts of the time horizon

$$\int_0^{T^M} q_t^M dt = R_0 = \int_0^{T^C} q_t^C dt = \int_\delta^{T^C+\delta} q_{t-\delta}^C dt = \int_\delta^{T^M} q_{t-\delta}^C dt > \int_\delta^{T^M} q_t^M dt.$$

The first two equalities use the terminal conditions (9.7) and (9.25), the third equality is a change in the integration variable, the fourth equality follows by definition of δ , and the inequality follows from equation (9.26), realizing that the inequality will be strict during some time intervals. Comparing the first integral to the last integral, $q_t \geq 0$ implies $\delta > 0$ and, thus, $T^M > T^C$. If extraction happens in finite time, the monopolist will extract more slowly.

9.5 Stock-dependent extraction costs and exogenous technological progress

Our earlier settings ignored two important features of contemporaneous resource extraction. First, the costs of extraction can depend on the amount of the resource left in the ground. Second, we have been experiencing massive technological progress over the past decades, reducing extraction costs. Generalizing the competitive setting, we assume that a representative firm faces the extraction costs $C(q_t, R_t, t)$. For now we assume $C_{q_t}(q_t, R_t, t) \equiv \frac{\partial C}{\partial q_t} > 0$ and $C_{R_t}(q_t, R_t, t) \equiv \frac{\partial C}{\partial R_t} < 0$. We will add further structure to this cost

function as we move along. The firm's optimal control problem is

$$\max \int_0^T (p_t q_t - C(q_t, R_t, t)) e^{-rt} dt \quad \text{s.t. } \dot{R} = -q_t, R_0 = \bar{R}, R_t \geq 0.$$

The current value Hamiltonian is

$$H^c(q_t, \mu_t) = p_t q_t - C(q_t, R_t, t) - \mu_t q_t$$

and the first order conditions are

$$\begin{aligned} \frac{\partial H^c(q_t, \mu_t)}{\partial q_t} = p_t - C_{q_t} - \mu_t &\stackrel{!}{=} 0 &\Rightarrow \mu_t = p_t - C_{q_t} \\ & &\dot{\mu}_t = \frac{d}{dt}(p_t - C_{q_t}) \\ \frac{\partial H^c(q_t, \mu_t)}{\partial R_t} = -C_{R_t} &\stackrel{!}{=} r\mu_t - \dot{\mu}_t &\Rightarrow \dot{\mu}_t = C_{R_t} + r\mu_t. \end{aligned}$$

Combining the two first order conditions into the Euler equation we find

$$\begin{aligned} \frac{d}{dt}(p_t - C_{q_t}) &= C_{R_t} + r(p_t - C_{q_t}) \\ \Rightarrow \frac{\frac{d}{dt}(p_t - C_{q_t})}{p_t - C_{q_t}} &= r - \underbrace{\frac{-C_{R_t}}{p_t - C_{q_t}}}_+. \end{aligned} \quad (9.27)$$

The left side of equation (9.27) represents the familiar growth rate of the resource rent $p_t - C_{q_t}$. The right side shows that, if extraction costs depend on the amount of the resource left in the ground, this resource rent no longer has to grow at the rate of interest. The interest rate represents an opportunity cost. By leaving a unit of the resource in the ground today, we forgo the interest rate r that we could earn by extracting today and depositing the profits on a bank account. However, under stock dependent extraction costs, we also have a direct benefit from leaving a unit of the resource in the ground. If we extract a unit less today, then extraction will be a little cheaper all along the future extraction path. This payoff from leaving the resource in the ground is subtracted from the opportunity cost (or the benefit of immediate extraction). As a result, the resource rent only has to grow at the rate of interest less the extraction cost savings that result from leaving a marginal unit in the ground (relative to marginal net revenue).

Next, we analyze the left side of equation (9.27) more closely. The numerator reads

$$\frac{d}{dt}(p_t - C_{q_t}(q_t, R_t, t)) = \dot{p}_t - \frac{\partial C_{q_t}}{\partial t} - \frac{\partial C_{q_t}}{\partial q_t} \dot{q}_t - \frac{\partial C_{q_t}}{\partial R_t} \dot{R}_t$$

First, the change of the resource rent from one period to the next is determined by the difference between the price change p_t and the direct time change of the marginal extraction cost $\frac{\partial C_{q_t}}{\partial t}$ resulting from potential technological progress. Technological progress would reduce extraction costs, thereby increasing the rent over time and reducing the price increase required to hold on to the resource. Second, if resource use falls over time, this flow reduction further reduces the flow dependent extraction costs ($\frac{\partial C_{q_t}}{\partial q_t} = \frac{\partial^2 C}{\partial q_t^2} > 0$ and $\dot{q}_t < 0$) reducing the required price increase. Finally, if *marginal* extraction costs increase with the depletion of resources, then the last term is positive and increases the required price increase ($\frac{\partial C_{q_t}}{\partial R_t} = \frac{\partial^2 C}{\partial q_t \partial R_t} < 0$ and $\dot{R}_t < 0$). This last contribution works into the opposite direction of the other terms, including the stock-dependent cost term on the right side of equation (9.27). As we deplete the resource, marginal extraction costs in the future will increase, partially offsetting the reduction of extraction costs over time represented by the earlier terms.⁷⁶

We illustrate these effects in an example where the cost function takes the form

$$C(q_t, R_t, t) = Q(q_t) \mathcal{R}^q(R_t) \mathcal{T}^q(t) + \mathcal{R}^R(R_t) \mathcal{T}^R(t).$$

The left summand controls marginal extraction costs and we assume $Q'(q_t) > 0$, $Q''(q_t) \geq 0$. We permit for the possibility that marginal extraction costs increase as we deplete the available stock of the resource, assuming $\mathcal{R}^{q'}(R_t) \geq 0$. Moreover, we permit for exogenous technological progress that reduces marginal extraction costs, assuming $\mathcal{T}^{q'}(t) \leq 0$. The second summand controls “per period fix costs” of extraction that are independent of the resource flow but depend on the remaining stock of the resource, $\mathcal{R}^{R'}(R_t) \leq 0$. E.g., as we deplete reservoirs close to the surface, we have to drill deeper. Some of

⁷⁶We emphasize the difference between this intuition and the mechanism underlying the right side of equation (9.27). The latter characterizes a benefit from leaving a unit of the resource in the ground. The present effect takes the decline of the resource stock as given ($\dot{R} < 0$) and reflects that, consequently, marginal extraction costs will rise (or fall less).

these additional costs for opening up the resource will be independent of the extraction flow. Again, we permit for exogenous technological progress that reduces these costs over time, $\mathcal{T}^{R'}(t) \leq 0$. We abbreviate relative changes of these cost factors as follows

$$\widehat{Q}(q_t) = \frac{Q''(q_t)}{Q'(q_t)} \quad \widehat{\mathcal{R}}^q(R_t) = -\frac{\mathcal{R}^{q'}(R_t)}{\mathcal{R}^q(R_t)} \quad \widehat{\mathcal{T}}^q(t) = -\frac{\mathcal{T}^{q'}(t)}{\mathcal{T}^q(t)}$$

all of which are defined to be positive. Then the Euler equation (9.27) takes the form

$$\dot{p}_t = C_{q_t} \left[\underbrace{\widehat{Q}(q_t)\dot{q}_t}_{-} - \underbrace{\widehat{\mathcal{T}}^q(t)}_{-} - \underbrace{\widehat{\mathcal{R}}^q(R_t)\dot{R}_t}_{+} \right] + r(p_t - C_{q_t}) + \underbrace{C_{R_t}}_{-}, \quad (9.28)$$

where the first sign assumes $\dot{q}_t < 0$. We recall the Euler equation (9.3) for constant marginal extraction costs

$$\dot{p}_t = r(p_t - c).$$

With constant marginal extraction costs, the price increase has to equal the interest on the rent. The generalization in equation (9.28) still contains this base for the price change. A reduction in marginal extraction costs over time due to a falling extraction rate ($\widehat{Q}(q_t)\dot{q}_t$) and exogenous technological progress ($-\widehat{\mathcal{T}}^q(t)$) reduce this price increase. Depletion of the resource tends to increase the marginal extraction costs as it becomes harder to access the resource. Partially offsetting the decrease in marginal costs, it increases the price increase required to make the firm willing to hold on to the resource ($-\widehat{\mathcal{R}}^q(R_t)\dot{R}_t$). Finally, leaving a unit of resource in the ground makes future extraction cheaper ($C_{R_t} < 0$), increasing the benefit from extracting tomorrow rather than today and reducing the necessary price increase.

Further specifying the above example, let us assume that all functions are exponentials, i.e., $Q(q_t) = A \exp(aq_t)$, $\mathcal{R}^q(R_t) = \exp(-bR_t)$, $\mathcal{T}^q(t) = \exp(-\delta_q t)$, $\mathcal{R}^R(R_t) = D \exp(-dR_t)$, $\mathcal{T}^R(t) = \exp(-\delta_R t)$, which satisfy our assumptions on functional form (given all parameters are positive). Then equation (9.28) becomes

$$\dot{p}_t - rp_t = C_{q_t} [a\dot{q}_t - b\dot{R}_t - (\delta_q + r)] + C_{R_t} \quad (9.29)$$

$$\text{with } C_{q_t} = a A \exp(aq_t - bR_t - \delta_q t)$$

$$\text{and } C_{R_t} = -b A \exp(aq_t - bR_t - \delta_q t) - d D \exp(-dR_t - \delta_R t)$$

The interpretation remains the same as for equation (9.28). Yet, equation (9.29) makes it easier to get an impression of short versus long-term dynamics. Exogenous technological progress, by reducing marginal extraction costs at rate δ_q , can reduce the price increase and even turn it negative (e.g. shale gas boom). However, as technological progress reduces the marginal extraction costs, it also degrades the very base C_{qt} by which it reduces the price increase. By contrast, as long as the price remains positive, the interest rate continues to push the price up also in the long run (the term rp_t on the left). Similarly, the stock dependence of the resource will initially reduce the price increase. However, this effect falls over time. In particular, the technological progress governing marginal extraction costs that reduces the price increase in the first channel, also reduces the benefit of leaving a resource in the ground; it thereby diminishes the C_{R_t} -channel's ability to reduce the price increase. Technological progress governing stock-dependent (per period) fix-costs of opening up a new reservoir only reduces the benefit of leaving more of the resource in the ground and unambiguously speeds up the price increase. In our example, technological progress eventually drives the right side of equation (9.29) to zero, implying that the left side approaches an exponential price increase in the long run. In summary, in the short-run there can be several forces at work that reduce the price increase of a scarce nonrenewable resource. In the long run, we expect an exponential price increase at the rate of interest – at least if resource use stretches sufficiently long.

We note that the present discussion focuses on the qualitative price path. These changes in the price path affect extraction over time and, thus, the initial price and extraction levels. We conjecture that a dampening of the initial price increase discussed above will tend to reduce the initial extraction rate, resembling a reduction of the interest rate in our quantitative examples of Section 8.

9.6 Taxes: Profits versus Revenue

Nonrenewable resources carry a rent and these rents present a convenient way to generate government revenue. We show that taxing profits transfers income from the resource extracting firm to the government without undoing Section 9.2's finding that the competitive market equilibrium is Pareto optimal. In contrast, taxing the firm's revenue would distort the extraction

and the market would fail to replicate the social planner's solution. We note that most nonrenewable resources are privately or state-owned and extracted under well-defined property rights. Thus, in contrast to other chapters, this section does not consider taxation as a means to “repair” the market, but merely as an instrument to generate government income.⁷⁷

Profit Tax. We consider the firm's optimization problem from Section 9.1

$$\max \int_0^T (1 - \tau)(p_t q_t - C(q_t))e^{-rt} dt \quad \text{s.t.} \quad \dot{R} = -q_t, R_0 = \bar{R}, R_t \geq 0, T \text{ free}$$

adding a profit tax, i.e., a tax τ proportional to the firm's net revenue $p_t q_t - C(q_t)$. To check whether the tax distorts the resource extraction path we derive the Euler equation

$$\frac{\frac{d}{dt}(1 - \tau)(p_t - C'(q_t))}{(1 - \tau)(p_t - C'(q_t))} = r \quad \Rightarrow \quad \frac{\frac{d}{dt}(p_t - C'(q_t))}{(p_t - C'(q_t))} = r.$$

We observe that the tax cancels out and the Euler equation coincides with equation (9.5) in the absence of a profit tax, which we have shown to coincide with the social planner solution. Moreover, the tax does not affect the resource's equation of motion or the terminal and transversality condition. The only impact the tax could have is on the second order condition if it was to exceed 100%. Then the firm would stop extraction. For a tax $\tau \in (0, 1)$, we find that the tax has no impact on the extraction path and, therefore, does not distort the socially optimal extraction path. An even simpler way to realize that the tax does not affect the socially optimal extraction path is by realizing that a multiplicative constant can be pulled out of the integral and does not change the maximization problem.⁷⁸

Revenue Tax. Now we assume that the regulator taxes the firm's revenue at a constant rate τ . The firm's optimization problem becomes

$$\max \int_0^\infty [(1 - \tau)p_t q_t - C(q_t)]e^{-rt} dt \quad \text{s.t.} \quad \dot{R} = -q_t, R_0 = \bar{R}, R_t \geq 0, T \text{ free.}$$

⁷⁷Unfortunately, the extraction of nonrenewable resources sometimes cause environmental degradation. These externalities are not always internalized together with the property right governing the resources themselves. In these situations we could consider a Pigovian tax in addition to the taxation discussed below, preferable of course a tax directly addressing the externality.

⁷⁸While this is true for a multiplicative constant, we note (yet again) that we have to be more careful with an additive constant because an additive constant changes the objective by multiplying the endogenous length of the time horizon.

The current value Hamiltonian is

$$H^c(q_t, \mu_t) = (1 - \tau)p_t q_t - C(q_t) - \mu_t q_t,$$

and we obtain the necessary conditions for an optimum

$$\begin{aligned} \frac{\partial H^c(q_t, \mu_t)}{\partial q_t} &= (1 - \tau)p_t - C'(q_t) - \mu_t \stackrel{!}{=} 0 \quad \Rightarrow \quad \mu_t = (1 - \tau)p_t - C'(q_t) \\ \frac{\partial H^c(q_t, \mu_t)}{\partial R_t} &= 0 \stackrel{!}{=} r\mu_t - \dot{\mu}_t \quad \Rightarrow \quad \dot{\mu}_t = r\mu_t. \end{aligned}$$

that give rise to the Euler equation

$$\frac{\frac{d}{dt} [(1 - \tau)p_t - C'(q_t)]}{(1 - \tau)p_t - C'(q_t)} = r \quad \Leftrightarrow \quad \frac{\frac{d}{dt} \left[p_t - \frac{C'(q_t)}{(1 - \tau)} \right]}{p_t - \frac{C'(q_t)}{(1 - \tau)}} = r,$$

which differs from our earlier Euler equation that characterizes the social optimum. The right version of the equation points out that a revenue tax has the same impact on the extraction path as an increase in marginal extraction costs. Unless extraction costs are zero, a revenue tax distorts the market equilibrium. We note that, once again, we could have obtained this insight more quickly by pulling the factor $(1 - \tau)$ out of the integral to obtain the equivalent optimization problem

$$\max \int_0^\infty \left[p_t q_t - \frac{C(q_t)}{1 - \tau} \right] e^{-rt} dt \quad \text{s.t.} \quad \dot{R} = -q_t, \quad R_0 = \bar{R}, \quad R_t \geq 0, \quad T \text{ free.}$$

We close by noting that the effective increase in marginal extraction costs can also imply that some of the resource will be left in the ground even though extraction would be socially beneficial.

10 Non-Convex Control Problems

This section is an introduction to non-convex control problems. These problems feature a non-concave Hamiltonian so that the necessary conditions don't generally identify a unique optimal trajectory. In general, we can observe multiple steady states and have to analyze various trajectories leading into different steady states. For our discussion, we pick the problem of optimally controlling phosphor inflow into shallow lakes. The dynamics of these lakes exhibits a sort of memory effect that is called hysteresis. For the optimally controlled system we find two saddle point stable steady states that can be identified with the lake being in an oligotrophic (clear) or a eutrophic (turbid) state. We show the existence of a so called Skiba point. For levels of the phosphor stock below this point it is optimal to pick a trajectory leading into the clean steady state, while for levels above it is optimal to settle with or move into a eutrophic state. The section builds on three papers published in a special issue of the Journal Environmental and Resource Economics (vol 26) on the economics non-convex ecosystems. The articles are Dasgupta & Mäler (2003), Brock & Starrett (2003), and Mäler, Xepapadeas & de Zeeuw (2003).

10.1 'Convex' Control Problems

Before we discuss a non-convex control problem, we briefly review what constitutes a convex problem. Let us recall the typical layout of the continuous time optimal control problem that we attacked with by means of the maximum principle.

$$\max \int_0^{\infty} u(c, x) e^{-\rho t} dt \quad (10.1)$$

$$\text{subject to } \dot{x} = f(c, x), \quad x(0) = x_0$$

yielding the current value Hamiltonian

$$H(c, x, \lambda) = u(c, x) + \lambda f(c, x)$$

and the necessary conditions for an optimum

$$\frac{\partial H}{\partial c_t} = 0, \quad \frac{\partial H}{\partial x_t} = \rho\lambda_t - \dot{\lambda}_t, \quad \frac{\partial H}{\partial \lambda_t} = \dot{x}_t,$$

and the transversality condition $\lim_{t \rightarrow \infty} e^{-\rho t} H(c, x, \lambda) = 0$

and under stronger conditions $\lim_{t \rightarrow \infty} e^{-\rho t} \lambda_t = 0$.

Paths $(c_t^*, x_t^*)_{t \in [0, \infty)}$ satisfying these conditions are candidates for an optimum. In a convex problem we have in addition sufficiency conditions (Mangasarian/Arrow) that state something like:

Let $\lambda \geq 0$ and $H(c, x, \lambda)$ be concave in (c, x) for all λ . Then the candidate(s) $(c_t^*, x_t^*)_{t \in [0, \infty)}$ are globally optimal paths. Moreover, if $H(c, x, \lambda)$ is strictly concave in (c, x) for all λ , then $(c_t^*, x_t^*)_{t \in [0, \infty)}$ is a unique optimal path.

Here we are slightly lax about the transversality conditions required in addition for sufficiency. Note that $H(c, x, \lambda)$ concave for $\lambda \geq 0$ is equivalent to $u(c, x)$ and $f(c, x)$ being concave in (c, x) . Why is such a control problem is generally referred to as “convex”? Here are our two explanations

1. Historically people were analyzing min-problems rather than max problems, so that problem (10.1) turns into

$$\min \int_0^{\infty} -u(c, x) e^{-\rho t} dt$$

with $-u$ convex rather than concave.

2. For a typical “capital stock investment (\dot{x}) versus consumption (c)” model with free disposal (\leq) where the constraint of motion is

$$\dot{x} \leq F(x) - c$$

we find that $F(x)$ concave implies that the set of feasible programs $\{(c_t, x_t)_{t \in [0, \infty)} \text{ s.th. } \dot{x} \leq F(x) - c\}$ is convex:

Let $(c_t^\circ, x_t^\circ)_{t \in [0, \infty)}$ and $(c_t^\diamond, x_t^\diamond)_{t \in [0, \infty)}$ be two feasible programs. Then $(\tilde{c}_t, \tilde{x}_t)_{t \in [0, \infty)}$ with

$$\begin{aligned}\tilde{c}_t &= \gamma c_t^\circ + (1 - \gamma) c_t^\diamond \\ \tilde{x}_t &= \gamma x_t^\circ + (1 - \gamma) x_t^\diamond\end{aligned}$$

is feasible as well for all $\lambda \in [0, 1]$.

Proof: By feasibility of $\{c_t^\circ, x_t^\circ\}_{t \in [0, \infty)}$ and $\{c_t^\diamond, x_t^\diamond\}_{t \in [0, \infty)}$ we have

$$\begin{aligned}\gamma \dot{x}_t^\circ &\leq \gamma F(x_t^\circ) - \gamma c_t^\circ \\ (1 - \gamma) \dot{x}_t^\diamond &\leq (1 - \gamma) F(x_t^\diamond) - (1 - \gamma) c_t^\diamond\end{aligned}$$

and adding these up we find

$$\begin{aligned}\dot{\tilde{x}}_t &= \gamma \dot{x}_t^\circ + (1 - \gamma) \dot{x}_t^\diamond \\ &\leq \gamma F(x_t^\circ) + (1 - \gamma) F(x_t^\diamond) - (\gamma c_t^\circ + (1 - \gamma) c_t^\diamond) \\ &\leq F(\gamma x_t^\circ + (1 - \gamma) x_t^\diamond) - (\gamma c_t^\circ + (1 - \gamma) c_t^\diamond) \\ &\leq F(\tilde{x}_t) - \tilde{c}_t.\end{aligned}$$

The second inequality is due to concavity of $F(x)$. Thus $(\tilde{c}_t, \tilde{x}_t)_{t \in [0, \infty)}$ is feasible.

10.2 The Shallow Lake Problem

Shallow lakes are sensitive to phosphorus inflow (loading) caused by fertilizer use in the agricultural sector. For low amounts of phosphorous the ecology of the shallow lake gives a positive feedback to an increase in phosphorous loading.⁷⁹ An increase in the phosphorous stock triggers the growth of algae and can cause a loss of oxygen in the water. The water becomes turbid and the ecosystem changes (to the distress of trouts, fishermen, and swimmers). Such a state is called eutrophic. A state of the lake with clear water is referred to as oligotrophic. In small ponds the change from an oligotrophic

⁷⁹Another way to think about it is that for low stocks and low flows of phosphorous some mechanisms in the lake hold back part of the phosphorous from going into the water and producing algae. This 'holding back effect' decreases as the stock increases (while the natural purification rate increases as the stock of phosphorous increases).

state to a eutrophic state can happen within a day, in lakes within weeks. The following analysis will explain the underlying dynamics.

We denote the stock of phosphorous (suspended in algae) in the lake by $x(t)$. The loading $c(t)$ denotes the amount of phosphorous flowing into the lake from the watershed, stemming from agricultural production. The (self-) purification rate of the lake is α . In addition, we have a function $f(x)$ capturing the feedback. It increases stronger for small phosphorous stocks than for large phosphorous stocks and is convex-concave. We assume $\lim_{t \rightarrow \infty} f'(x) = 0$. An example that we will use for numeric analysis is: $f(x) = \frac{x^2}{1+x^2}$. The lake dynamics are governed by the equation of motion

$$\dot{x}_t = c_t - \alpha x_t + f(x_t) \equiv X(c_t, x_t) . \quad (10.2)$$

Thus, a lake with a constant loading c is in equilibrium if

$$\begin{aligned} \dot{x}_t &= 0 \\ \Leftrightarrow \alpha x_t - c_t &= f(x_t) \end{aligned} \quad (10.3)$$

$$\Leftrightarrow c_t = \alpha x_t - f(x_t) \equiv h(x) . \quad (10.4)$$

Figure 8 depicts the two sides of equation (10.3) for different constant loadings c . Without any phosphorous inflow from the watershed we have a unique equilibrium at a zero phosphorous stock. With a small inflow (dashed line) we find three equilibria and history determines which one the lake is in. Note that you can read off the phosphorous loading as the absolute of the intersection with the y -axis. Assume we start with a zero phosphorous stock and slowly load the lake. we increase the loading slowly so that we move in or close to the equilibrium. Then, the equilibrium stock moves up the $f(x)$ curve as we increase this almost constant slowly increasing loading until we reach x_1 . If we further increase the loading, the equilibrium with the low phosphorous stock disappears. The dotted line representing net loading now lies underneath the $f(x)$ line. By equation (10.2) we know

$$\alpha x_t - c < f(x_2) \Rightarrow \dot{x}_t > 0 .$$

Thus, the lake will move all the way to \bar{x} . This move in Figure 8 corresponds to the move from a clear to a turbid lake or from an oligotrophic to a eutrophic state. Note that not only a small difference in the loading can cause a huge difference in the equilibrium phosphorous stock, but also a decrease of the

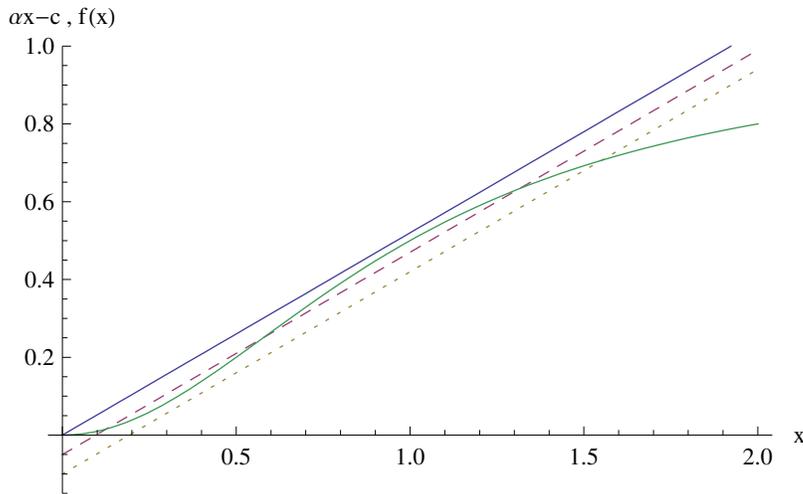
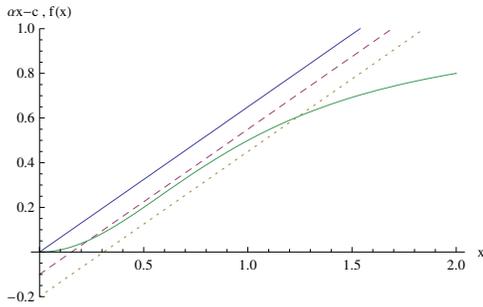


Figure 8: Shallow lake model. Feedback $f(x)$ curve and ‘net purification’ $\alpha x - c$ curve for $c = 0$ (straight line), $c = .05$ (dashed line) and $c = .1$ (dotted line). The purification parameter determines the slope of the line, here it is $\alpha = .52$. The feedback function is $f = \frac{x^2}{1+x^2}$.

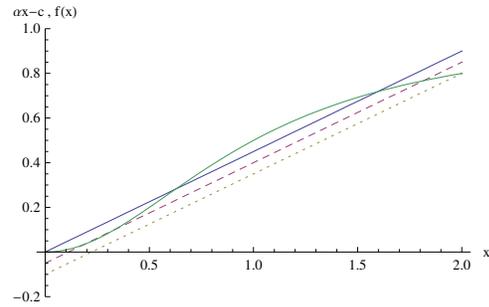
loading does not immediately take us back to x_1 . If we start reducing the loading we slowly move down the $f(x)$ curve until we reach x_2 . This point corresponds to a significantly lower loading than the one corresponding to x_1 , but still to a higher phosphorous stock. Only if we decrease the loading further we move back to the lower part of the $f(x)$ curve ($\alpha x_t - c > f(x_2) \Rightarrow \dot{x}_t < 0$) and reach a new equilibrium at \underline{x} . This effect is called *hysteresis*.

The scenario described above is the one that we would like to analyze in more detail and from an economic perspective. However, we point out that two other qualitatively different scenarios are possible if the purification rate of the lake is higher respectively lower than in the above scenario. With a high enough purification rate α we would ultimately reach the scenario depicted in Figure 9.a) where we are back to always finding a unique steady state and a continuous adjustment of the equilibrium phosphorous stock to changes in the loading. On the other hand, if the purification rate is sufficiently small we end up with the scenario depicted in Figure 9.b). Here, even a reduction of the loading to zero cannot get us out of a eutrophic state. A move from the lower branch of $f(x)$ to the upper branch is *irreversible*.

Heading for the economic analysis let us translate Figure 8 into a diagram



9.a) Unique Oligotrophic Equilibrium



9.b) Lake with Irreversible Eutrophication

Figure 9: Shallow lake model. Feedback $f(x)$ curve and ‘net purification’ $\alpha x - c$ curve. If the purification rate is high enough the lake will always have a unique equilibrium (here $\alpha = .65$). If the purification rate is small enough the lake will not be able to pass back from the eutrophic state to the oligotrophic state even if phosphor inflow seizes completely (here $\alpha = .4$). The feedback function $f = \frac{x^2}{1+x^2}$ is the same as in Figure 8.

that corresponds to equation (10.4) and depicts the difference between the purification and the feedback of the lake (i.e. the difference between the solid straight line and the $f(x)$ curve in Figure 8). In Figure 10 we find for any constant loading c the corresponding equilibria of the lake on a horizontal line through c . By adding the off-equilibrium phosphorous stock dynamics to Figure 10, we can also observe that the first and the third equilibrium are stable while the intermediate equilibrium is unstable. For this purpose we take the derivative of the equation of motion for x_t , defined in equation (10.2) as $X(c_t, x_t)$, with respect to the loading c_t and find

$$\frac{\partial X(c_t, x_t)}{\partial c_t} = 1.$$

Given, the change of the stock is zero on the line, we know that the stock is increasing above the curve and decreasing below. In Figure 11 we plot the corresponding vector field of $X(c, x)$ where the length of the arrows is proportional to the magnitude of \dot{x} at the corresponding point. We can observe how the off-equilibrium dynamics drive us from x_1 over to the high phosphorous stock equilibrium on the right branch of the curve if we increase c a little bit more.

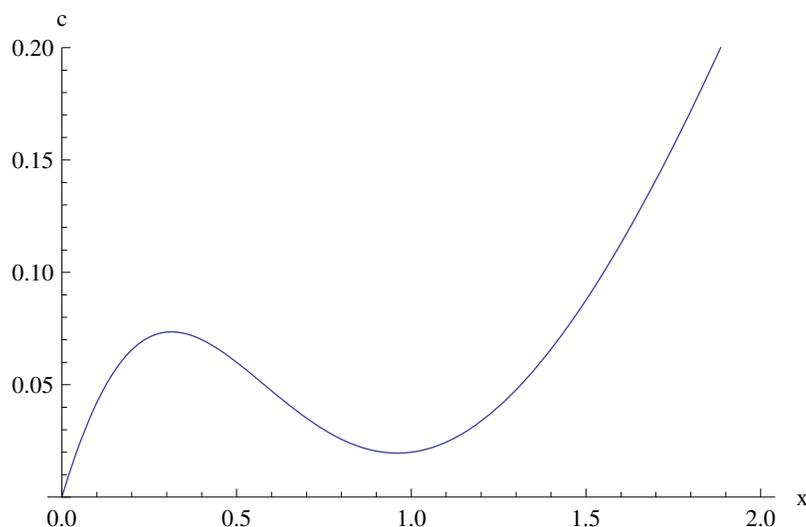


Figure 10: Shallow lake model. Net purification (after feedback). Same specifications as in Figure 8.

10.3 Optimal Control of the Shallow Lake

So far we were only contemplating how the lake reacts to different loadings. Now we will analyze how to obtain the optimal loading. We face a trade-off between the recreational value of a clear lake and its use value as a ‘dumping site’ for phosphorous which is implied by the use of fertilizer in the agricultural sector. We will assume the following specification of welfare

$$W = \int_0^{\infty} (\ln c - \beta x^2) e^{-\rho t} dt.$$

Here β is a scaling parameter for the (quadratic) appreciation of a clear lake and ρ is the rate of pure time preference. The logarithmic specification of the use value of the lake as a dumping site for fertilizer implies that a social planner will never choose a zero loading, e.g. because the complete abandonment of fertilizer could cause a food shortage that outweighs any other concern. Optimizing W with respect to the constraint given in equation (10.2) and some initial stock x_0 we find

$$H(c, x, \lambda) = \ln c - \beta x^2 + \lambda(c - \alpha x + f(x))$$

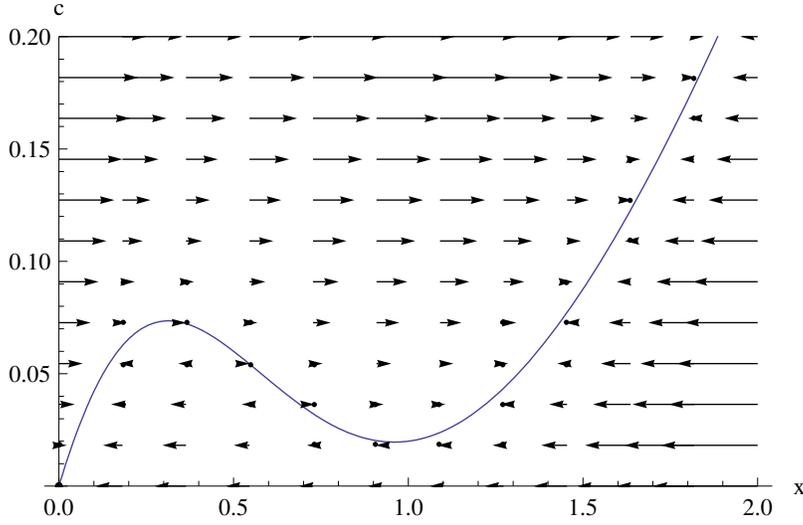


Figure 11 : The magnitude and direction of the arrows indicates the change of the phosphorous stock \dot{x} . Same specifications as in figures 8 and 10.

and

$$\begin{aligned}
 \frac{\partial H}{\partial c_t} &= \frac{1}{c} + \lambda \stackrel{!}{=} 0 \quad \Rightarrow \lambda = -\frac{1}{c} \quad \Rightarrow \dot{\lambda} = \frac{\dot{c}}{c^2} \\
 \frac{\partial H}{\partial x_t} &= -2\beta x + \lambda(-\alpha + f'(x)) \stackrel{!}{=} \rho\lambda_t - \dot{\lambda}_t \\
 \Rightarrow -2\beta x + \frac{\alpha}{c} - \frac{f'(x)}{c} &= -\frac{\rho}{c} - \frac{\dot{c}}{c^2} \\
 \Rightarrow \dot{c} &= 2\beta x c^2 - (\alpha + \rho - f'(x))c \equiv C(c, x). \tag{10.5}
 \end{aligned}$$

Thus, stationarity of the optimal loading holds for $\dot{c} = 0$ which is equivalent to

$$\begin{aligned}
 c = 0 \quad \text{or} \quad c &= \frac{\alpha + \rho - f'(x)}{2\beta x} \equiv g(x) \\
 \Rightarrow \alpha + \rho - f'(x) &= 2\beta x c \quad \text{if } c > 0. \tag{10.6}
 \end{aligned}$$

Figure 12 adds the $\dot{c} = 0$ curve into the lake model of Figure 10. While Figure 12 depicts the scenario that we would like to analyze, some qualitatively different figures can result as well. If we increase the discount rate sufficiently,

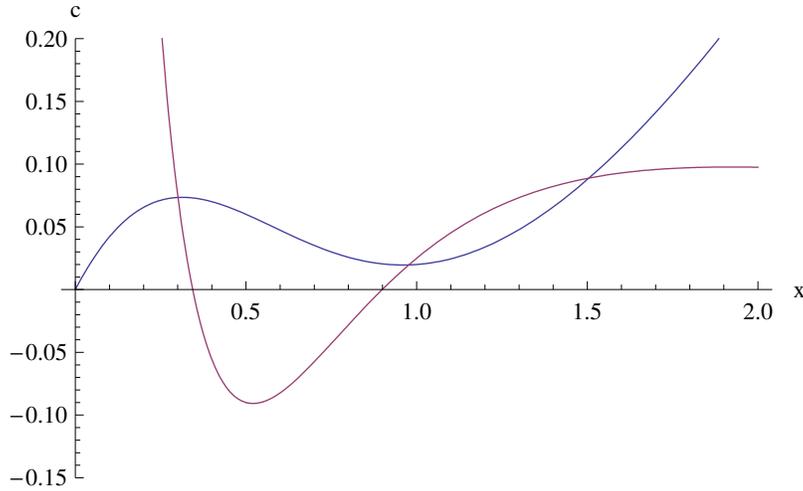
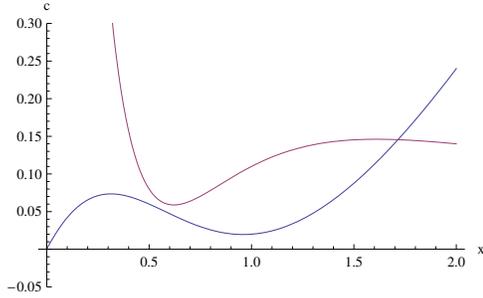


Figure 12: $\dot{x} = 0$ and $\dot{c} = 0$ curves for the shallow lake model. Here the purification parameter $\alpha = .52$ is chosen to yield a reversible lake dynamics with hysteresis. The feedback function is $f = \frac{x^2}{1+x^2}$. The other parameters are chosen as $\beta = 1$ characterizing the appreciation of the clear lake and a discount rate of $\rho = 3\%$. These specifications result in three intersections of the depicted curves.

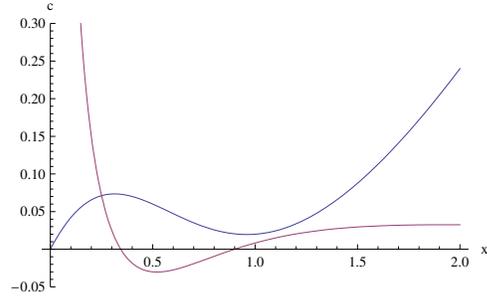
we end up with a unique intersection in the eutrophic region (Figure 13.a). On the other hand, if we increase the appreciation of the clear lake expressed by the parameter β , we end up with a unique equilibrium in the oligotrophic region (Figure 13.b). In order to analyze the dynamics off the steady state equilibria we analyze how the change in loading $C(c, x)$ defined in equation (10.5) depends on the loading and find

$$\left. \frac{\partial C(c_t, x_t)}{\partial c_t} \right|_{\dot{c}=0} = -(\alpha + \rho - f'(x)) + 4\beta xc = 2\beta xc > 0$$

where we used equation (10.6) assuming $c > 0$. Thus, the optimal loading increases above the $\dot{c} = 0$ -line and decreases below of it. With that information we can draw the phase diagrams in figures 14 and 15.



13.a) Unique Eutrophic Equilibrium



13.b) Unique Oligotrophic Equilibrium

Figure 13: Alternative $\dot{c} = 0$ curves yielding unique steady states. Figure 13.a) is obtained by increasing the discount rate to $\rho = 20\%$. Figure 13.b) is obtained by increasing the appreciation of the clear lake to $\beta = 3$. The other parameter are the same as in Figure 12 ($\alpha = .52$, $\rho = 3\%$ and $\beta = 1$ respectively $\rho = 3\%$).

10.4 A Close Up of the Steady States

Linearizing the dynamic system characterized by equations (10.2) and (10.5)

$$\begin{aligned} \dot{x} &= c - \alpha x + f(x) \\ \dot{c} &= 2\beta x c^2 - (\alpha + \rho - f'(x)) c \end{aligned}$$

in the steady states results in

$$\begin{aligned} \dot{x} &= (f'(x) - \alpha) \quad dx + \quad \quad \quad dc \\ \dot{c} &= (cf''(x) + 2\beta c^2) \quad dx + (4cx\beta - (\alpha + \rho) + f'(x)) \quad dc \\ \Leftrightarrow \begin{pmatrix} \dot{x} \\ \dot{c} \end{pmatrix} &= \begin{pmatrix} f'(x) - \alpha & 1 \\ cf''(x) + 2\beta c^2 & 4cx\beta - (\alpha + \rho) + f'(x) \end{pmatrix} \begin{pmatrix} dx \\ dc \end{pmatrix} \\ \Leftrightarrow \begin{pmatrix} \dot{x} \\ \dot{c} \end{pmatrix} &= \underbrace{\begin{pmatrix} f'(x) - \alpha & 1 \\ cf''(x) + 2\beta c^2 & 2cx\beta \end{pmatrix}}_{\equiv J} \begin{pmatrix} dx \\ dc \end{pmatrix} \end{aligned}$$

where we used equation (10.6) which will be used in the next line again. Then

$$\text{tr}(J) = f'(x) - \alpha + 2cx\beta = \rho$$

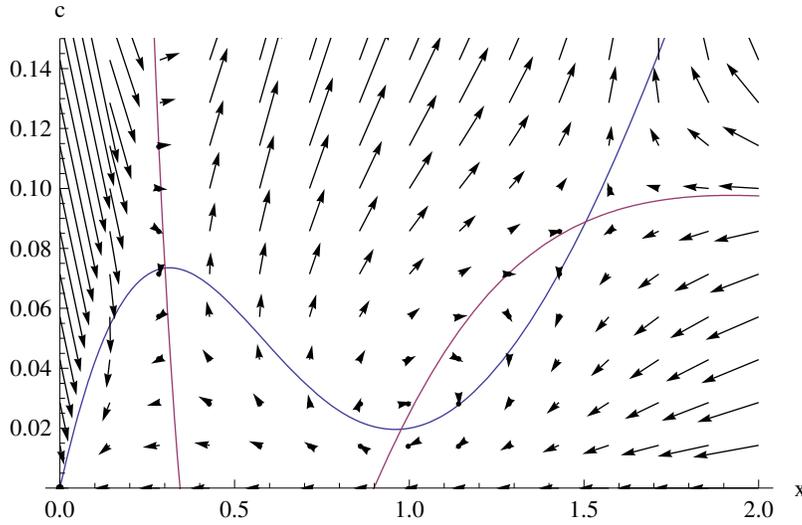


Figure 14 : Same as Figure 12 with off-equilibrium dynamics added. The magnitude and direction of the arrows correspond to the according change in phosphorous stock and loading. Parameters are $\alpha = .52$, $\beta = 1$ and $\rho = 3\%$.

and, like always, the equilibrium can be at most saddle point stable (at least one Eigenvalue has to be positive). The determinant of J is unfortunately a little harder to sign

$$\det(J) = c [(f'(x) - \alpha)2x\beta - (f''(x) + 2\beta c)] .$$

In order to sign the determinant a worthwhile try is comparing the above expression to the ratio of the slopes of h and g in the steady state. Obviously, which of the curves is steeper alternates over the three steady states (which we suspect to alternate in stability). Moreover, which of the two curves cuts from above respectively below determines the pattern of the arrows in the phase diagram around the steady state. The derivatives are (using equation 10.6)

$$g'(x) = -\frac{xf''(x) + \alpha + \rho - f'(x)}{2\beta x^2} = -\frac{f''(x) + 2\beta c}{2\beta x}$$

$$h'(x) = \alpha - f'(x)$$

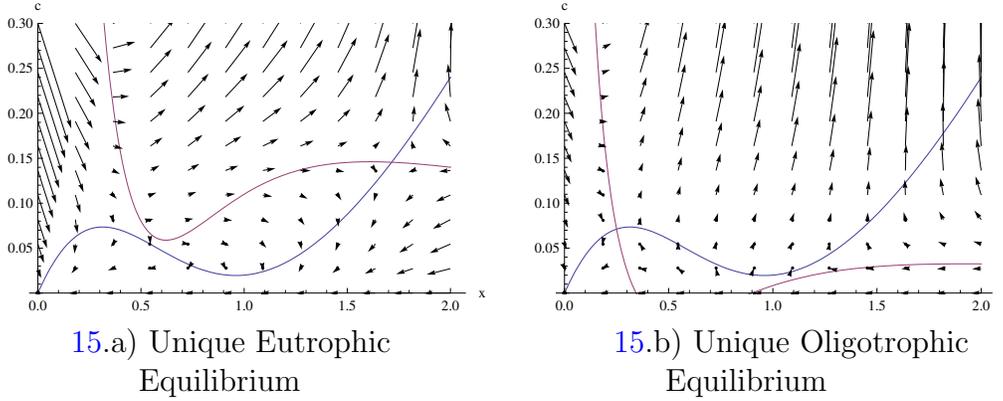


Figure 15 : Same as Figure 13 with off-equilibrium dynamics added. The magnitude and direction of the arrows correspond to the according change in phosphorous stock and loading. Figure 13.a) is obtained by increasing the discount rate to $\rho = 20\%$. Figure 13.b) is obtained by increasing the appreciation of the clear lake to $\beta = 3$. The other parameter are the same as in Figure 14 ($\alpha = .52$, $\rho = 3\%$ and $\beta = 1$ respectively $\rho = 3\%$).

implying that

$$\begin{aligned}
 &g'(x) > h'(x) \\
 \Leftrightarrow &-f''(x) - 2\beta c > 2\beta x(\alpha - f'(x)) \\
 \Leftrightarrow &\frac{\det(J)}{c} > 0
 \end{aligned}$$

Hence $\text{Det}(J)$ is negative for x^* and x^{***} and positive for the steady state x^{**} in between. In consequence x^* and x^{***} are saddle point stable as both Eigenvalues

$$\lambda_{1,2} = \frac{\text{tr}(J)}{2} \pm \sqrt{\left(\frac{\text{tr}(J)}{2}\right)^2 - \det(J)}$$

are real and λ_1 is positive while λ_2 is negative. For x^{**} the Eigenvalues are either both real and positive if $\text{tr}(J)^2 \geq 4 \det(J)$, or both complex otherwise. In the first case we have an unstable node, in the second we have an unstable spiral (the real part of the two complex Eigenvalues $\frac{\text{tr}(J)}{2} = \frac{\rho}{2}$ is positive).

In the shallow lake example the condition $\text{tr}(J)^2 \geq 4 \det(J)$ translates into

$$\begin{aligned} \rho^2 &\geq c(-2c\beta + 6x\beta(2cx\beta + \rho) - f''(x)) \\ \Leftrightarrow \rho^2 &\geq c[(f'(x) - \alpha)2x\beta - (f''(x) + 2\beta c)]. \end{aligned}$$

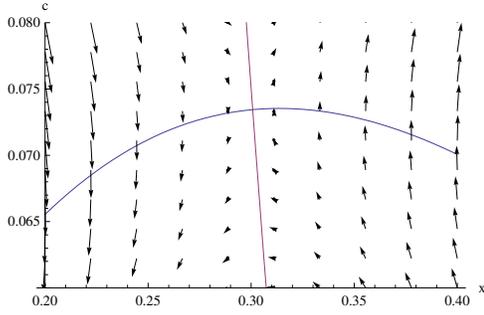
Numerical calculations for the $f(x) = \frac{x^2}{1+x^2}$ example show that both cases can occur. In the example plotted in figures 12 and 14 where $\alpha = 0.52$, $\beta = 1$, and $\rho = .03$ we find that the Eigenvalues corresponding to x^* at coordinates $(x, c) = (0.30, 0.07)$ are $\lambda_1^* = 0.32$ and $\lambda_2^* = -0.29$ with the respective Eigenvectors $(-0.95, -0.32)$ tangent to the unstable manifold and $(-0.96, 0.27)$ tangent to the stable manifold. For x^{**} at coordinates $(x, c) = (0.98, 0.02)$ we find the complex Eigenvalues $\lambda_1^{**} = (0.02 + 0.09i)$ and $\lambda_2^{**} = (0.02 - 0.09i)$ with the respective complex Eigenvectors $(1, 0.02 + 0.09i)$ and $(1, 0.02 - 0.09i)$. Thus, as we might guess from the phase diagram in Figure 14 we have an unstable spiral at x^{**} . Finally at x^{***} with coordinates $(x, c) = (1.5, 0.09)$ we find our second saddle point stable equilibrium with Eigenvalues $\lambda_1^{**} = 0.24$ and $\lambda_2^{**} = -0.21$ and corresponding Eigenvectors $(-0.90, -0.43)$ tangent to the unstable manifold and $(-1, -0.03)$ tangent to the stable manifold. Figure 16 depicts a close up of the three equilibria that we analyzed above.

Figure 17 gives an example where x^{**} is an unstable node rather than an unstable spiral. The corresponding parameter values are $\beta = 5$ and $\rho = 30$ (which might well be considered non-economic). From the diagram 16.b) it is not immediately clear whether or not the trajectories spiral out. However, calculating the corresponding Eigenvalues yields $\lambda_1^{**} = 0.18$ and $\lambda_2^{**} = 0.12$ with corresponding Eigenvectors $(-0.995, -0.097)$ and $(-0.999, -0.032)$.

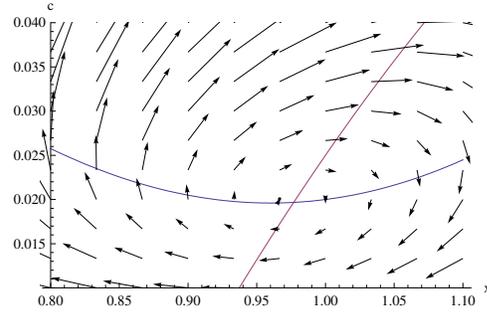
So far we have omitted one steady state. Recall that $\dot{c} = 0$ also had the solution $c = 0$. Thus, $(0, 0)$ is a fourth steady state. Evaluating the matrix J for $x = 0$ and $c = 0$ yields

$$J = \begin{pmatrix} f'(0) - \alpha & 1 \\ 0 & f'(0) - (\alpha + \rho) \end{pmatrix}.$$

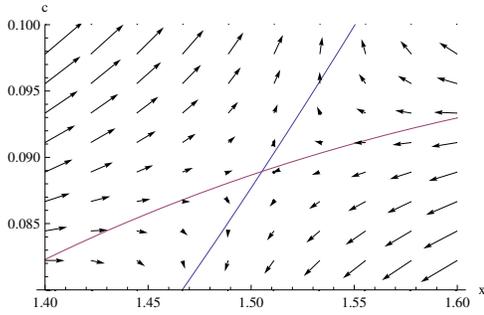
Note that we have to take the first form of J in equation (10.7) because in equation (10.7) we used equation (10.6) which is only true for the $c > 0$ branch of $\dot{c} = 0$. We know that (for the case we are analyzing) $\alpha > f'(0)$ so that the trace is negative and the determinant is positive. Hence $(0, 0)$ is a stable equilibrium. Precisely, the Eigenvalues can be read right off the



16.a) First Steady State



16.b) Second Steady State



16.c) Third Steady State

Figure 16: Close-up on the steady states in Figure 14. While the first and the third are saddle points, the second is unstable and spirals out. Note that the c -scale is ten times magnified with respect to the x -scale.

matrix J as $\lambda_1 = f'(0) - \alpha$ and $\lambda_2 = f'(0) - (\alpha + \rho)$. However, as we will see in the next section this steady state is not optimal.

10.5 The Optimal Consumption Paths

In a convex setting with a strictly concave Hamiltonian we know that any candidate that satisfies the necessary conditions is the unique global optimum of our control problem. In particular, if we have a path that converges into a steady state that satisfies the necessary conditions we can discard the other trajectories. However, in a non-convex setting uniqueness can no longer be taken for granted. Thus, we have to figure out which of the candidate paths leading to which of the steady states is (among the) best. In addition, we

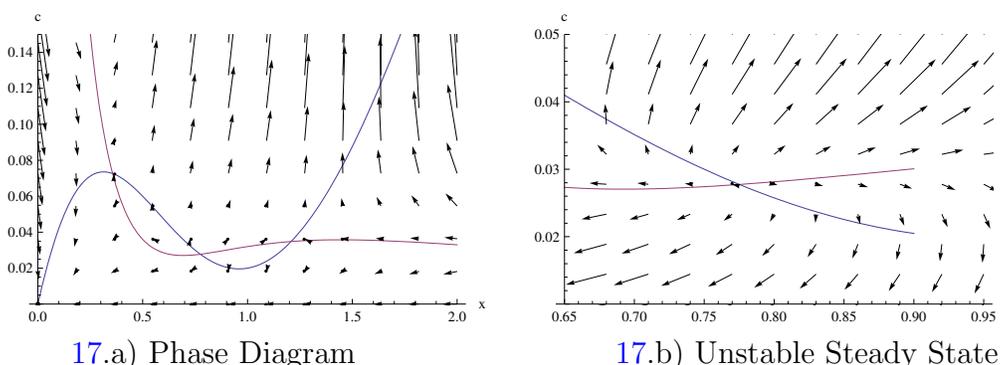


Figure 17: Example for a scenario where the unstable steady state x^{**} is an unstable node rather than a spiral. The corresponding parameters are $\alpha = .52$ as before, but a significant increase of $\beta = 5$ and $\rho = 30$.

would have to show that the diverging trajectories are not better than those converging into a steady state. We will only give the following economic intuition. We can observe from the phase diagram that all the paths not converging into a steady state approach (∞, ∞) . As the stock of phosphorous and the loading go to infinity, the quadratic loss from the increase in the stock will outgrow the logarithmic gains from loading the lake. Thus, unlimited dumping will not be optimal with the given benefit functions.⁸⁰ The economic intuition that paths converging to the stable steady state $(0, 0)$ cannot be optimal is straight forward. With a logarithmic utility from loading the lake it cannot be optimal for the social planner to reduce c to zero. Formally we can either show that the transversality condition is violated, or that utility along paths converging into the origin always yield lower welfare than paths converging into a steady state. First, let us assume that the transversality condition $\lim_{t \rightarrow \infty} e^{-\rho t} \lambda_t = 0$ is a necessary condition for our problem. It is easily observed that the Eigenvector to $\lambda_1 = f'(0) - \alpha$ is $(1, 0)$ implying that along this stable manifold the costate $\lambda = -\frac{1}{c}$ is negative infinity all the way and does not converge. For other trajectories, the loading c falls with the speed determined by the second Eigenvalue which, thus, is proportional to

⁸⁰This is only an intuition. The argument is not rigorous in neither considering the ratio of convergence speeds of c and x nor discounting. For a formal analysis in a similar setting see Wagener (2003).

$e^{-(\rho+(\alpha-f'(0))t)}$. Therefore

$$\lim_{t \rightarrow \infty} e^{-\rho t} \lambda(t) \propto - \lim_{t \rightarrow \infty} \frac{e^{-\rho t}}{e^{-\rho - (\alpha - f'(0)) t}} = - \lim_{t \rightarrow \infty} e^{(\alpha - f'(0)) t} = -\infty$$

and the necessary condition for an optimum is not satisfied.

Another way to show that trajectories converging to $(0, 0)$ cannot be optimal follows more directly the intuition that the logarithmic utility for loading the lake should not go to negative infinity for an optimal path. Instead of using the transversality condition to show that the paths converging to $(0, 0)$ are not satisfying the necessary conditions for an optimum, we simply show that the trajectory following the stable manifold leading into x^* always yields higher welfare. First, using the same argument as above, eliminate the trajectories starting with (and keeping) a zero loading because they yield negatively infinite utility at every point of time. Then, on any other trajectory in Figure 15 converging to the origin, let \bar{c} be an arbitrarily small value of the loading that the trajectory takes on at some point of time t_0 . Obviously, we can choose \bar{c} arbitrarily close to zero (by picking t_0 accordingly high). Without loss of generality set $t_0 = 0$. We denote the remaining part of the trajectory following the dynamic system 10.2 and 10.5 by $(x^o(t), c^o(t))_{t \in (0, \infty)}$ leading from \bar{c} and the stock at t_0 into the origin. Then the welfare along this path has the following upper bound

$$\int_0^\infty [\ln c^o(t) - \beta x^o(t)^2] e^{-\rho t} dt \leq \int_0^\infty \ln \bar{c} e^{-\rho t} dt = \frac{\ln \bar{c}}{\rho}. \quad (10.8)$$

Now we compare the welfare along this path with the welfare obtained from choosing the loading at t_0 in a way that brings us onto the stable manifold leading into the steady state (x^*, c^*) . Let us denote the according path by $(x^s(t), c^s(t))_{t \in (0, \infty)}$. Along the trajectory we know that $x^s(t) < x^*$ and $c^s(t) > c^*$ for all times so that we obtain the following lower bound for welfare along $(x^s(t), c^s(t))_{t \in (0, \infty)}$

$$\begin{aligned} \int_0^\infty [\ln c^s(t) - \beta x^s(t)^2] e^{-\rho t} dt &\geq \int_0^\infty [\ln c^* - \beta x^{*2}] e^{-\rho t} dt \\ &= \frac{\ln c^* - \beta x^{*2}}{\rho}. \end{aligned} \quad (10.9)$$

The steady state values x^* and c^* in equation (10.9) are some given finite numbers. On the other hand, equation (10.8) has to hold for arbitrarily small

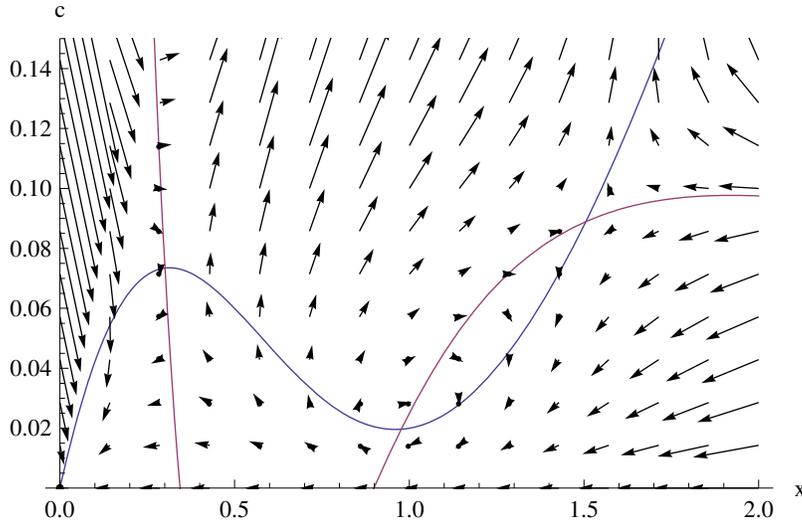


Figure 18: Stable manifolds for the saddle point stable steady states in Figure 12 and 14 ($\alpha = .52$, $\beta = 1$ and $\rho = 3\%$).

\bar{c} . Because the logarithm approaches negative infinity as \bar{c} approaches zero, we can pick a \bar{c} small enough to assure that

$$\frac{\ln \bar{c}}{\rho} < \frac{\ln c^* - \beta x^*}{\rho}.$$

Thus, the welfare along $(x^o(t), c^o(t))_{t \in (0, \infty)}$ is strictly less than the welfare along $(x^s(t), c^s(t))_{t \in (0, \infty)}$ and a trajectory leading into the origin cannot be optimal.

We are left with trajectories converging to the steady states x^* , x^{**} , and x^{***} . For the unstable steady state x^{**} that “trajectory” is only the point itself. If we draw the stable manifolds leading into the steady states x^* and x^{***} we obtain Figure 18. We denote the stable manifold leading into the oligotrophic steady state x^* by $T1$ and that leading into the eutrophic steady state x^{***} by $T2$. Under a slight abuse of notation we also denote by $T1(x)$ and $T2(x)$ the loading along trajectories $T1(x)$ and $T2(x)$ corresponding to the phosphorous stock x .⁸¹ Increasing the phosphorous stock from x^* moving

⁸¹The loading corresponding to some phosphorous stock along the trajectories is not everywhere unique because of the spiraling nature. For reasons that become obvious below, we only use the lower branch of the spiral of $T1(x)$ and the

along $T1$ (in inverted time) we find an intersection with the $\dot{x} = 0$ line. Let \bar{x}^s denote this (first respectively last in real time) intersection of trajectory $T1$ with the $\dot{x} = 0$ curve. Observe that because $\dot{x} = 0$ and $\dot{c} \neq 0$ at this point the trajectory is vertical. Similarly, decrease the phosphorous stock from x^{**} moving along $T2$ (in inverted time) and denote the first upcoming intersection with the $\dot{x} = 0$ line by \underline{x}^s .

For a phosphorous stock below \underline{x}^s there is only one candidate trajectory leading into a steady state. Thus, for an initial stock $x_0 < \underline{x}^s$ we take the path $T1$ that leads us into the oligotrophic steady state. Above \bar{x}^s there is also just one trajectory leading into a steady state. Thus, for an initial stock $x_0 > \bar{x}^s$ we take the path $T2$ that leads us into the eutrophic steady state. However, between the points \underline{x}^s and \bar{x}^s we can either choose a loading that puts us on the trajectory leading into the oligotrophic steady state, or we can choose a loading to put us on the trajectory leading into the eutrophic steady state. Both trajectories satisfy all the necessary conditions for an optimum.⁸² Close to the unstable steady state we even have the possibility to jump on either of the stable manifold at different points of the spiral (at an ‘earlier’ point or a ‘later’ point of the same trajectory that spirals out toward the saddle point stable steady state). First, we will show that it is always optimal to choose the loading in a way to be on the outermost (or ‘latest’) branch of the spiral leading into a given steady state. Then, we show that there exists a stock of phosphorous between \underline{x}^s and \bar{x}^s such that for an initial stock above that point it is optimal to converge into the eutrophic steady state, while for an initial stock below that point it is optimal to converge into the oligotrophic steady state. Such a point x^s is called a *Skiba point* after (Skiba 1978) contribution first showing the existence of such a point in a convex-concave production problem.

Proposition 6: It is not optimal to pick a point on the spiraling manifold leading into either of the steady states that comes back to the same phosphorous stock (at a later point on the trajectory with a different loading).

The proof of the proposition makes use of the following observation.

upper branch of the spiral of $T2$ when we write $T2(x)$.

⁸²Note that the transversality condition trivially holds in the steady states because $\lambda = -\frac{1}{c}$ is constant.

Lemma 3: Let a trajectory $(x^c, c^c) \equiv (x^c(t), c^c(t))_{t \in [0, \infty)}$ satisfy the necessary conditions for an optimum including the transversality condition $\lim_{t \rightarrow \infty} e^{-\rho t} H(x^c(t), c^c(t), \lambda^c(t)) = 0$. Then

$$H(x^c(0), c^c(0), \lambda^c(0)) = \rho \underbrace{\int_0^{\infty} [\ln c^c(t) - \beta x^c(t)^2] e^{-\rho t} dt}_{\equiv W(x^c(0), c^c(0))} .$$

Note that $\lambda(t)$ along the trajectory is uniquely determined by $\lambda(t) = -\frac{1}{c(t)}$. Moreover note that $W(x^c(0), c^c(0))$ is well defined (whenever it exists) as for any point in the (x, c) space there is (at most) a unique trajectory passing through it that satisfies the necessary conditions.

Proof of lemma 3: Let $H^{ps}(c, x, \lambda, t)$ denote the present value Hamiltonian. Then along a path satisfying the necessary conditions for an optimum it holds

$$\begin{aligned} \frac{dH^{ps}}{dt} &= \frac{dH^{ps}}{dc} \frac{dc}{dt} + \frac{dH^{ps}}{dx} \frac{dx}{dt} + \frac{dH^{ps}}{d\lambda} \frac{d\lambda}{dt} + \frac{\partial H^{ps}}{\partial t} \\ &= 0 + \frac{dH^{ps}}{dx} \frac{dH^{ps}}{d\lambda} + \frac{dH^{ps}}{d\lambda} \left(-\frac{dH^{ps}}{dx} \right) + \frac{\partial H^{ps}}{\partial t} = \frac{\partial H^{ps}}{\partial t} . \end{aligned}$$

Moreover at $t = 0$ the present value and the current value Hamiltonian coincide so that

$$\begin{aligned} H(c^c(0), x^c(0), \lambda^c(0)) &= H^{ps}(c^c(0), x^c(0), \lambda^c(0), 0) \\ &= \lim_{t \rightarrow \infty} H^{ps}(c^c(t), x^c(t), \lambda^c(t), t) - \int_0^{\infty} \frac{\partial H^{ps}}{\partial t} dt \\ &= \lim_{t \rightarrow \infty} e^{-\rho t} H(c^c(t), x^c(t), \lambda^c(t)) - \\ &\quad \int_0^{\infty} \frac{\partial}{\partial t} [\ln(c^c(t) - \beta x^c(t)^2)] e^{-\rho t} dt \\ &= 0 + \int_0^{\infty} \rho [\ln(c^c(t) - \beta x^c(t)^2)] e^{-\rho t} dt \\ &= \rho W(x^c(0), c^c(0)) \end{aligned}$$

□

Note that the above relation only makes sense because a point $(x^c(0), c^c(0))$ in phase space uniquely determines $W(x^c(0), c^c(0))$ (see above) as well as $H(c^c(0), x^c(0), \lambda^c(0))$ by optimality implying $\lambda^c(0) = -\frac{1}{c^c(0)}$. If we analyze a change in the phase space location (x, c) we have to pick and vary $\lambda^c(0)$ according to the optimality condition $\lambda^c(0) = -\frac{1}{c^c(0)}$ for the above relation to hold.

Proof of Proposition 6: Let $c_0 = c^c(0)$ and $x_0 = x^c(0)$. By lemma 3 we have

$$\begin{aligned} \frac{\partial W}{\partial c_0} &= \frac{1}{\rho} \frac{dH}{dc_0} \Big|_{\bar{x}_0, \lambda(c_0)} = \frac{1}{\rho} \underbrace{\frac{\partial H}{\partial c_0}}_{=0} + \frac{1}{\rho} \frac{\partial H}{\partial \lambda_0} \frac{d\lambda_0}{dc_0} \\ &= \frac{1}{\rho} \frac{d\lambda}{dc_0} X(c_0, x_0) = \frac{1}{\rho} \frac{1}{c_0^2} X(c_0, x_0) \end{aligned} \tag{10.10}$$

whenever $\lim_{t \rightarrow \infty} e^{-\rho t} H(x^c(t), c^c(t), \lambda^c(t))$. Expression (10.10) is positive above the $\dot{x} \equiv X(c, x) = 0$ line and negative below it. Thus, for a given initial stock we can increase welfare by moving further out on the spiral away from the $\dot{x} = 0$ line. Thus, if any of the intersections of the ‘out-spiraling’ stable manifolds with the $x = x_0$ line is optimal, then it must be on either of the outermost branches of the spiral. For a manifold leading into a steady state with finite value of H the condition $\lim_{t \rightarrow \infty} e^{-\rho t} H(x^c(t), c^c(t), \lambda^c(t))$ is obviously satisfied.⁸³ \square

Lemma 4: With an initial phosphorous stock of \underline{x}^s it is optimal to take trajectory $T1$ into the oligotrophic steady state. With an initial phosphorous stock of \bar{x}^s it is optimal to take trajectory $T2$ into the eutrophic steady state.

Proof of Lemma 4: In \underline{x}^s either $T1$ or $T2$ must be optimal. By equation (10.10) the Hamiltonian takes on a higher value at $(\underline{x}^s, T1(\underline{x}^s))$ than at $(\underline{x}^s, T2(\underline{x}^s))$ because moving down from $T2(\underline{x}^s)$ where $X(c, x) = 0$ the Hamiltonian increases. Thus, by lemma 3 we know that the path from $(\underline{x}^s, T1(\underline{x}^s))$ along the trajectory $T1$ into the oligotrophic steady state yields

⁸³Why can't we further increase welfare by moving away from the $\dot{x} = 0$ curve beyond the stable manifolds leading into steady states x^* or x^{***} ?

a higher welfare than the path $(\underline{x}^s, T2(\underline{x}^s))$ along the trajectory $T2$ into the eutrophic steady state, i.e. $W(T1(\underline{x}^s), \underline{x}^s) > W(T2(\underline{x}^s), \underline{x}^s)$. Similarly, at \bar{x}^s the Hamiltonian is larger in $(\bar{x}^s, T2(\bar{x}^s))$ than in $(\bar{x}^s, T1(\bar{x}^s))$ as by equation (10.10) the Hamiltonian increases from moving up from $T1(\bar{x}^s)$ where $X(c, x) = 0$. Thus, in \bar{x}^s welfare along trajectory $T2$ is higher, i.e. $W(T2(\bar{x}^s), \bar{x}^s) > W(T1(\bar{x}^s), \bar{x}^s)$. \square

The question which trajectory is optimal in the interval $(\underline{x}^s, \bar{x}^s)$ can be answered by calculating the so called Skiba point. We will only show that a Skiba point x^s with the following properties exists.

Proposition 7: There exists a point $x^s \in (\underline{x}^s, \bar{x}^s)$ such that

- for an initial stock $x_0 < x^s$ it is optimal to take trajectory $T1$ into the oligotrophic steady state.
- for an initial stock $x_0 > x^s$ it is optimal to take trajectory $T2$ into the eutrophic steady state.
- for an initial stock $x_0 = x^s$ the social planner is indifferent between trajectories $T1$ and $T2$.

Proof of Proposition 7: For $x \in [\underline{x}^s, \bar{x}^s]$ define the function $\Delta W(x) = W(T1(x), x) - W(T2(x), x)$. In the proof of Lemma 4 we have shown

$$\Delta W(\underline{x}^s) = W(T1(\underline{x}^s), \underline{x}^s) - W(T2(\underline{x}^s), \underline{x}^s) > 0 \quad \text{and}$$

$$\Delta W(\bar{x}^s) = W(T1(\bar{x}^s), \bar{x}^s) - W(T2(\bar{x}^s), \bar{x}^s) < 0.$$

Moreover, we know that $\frac{d}{dx}W = \lambda$ and thus $\frac{d}{dx}\Delta W = \frac{d}{dx}W(T1(x), x) - \frac{d}{dx}W(T2(x), x) = \lambda(T1(x)) - \lambda(T2(x)) = \frac{1}{T2(x)} - \frac{1}{T1(x)} < 0$. Hence somewhere between \underline{x}^s and \bar{x}^s there is a unique point x^s satisfying $\Delta W(x^s) = 0$. Here $W(T1(x), x) = W(T2(x), x)$ and both trajectories are optimal. For $x < x^s$ we have $\Delta W(x) > 0$ and, thus, $W(T1(x), x) > W(T2(x), x)$ and $T1$ optimal. For $x > x^s$ we have $\Delta W(x) < 0$ and, thus, $W(T2(x), x) > W(T1(x), x)$ and $T2$ optimal. \square

That finishes our analysis of the socially optimal control of the shallow lake. See Mäler et al. (2003) and Kossioris, Plexousakis, Xepapadeas & de Zeeuw K.-G. Mäler (2008) for a game theoretic extension and ?? [Wagener 2003] for relating Skiba points and bifurcations.

11 Dynamic Programming in Continuous Time

This section introduces dynamic programming in continuous time. First, we take the continuous time limit of the dynamic programming equation we used in discrete time. We arrive at a partial differential equation called the Hamilton Jacobi Bellman equation. Second, we derive the Hamilton Jacobi Bellman equation more formally as a sufficient condition for an optimum. We then relate the continuous time dynamic programming approach to the Maximum Principle. Finally, we solve a linear quadratic example as an example of an infinite horizon autonomous problem. For these cases we can transform the partial differential equation into an ordinary differential equation.

11.1 The Continuous Time Limit of the Dynamic Programming Equation

Recall the dynamic programming equation (DPE) we encountered in discrete time. We can write the DPE as

$$J(x_t, t) = \max_{y_t} U(x_t, y_t, t) + J(x_{t+1}, t + 1)$$

s.t. $x_{t+1} = g(x_t, y_t, t)$, x_0 given .

In order to take the continuous time limit we replace the unit time step by the time step Δt . In this step we have to be careful to identify those terms that are densities and have to be multiplied with the time step.⁸⁴ In the evaluation function the utility is a density characterizing utils per time step (or alternatively profits per time step), while the value function is an absolute measure of welfare (or profits). For the constraint we pick a convenient form for the discrete time equation of motion that can be turned straight forwardly into its continuous time analog

$$J(x_t, t) = \max_{y_t} U(x_t, y_t, t) \Delta t + J(x_{t+\Delta t}, t + \Delta t)$$

s.t. $x_{t+\Delta t} = x_t + f(x_t, y_t, t) \Delta t$, x_0 given .

⁸⁴We already observed that we had to turn “hidden ones” into Δt in section 7 when we took the continuous time limit for the equation of motion for a pollution stock (see 7.3).

You can verify that for example the equation of motion for our stock pollution problem in equation (7.3) in section 7.3 is of this form. In the continuous time limit the constraint turns into the same form for the equation of motion as in our continuous time formulation of the intertemporal optimization problem in section 7

$$\dot{x} = \lim_{t \rightarrow \infty} \frac{x_{t+\Delta t} - x_t}{\Delta t} = f(x_t, y_t, t) .$$

Expanding the value function in Δt delivers

$$\begin{aligned} J(x_t, t) = \max_{y_t} U(x_t, y_t, t) \Delta t + J(x_t, t) + \frac{\partial J(x_t, t)}{\partial x_t} \frac{x_{t+\Delta t} - x_t}{\Delta t} \Delta t \\ + \frac{\partial J(x_t, t)}{\partial t} \Delta t + o(\Delta t) . \end{aligned}$$

Substituting the constraint into the dynamic programming equation and canceling the value function leaves us with

$$0 = \max_{y_t} U(x_t, y_t, t) \Delta t + \frac{\partial J(x_t, t)}{\partial x_t} f(x_t, y_t, t) \Delta t + \frac{\partial J(x_t, t)}{\partial t} \Delta t + o(\Delta t) .$$

We observe that the term involving the (partial) time derivative of the value function is independent of the control. Rearranging it to the other side, dividing the equation by Δt , and taking the limit $\Delta t \rightarrow 0$ yields

$$-\frac{\partial J(x_t, t)}{\partial t} = \max_{y_t} U(x_t, y_t, t) + \frac{\partial J(x_t, t)}{\partial x_t} f(x_t, y_t, t) . \quad (11.1)$$

Equation (11.1) is called the Hamilton-Jacobi-Bellman (HJB) equation. It is a first order partial differential equation, i.e. it depends of the first partial derivatives in more than one variable. With a finite time horizon we generally have a boundary condition of the form $J(x_T, T) \stackrel{!}{=} \Gamma(x)$ with $\Gamma(x)$ characterizing the scrap value of the stock at the end of the planning horizon. Note that one of the assumptions involved in our derivation is that the value function is continuously differentiable, otherwise the above first order approximation would be meaningless.

11.2 A Neat Derivation of the HJB Equation as a Sufficient Condition for an Optimum

Let us define the maximized Hamiltonian as

$$\begin{aligned}\hat{H}(x_t, \lambda_t, t) &= \max_y H(x_t, y, \lambda_t, t) \\ &= \max_y U(x_t, y, t) + \lambda_t f(x_t, y, t) .\end{aligned}$$

Then, we can restate the Hamilton Jacobi Bellman equation (11.1) in terms of the maximized Hamiltonian as⁸⁵

$$\frac{\partial J(x, t)}{\partial t} + \hat{H}\left(x, \frac{\partial J(x, t)}{\partial x}, t\right) = 0 . \quad (11.2)$$

We assume that there exists a continuously differentiable function $J(x, t)$ that solves the partial differential equation (11.2) (which coincides with equation 11.1) subject to the boundary condition $J(x_T, T) \stackrel{!}{=} \Gamma(x)$. The function $\Gamma(x)$ represents the scrap value of the stock variable and is more formally known as the Mayer term. Integrating the total time derivative of the continuously differentiable function $J(x_t, t)$ along a feasible path $(x_t, y_t)_{t \in [t_0, T]}$ over time implies

$$J(x_T, T) = J(x_{t_0}, t_0) + \int_{t_0}^T \left(\frac{\partial J(x_t, t)}{\partial x} f(x_t, y_t, t) + \frac{\partial J(x_t, t)}{\partial t} \right) dt .$$

⁸⁵We can define the left hand side of equation (11.2) as a new Hamiltonian arising under a suitable (canonical) transformation of variables. In these new variables the differential equations defining the necessary condition of the maximum principle introduced in section 7 become trivial and the new state and co-state variables are constant. Transforming the Hamiltonian canonical equations, i.e. the differential equation for the state and the co-state variable in the maximum principle, in a way that makes the differential equations trivial is a motivation for the Hamilton-Jacobi equation in Physics and was derived before Richard Bellman developed it in the current context.

Moreover, along any such feasible path we find that

$$\begin{aligned}
& \int_{t_0}^T U(x_t, y_t, t) dt + J(x_T, T) \\
&= \int_{t_0}^T U(x_t, y_t, t) + \frac{\partial J(x_t, t)}{\partial x} f(x_t, y_t, t) + \frac{\partial J(x_t, t)}{\partial t} dt + J(x_{t_0}, t_0) \\
&= \int_{t_0}^T H\left(x_t, y_t, \frac{\partial J(x_t, t)}{\partial x}, t\right) + \frac{\partial J(x_t, t)}{\partial t} dt + J(x_{t_0}, t_0) \\
&\leq \int_{t_0}^T \underbrace{\hat{H}\left(x_t, \frac{\partial J(x_t, t)}{\partial x}, t\right)}_{=0 \text{ because } J \text{ solves HJB}} + \frac{\partial J(x_t, t)}{\partial t} dt + J(x_{t_0}, t_0) \\
&= J(x_{t_0}, t_0),
\end{aligned}$$

with equality for the optimal control path y^* . We have shown that a path $(x_t^*, y_t^*)_{t \in [t_0, T]}$ satisfying $H\left(x_t^*, y_t^*, \frac{\partial J(x_t^*, t)}{\partial x_t^*}, t\right) = \hat{H}\left(x_t^*, \frac{\partial J(x_t^*, t)}{\partial x_t^*}, t\right)$ maximizes the objective function within the set of all feasible paths. Moreover, the function $J(x_{t_0}, t_0)$ solving the Hamilton Jacobi Bellman equation for a given terminal scrap value function $\Gamma(x)$ gives us the maximal value of our program.

11.3 Relation to the Maximum Principle

There is an immediate relation between the Maximum principle discussed in section 7 and the HJB equation. Let us define the value of a marginal unit of the stock at time t by

$$\lambda_t = \frac{\partial J(x_t, t)}{\partial x_t}. \quad (11.3)$$

Then, we can rewrite the HJB equation (11.2) in terms of the maximized Hamiltonian as

$$-\frac{\partial J(x_t, t)}{\partial t} = \hat{H}(x_t, \lambda_t, t). \quad (11.4)$$

The Hamiltonian characterizes the (partial/direct) time change of the value function. Equation (11.4) confirms our earlier interpretation of the Hamilto-

nian in section 7 as the overall contribution to welfare in terms of immediate gratification and change in stock value in period t . Precisely, equation (11.2) states that the decrease in the value function from t to $t+dt$ is proportional to the Hamiltonian, i.e. to the lost contribution to the overall value by omitting period t . Finally, note that a free terminal state in the maximum principle corresponds to a terminal value function $\Gamma(x)$ that is independent of x in the HJB formulation. Then we have $\lambda_t = \frac{\partial J(x_T, T)}{\partial x_T} = \frac{\partial \Gamma(x)}{\partial x} = 0$ confirming our earlier interpretation of the transversality condition that if the terminal state was free, the stock at the terminal time should be of no value to us.

We can also derive the co-state equation of motion of the Maximum Principle from the HJB. Differentiate equation (11.2) with respect to the state variable to obtain

$$-\frac{\partial^2 J(x_t, t)}{\partial x_t \partial t} = \left. \frac{\partial \hat{H}(x_t, \lambda_t, t)}{\partial x_t} \right|_{\lambda_t = \frac{\partial J(x_t, t)}{\partial x_t}} + \left. \frac{\partial \hat{H}(x_t, \lambda_t, t)}{\partial \lambda_t} \right|_{\lambda_t = \frac{\partial J(x_t, t)}{\partial x_t}} \frac{\partial^2 J(x_t, t)}{\partial x_t^2}$$

Under the assumption that J is twice continuously differentiable we can exchange the order of the time and the state partial derivatives of J . From the envelope theorem we know that the partial of the maximized Hamiltonian (suppressing the control) simply gives us back the partial of the full Hamiltonian (containing the control has an explicit variable), so that

$$-\frac{\partial^2 J(x_t, t)}{\partial t \partial x_t} = \left. \frac{\partial \hat{H}(x_t, \lambda_t, t)}{\partial x_t} \right|_{\lambda_t = \frac{\partial J(x_t, t)}{\partial x_t}} + f(x_t, y_t, t) \frac{\partial^2 J(x_t, t)}{\partial x_t^2}. \quad (11.5)$$

Moreover, along a feasible path, the total time derivative of λ_t is by definition (11.3)

$$\begin{aligned} \dot{\lambda}_t &= \frac{d}{dt} \frac{\partial J(x_t, t)}{\partial x_t} = \frac{\partial^2 J(x_t, t)}{\partial t \partial x_t} + \frac{\partial^2 J(x_t, t)}{\partial x_t^2} \dot{x}_t \\ &= \frac{\partial^2 J(x_t, t)}{\partial t \partial x_t} + \frac{\partial^2 J(x_t, t)}{\partial x_t^2} f(x_t, y_t, t) \end{aligned}$$

Combining the latter equation with equation (11.5) we obtain the equation of motion for the co-state variable featuring as a necessary condition in the

Maximum Principle

$$\dot{\lambda}_t = -\frac{\partial H(x_t, y_t, \lambda_t, t)}{\partial x_t} . \quad (11.6)$$

We used again the envelope theorem in order to replace the maximized Hamiltonian by the original Hamiltonian evaluated along the optimal path. Obviously, equation (11.6) has to hold jointly with the condition that we maximize the Hamiltonian over the control.

11.4 Autonomous Problems with Infinite Time Horizon, Example and General Approach

In general the HJB equation is difficult to solve. We obtain a significant simplification if we can separate the time dependent part in equation (11.4) from the state dependent part. The following is one of the simplest non-trivial examples:

$$\max \int_0^\infty -\frac{ay^2 + bx^2}{2} \exp[-\delta t] dt \quad \text{s.t.} \quad \dot{x}_t = y ,$$

x_0 given, $a, b > 0$, and the scrap value assumed to be zero. Then the HJB becomes

$$\begin{aligned} -J_t(x, t) &= \max_y -\frac{ay^2 + bx^2}{2} \exp[-\delta t] + J_x(x, t)y \\ \Leftrightarrow -J_t(x, t) \exp[\delta t] &= \max_y -\frac{ay^2 + bx^2}{2} + J_x(x, t) \exp[\delta t]y . \end{aligned}$$

We observe that the only time dependence in the differential equation is the exponential adjacent to the value function. Thus, we try the form $J(x, t) = V(x) \exp[-\delta t]$:

$$\begin{aligned} \delta V(x) \exp[-\delta t] \exp[\delta t] &= \max_y -\frac{ay^2 + bx^2}{2} + V_x(x) \exp[-\delta t] \exp[\delta t]y \\ \Leftrightarrow \delta V(x) &= \max_y -\frac{ay^2 + bx^2}{2} + V_x(x)y . \end{aligned} \quad (11.7)$$

The result is an ordinary differential equation (ODE) in the state variable. Without the max operator equation (11.7) would be a first order linear differential equation as the derivative of the value function enters the equation only linearly. Unfortunately, carrying out the maximization transforms it into a non-linear first order ODE:

$$\begin{aligned} & \max_y -\frac{ay^2 + bx^2}{2} + V_x(x)y \\ \Rightarrow -ay + V_x(x) &= 0 \quad \Rightarrow \quad y = \frac{V_x(x)}{a}, \end{aligned} \quad (11.8)$$

implying the (inhomogenous) differential equation

$$\begin{aligned} \delta V(x) &= -\frac{V_x(x)^2}{2a} - \frac{bx^2}{2} + \frac{V_x(x)^2}{a} \\ \Leftrightarrow \delta V(x) - \frac{V_x(x)^2}{2a} &= -\frac{bx^2}{2}. \end{aligned} \quad (11.9)$$

Equation (11.9) is called a Riccati differential equation and it is about the simplest Riccati differential equation you can get in this context. The easiest approach to solving this equation is a good guess. Recall that we already know from the discrete time case that a good shot for a trial is a quadratic value function $V(x) = \alpha x^2$, transforming equation (11.9) into

$$\begin{aligned} \delta \alpha x^2 - \frac{2\alpha^2 x^2}{a} &= -\frac{bx^2}{2} \\ \Leftrightarrow \left(\delta \alpha - \frac{2\alpha^2}{a} + \frac{b}{2} \right) x^2 &= 0, \end{aligned}$$

which is satisfied for all x whenever

$$\begin{aligned} \alpha^2 - \frac{a\delta}{2}\alpha &= +\frac{ab}{4} \\ \Leftrightarrow \alpha &= \frac{a}{4} \left(\delta \pm \sqrt{\delta^2 + \frac{4b}{a}} \right). \end{aligned}$$

So our trial solution worked and

$$V(x) = \alpha x^2 = \frac{a}{4} \left(\delta \pm \sqrt{\delta^2 + \frac{4b}{a}} \right) x^2.$$

Given that our value function has to be everywhere (weakly) negative we know that the negative root is the correct one and we have

$$J(x, t) = V(x) \exp[-\delta t] = \frac{a}{4} \left(\delta - \sqrt{\delta^2 + \frac{4b}{a}} \right) x^2 \exp[-\delta t] .$$

Furthermore we obtain the optimal control in feedback form from equation (11.8)

$$y(x) = \frac{V_x(x)}{a} = \frac{2\alpha x}{a} = \frac{1}{2} \left(\delta - \sqrt{\delta^2 + \frac{4b}{a}} \right) x .$$

We obtain the solution for the time path of the state variable by integrating the equation of motion for the optimal control

$$\dot{x}_t = \frac{2\alpha x}{a} \quad \Rightarrow \quad x_t = x_0 \exp \left[\frac{2\alpha t}{a} \right] ,$$

which converges to zero given that $\alpha < 0$.

Our simplistic equation of motion $\dot{x}_t = y$ bought us equation (11.9), the fact that the value function only depends on the stock quadratically, and that we only have to solve a single quadratic equation for α . You can add linear terms to the objective function and make the equation of motion for the state a more general linear function and still follow the same procedure as above. Then the Ricatti differential equation (11.9) will look slightly more complicated and you will have to set up a trial solution that also incorporates linear and affine parts. Collecting terms by the power of x gives you three generally dependent algebraic equations that you can solve for the affine, linear, and quadratic coefficient in the value function.

The transformation of the partial differential equation into an ordinary differential equation in the example works more generally for autonomous problems with an infinite time horizon:

$$\max \int_0^\infty u(x_t, y_t) \exp[-\delta t] dt \quad \text{s.t. } \dot{x}_t = f(x_t, y_t) ,$$

x_0 given. Define $J(x_t, t) = V(x) \exp[-\delta t]$ and, assuming once more sufficient differentiability of the value function, use it as a trial solution for the HJB.

We obtain

$$\begin{aligned} \delta V(x) \exp[-\delta t] &= \max_y u(x_t, y_t) \exp[-\delta t] + V_x(x) \exp[-\delta t] f(x_t, y_t) \\ \Rightarrow \delta V(x) &= \max_y u(x_t, y_t) + V_x(x) f(x_t, y_t) , \end{aligned} \quad (11.10)$$

which once more is an ordinary differential equation. Equation (11.10) has an immediate and insightful interpretation. It tells us that along an optimal path the overall welfare is proportional to the value of the sum of the immediate welfare gain in terms of u and the change in stock weighted with the proper accounting price $V_x(x)$. The proportionality factor is our pure rate of time preference.

So far we have not paid any attention to the infinite horizon boundary and transversality conditions. The issues with necessary and sufficient conditions in the infinite time horizon are similar to those for the maximum principle. We don't generally obtain a necessary transversality conditions without further assumptions. The sufficient condition for an optimum stated in section 7 for the maximum principle translates by rewriting the shadow value in terms of the current value value function into

$$\lim_{t \rightarrow \infty} \exp(-\delta t) \frac{\partial V(x)}{\partial x} x_t \geq 0$$

for all feasible paths. However, I have not yet identified a careful proof under which circumstances it is a sufficient condition in continuous time dynamic programming. The proof of the HJB as a sufficient condition in section 11.2 suggests another sufficient condition. Observe that the inequality holds also in the limit of an infinite time horizon if

$$\lim_{t \rightarrow \infty} J(x_t, t) = 0$$

for all feasible paths. Translated into the time independent function V we have the sufficient condition

$$\lim_{t \rightarrow \infty} \exp[-\delta t] V(x_t) = 0 .$$

12 Event Uncertainty

A class of problem involving uncertainty can be modeled using either deterministic optimal control or stochastic control of a jump process – defined as a stochastic process whose realization can be a discontinuous function of time. We illustrate both of these approaches using a problem in which the probability of a random event depends on the value of a state variable that the decision maker controls. For example, the probability of a climate-related catastrophe might depend on the stock of GHGs. At a point in time this stock is predetermined, so the decision maker is not able to affect the probability of the catastrophe over the next short interval of time. However, by choosing an emissions path, the decision maker chooses the trajectory of the stock of GHGs, and thereby influences the probability of catastrophe in the future.

12.1 Preliminaries

The building block of this kind of model is the hazard function, which gives the probability of an event (such as a catastrophe) over the next small interval of time, conditional on the event not having yet occurred. Denote τ as a random variable, the time at which this event occurs, and denote the cumulative distribution function as $F(t) = \Pr\{\tau \leq t\}$. We assume that this function is differentiable, so $f(t) \equiv \frac{dF(t)}{dt}$ is the probability density function. For small dt , $f(t)dt$ is approximately the probability that the event occurs during the interval of time $(t, t + dt)$.

If A and B are two random events, then using the rule for conditional probabilities we have

$$\Pr(A | B) = \frac{\Pr(A \cap B)}{\Pr(B)}.$$

This formula states that the probability of event A , given that event B has occurred, equals the probability that both A and B occur divided by the unconditional probability that B occurs. Think of event A as being “the disaster occurs over the interval $(t, t + dt)$ ” and event B being “the probability does not occur by time t ”. Applying the formula for conditional

probabilities, we have

$$\begin{aligned} & \Pr \{ \text{event occurs during } (t + dt) \mid \text{event has not occurred by } t \} \quad (12.1) \\ & = h(t)dt \equiv \frac{f(t)dt}{1 - F(t)}. \end{aligned}$$

The function $h(t)$ is the hazard rate. If we interpret $dt = 1$ as an “instant”, then the hazard rate equals the probability that the event occurs during the next instant, given that it has not yet occurred.

One of the earliest applications of event uncertainty involves the problem of life-time consumption when death occurs at a random time (cite Yari). For this problem, the probability of the “event”, death, is exogenous: the decision maker cannot alter the future risk, but he can adjust his consumption decisions to take the risk into account. We use this problem to illustrate the method of analysis, before turning to the problem of interest, in which the decision maker affects future risk.

If death occurs at time t , the present discounted value of the agent’s welfare at the initial time 0 is

$$\int_0^t e^{-rs} U(c_s) ds + d(t) B(k(t)). \quad (12.2)$$

In this problem, the instantaneous flow of utility is the increasing and concave function $U(c)$ and the pure rate of time preference (the discount rate) is r . The amount of wealth that the decision maker bequeaths to his heirs is $k(t)$, and $B(k)$ is the utility that the decision maker obtains from this bequest; $d(t)$ discounts utility of the bequest at time t back to the first period. The function $d(t)$ could have a variety a shapes. For example, it might be small for t close to 0 and close to T but large for intermediate values of t ; this shape would arise if the decision maker does not worry about leaving a bequest before he has a family (small t) or after the family is grown (large t), but is concerned about the bequest during his middle age, when his family is young.

The decision maker is able to invest wealth at a constant rate of return i and receives an exogenous income stream $Y(t)$, so his wealth obeys the differential equation

$$\dot{k}(t) = ik(t) + Y(t) - c(t), \quad k(0) = k_0, \text{ given.} \quad (12.3)$$

The probability distribution function for the random time of death, t , is $F(t)$, with $F(0) = 0$ and $F(T) = 1$. The first equality states that the decision maker is certainly alive at time 0, and the second states that he is certainly dead by time T . The decision maker is risk neutral and maximizes the expectation of lifetime utility

$$\begin{aligned} & \max_{\{c\}} E_t \int_0^t e^{-rs} U(c_s) ds + d(t) B(k(t)) \\ & = \max_{\{c\}} \int_0^T \left[\int_0^t e^{-rs} U(c_s) ds + d(t) B(k(t)) \right] f(t) dt. \end{aligned} \quad (12.4)$$

The term in square brackets is lifetime utility conditional on death occurring at time t ; we multiply this term times the probability of death at time t , $f(t)$, and integrate over t to obtain the expectation of lifetime utility.

Denote

$$z(t) = \int_0^t e^{-rs} U(c_s) ds \implies \frac{dz}{dt} = e^{-rt} U(c_t) \quad (12.5)$$

and integrate the first term on the right side of equation 12.4 by parts, using $f(t) = F'(t)$ to obtain

$$\begin{aligned} \int_0^T z(t) f(t) dt &= \int_0^T z(t) F'(t) dt = z(t) F(t) \Big|_0^T - \int_0^T F(t) \frac{dz}{dt} dt \\ &= \int_0^T e^{-rt} U(c_t) dt - \int_0^T F(t) e^{-rt} U(c_t) dt \\ &= \int_0^T (1 - F(t)) e^{-rt} U(c_t) dt. \end{aligned}$$

The third equality uses the fact that $z(0) = 0$ and $F(T) = 1$. Using the expression after the last equality, we write the expected payoff in equation 12.4 as

$$\max_{\{c\}} \int_0^T [(1 - F(t)) e^{-rt} U(c_t) + d(t) B(k(t)) f(t)] dt. \quad (12.6)$$

At each point in time, t , if the agent is alive he obtains the present value utility of consumption $e^{-rt} U(c_t)$, and if he dies at that point in time he obtains the present value utility of the bequest $d(t) B(k(t))$. The probability of the first event is $(1 - F(t))$ and the probability of the second event is $f(t)$.

The expected payoff is the integral over time of the two payoffs, weighted by their probabilities. Equation 12.3 is the constraint to the optimization problem.

These manipulations transform the original problem involving risk into an almost standard deterministic problem. Recall that in a deterministic problem the optimal control rule can be written either as a function of time, i.e. as an open loop control rule, or as function of the state variable (and possibly also of time), i.e. as a feedback rule. The two solutions are equivalent because in a deterministic setting the decision maker does not acquire new information as time goes on. The control problem consisting of expression 12.6 and the constraint 12.3 appears to be a deterministic problem, because the random time of death has been “concentrated out”, i.e. removed by taking expectations.

Thus, it appears that the decision maker can choose at the initial time $t = 0$ the entire trajectory of consumption. This appearance is correct, provided that we interpret the optimal consumption trajectory as conditioned on the event that the decision maker is still alive when he attempts to carry out the proposed plan. This caveat may seem like a semantic quibble, but it is particularly important when we consider an infinite horizon version of this problem.

It is also worth noting that the optimal solution to this problem is *time consistent*. In this context, time consistency means that if the decision maker follows the plan that is optimal at time $t = 0$ up to some time $t = s > 0$ and then reoptimizes, the optimal solution is the continuation of the plan that he chose at time $t = 0$. If he is still alive at time $s > 0$ then his objective is

$$\max_{\{c\}} \frac{1}{1 - F(s)} \int_s^T [(1 - F(t)) e^{-rt} U(c_t) + d(t) B(k(t)) f(t)] dt.$$

The information that he is still alive at time s causes the density $f(t)$ to be replaced by the conditional density $\frac{f(t)}{1 - F(s)}$ and the probability $1 - F(t)$ to be replaced by a conditional probability $\frac{1 - F(t)}{1 - F(s)}$ but these changes amount to nothing more than dividing through by a constant, and therefore they do not alter the optimal plan.

Define the integrand as

$$G(t, k, c) = (1 - F(t)) e^{-rt} U(c_t) + d(t) B(k(t)) f(t)$$

and write the Hamiltonian as

$$H = G + \lambda (ik(t) + y(t) - c(t)).$$

The first order condition for maximization of the Hamiltonian is

$$\frac{\partial H}{\partial c} = G_c - \lambda = e^{-rt} (1 - F(t)) U'(c_t) = \lambda, \quad (12.7)$$

which states that the present value of marginal utility at time t , multiplied by the probability that the decision maker is still alive at time t , equals the shadow value of wealth. The costate equation is

$$\dot{\lambda} = -G_k - i\lambda. \quad (12.8)$$

Using the menu for analyzing this kind of problem, developed in Chapter 7, we differentiate equation 12.7 with respect to time, use equation 12.8 to eliminate $\dot{\lambda}$, and then use equation 12.7 to eliminate λ . After dividing through by c , we have the optimal growth rate in consumption equal to

$$\frac{\dot{c}}{c} = \frac{(i - r - h)}{RRA(c)} + \frac{h}{RRA(c)} \left[\frac{d(t)B'(k)}{e^{-rt}U'(c)} \right], \quad (12.9)$$

where $RRA(c) \equiv \frac{-U''c}{U'} > 0$ is the relative risk aversion and the hazard rate h is defined in equation 12.1. For example, with $U(c) = \ln c$, $RRA(c) = 1$. [Chapter 3.]

In the risk-free analog to this problem where the time of death is known to be T , the Euler Equation simplifies to

$$\frac{\dot{c}}{c} = \frac{(i - r)}{RRA(c)}. \quad (12.10)$$

In this risk-free case, the transversality condition is $\lambda(T) = d(T)B'(k_T)$, which states that the shadow value of wealth at the time of death equals the present value of the marginal value of the bequest. In the problem with risk, equation 12.7 implies, using $F(T) = 1$, that $\lambda(T) = 0$. With risk, the consumption profile insures that the shadow value of wealth is 0 at $t = T$. This difference in transversality condition, between the two problems, reflects the fact that in the absence of risk, utility from the bequest is obtained only at $t = T$. In the presence of risk, the utility of the bequest affects the expected

flow of utility (the function G) at every point in time where $f(t) > 0$, i.e. where the hazard is positive. This effect is apparent in the second term on the right side of equation 12.9, which is absent in the risk-free analog, equation 12.10. The term in square brackets in equation 12.9 is the ratio of the present value of the marginal utility of the bequest, to the present value of the marginal utility of consumption.

The second term in the Euler Equation 12.9 is positive, but because the risk of death changes the entire consumption profile, risk changes the level at which we evaluate the derivatives in the equation. Therefore, we cannot simply compare the right sides of equation 12.9 and 12.10 to determine the effect of risk. The two expressions are evaluated at different levels of c and k . The effect of uncertainty, on the optimal consumption profile in the general form of the problem is therefore not easy to determine.

In the special case where $B \equiv 0$, the second term in equation 12.9 vanishes, and the role of risk is transparent: an increase in risk (larger h) is equivalent to a time-varying increase in the pure rate of time preference, r . A higher discount rate causes the decision maker to value the future less because of greater impatience; a larger hazard rate causes the decision maker to value the future less because there is a greater chance that he will not be alive to enjoy it. Thus, increases in either r or h have the same qualitative effect: shifting consumption from the future to the present.⁸⁶ This observation is probably the most important insight from this branch of the literature, with exogenous risk. It is common to invoke the risk of a catastrophe that wipes out civilization, as an ethical basis for discounting the future, as in the Stern Review (cite here). If there is a chance that civilization will disappear, it seems reasonable to take this possibility into account when making social investments, such as those designed to protect against climate change.

12.2 Endogenous risk

The key feature in the problem studied in Chapter 12.1 is that risk there is exogenous: the decision maker cannot influence the hazard rate. There are many situations in which the decision maker influences the hazard rate for an

⁸⁶Problem: Show that in the consumption problem without risk, when the bequest function is 0, a higher level of the discount rate increases the current level of consumption at any time and at any level of wealth $k > 0$.

event. For example, consider the problem of machine maintenance, where the random event is the breakdown of a machine. One possibility is that the hazard rate at a point in time depends on how carefully the decision maker operates the machine at that instant. Another possibility is that the hazard rate depends on how well-maintained the machine has been throughout its history. In both of these cases, the hazard rate is endogenous. In the first case, however, the hazard rate depends on current actions, and in the second case it depends on the history of actions. The second case is more complex because it requires an additional state variable that captures the effect of past actions. Of course, the actual hazard may depend on both current actions and the history of actions. For example, the probability that a car breaks down during a particular trip (corresponding to an instant, in a continuous time setting) depends both on how it is driven during the trip (the next instant) and how it has been driven over the past years.

Here we study the problem where the risk of catastrophe depends on the stock of pollution. The key to simplifying this model is the assumption that the post-catastrophe payoff is 0. This assumption simplifies the problem for the same reason that setting the bequest $B(k) \equiv 0$ simplifies the uncertain-time-of-death model above. After we examine the case with post-catastrophe payoff equal to 0, we consider more complicated models where the payoff depends on the stock of pollution at the time of the catastrophe.

We denote the stock of pollution as x and the flow of emissions as y . The stock of pollution evolves according to

$$\dot{x} = y - g(x), \quad x(0) = x_0, \text{ given} \quad (12.11)$$

where the (not necessarily linear) $g(x)$ is the natural decay rate. We assume that $g(x)$ is continuous in x and that the optimal y is a continuous function of time, so that \dot{x} is a continuous function of time. The decision maker's flow of utility is $U(x, y)$, increasing and concave in emissions, y , and weakly decreasing and concave in pollution, x . A larger flow of pollution means that more resources are devoted to consumption which increases utility, and fewer resources are devoted to abatement. A larger stock of pollution decreases the flow of welfare, in addition to increasing the risk of a catastrophe. The pure rate of time preference is r .

The hazard rate at time t depends on the stock of pollution at that time. With some abuse of notation, we write the hazard rate as $h(t) = h(x(t))$,

with $h(x)$ a differentiable function. The time of catastrophe is a random variable. If the catastrophe occurs at time t , the value of the stream of welfare, discounted back to time 0, is

$$z(t) \equiv \int_0^t e^{-rs} U(x_s, y_s) ds. \quad (12.12)$$

Our use of $z(t)$ to denote this level of welfare emphasizes the similarity between the problems here and in the previous section.

Now consider an arbitrary pollution trajectory $x(t)$. This trajectory induces a trajectory for $F(t)$, the probability that the catastrophe has occurred by time t . By assumptions above, the hazard function is differentiable in x , which (in equilibrium) is a continuous function of time, so the density $f(t) = \frac{dF}{dt}$ exists. Therefore, we can write the planner's maximand as

$$E_t z(t) = \int_0^\infty z(t) \frac{dF}{dt} dt = \int_0^\infty z(t) f(t) dt. \quad (12.13)$$

By choosing the emissions trajectory, the planner chooses the pollution trajectory, and thereby chooses the function $f(t)$; here risk is endogenous. We assume that the planner is not able to bring the level of risk arbitrarily close to 0, or that it would be prohibitively expensive to do so. For example, eliminating risk might require the elimination of the pollution stock, which might be possible to achieve only asymptotically, and then only if consumption remains at 0 forever. If the marginal utility of consumption becomes large as consumption goes to 0, such a program is not optimal. Thus, $f(t) \geq \varepsilon$ for some $\varepsilon > 0$; these inequalities imply that $\lim_{t \rightarrow \infty} F(t) = 1$: with probability 1, the catastrophe eventually occurs. The catastrophe has not yet occurred at time 0, i.e. $F(0) = 0$.

Integrating the last expression in equation 12.13 by parts (as in the previous section), and using $\lim_{t \rightarrow \infty} F(t) = 1$, we write the expected payoff as

$$\int_0^\infty (1 - F(t)) e^{-rt} U(x_t, y_t) dt. \quad (12.14)$$

We want to choose an emissions path to maximize this expression. This expression is not an adequate basis for optimization, however, because it does not reflect the fact that $F(t)$ depends on the entire trajectory of $\{x_s\}_0^t$. To model this dependence, we define the survival probability, $S(t) = 1 - F(t)$,

the probability that the catastrophe has not occurred by time t . Taking the derivative with respect to time, we find that the rate of decrease of the survival probability equals the hazard rate:

$$-\frac{\frac{dS(t)}{dt}}{S(t)} = \frac{\frac{dF(t)}{dt}}{1 - F(t)} = h(x(t)).$$

Here we use the definition of the hazard rate and the fact that it is a function of the current pollution stock. In order to facilitate the interpretation of subsequent expressions, we also define $w = -\ln S$, which implies

$$S(t) = e^{-w(t)} = 1 - F(t) \quad (12.15)$$

and

$$\frac{dw(t)}{dt} = h(x(t)), \quad w(0) = 0. \quad (12.16)$$

The initial condition, $w(0) = 0$ states that the catastrophe has not occurred by the initial time, $t = 0$, i.e. the probability of surviving until time 0 is 1. This assumption is innocuous. If the catastrophe had occurred by time 0, there is no decision problem. The problem makes sense only if it begins before the catastrophe occurs.

With these definitions, we can write the maximand, expression 12.14, as

$$\int_0^{\infty} e^{-rt-w(t)} U(x_t, y_t) dt. \quad (12.17)$$

The inclusion of risk “resembles” an increase in the discount rate: the presence of risk changes the discount factor from e^{-rt} to $e^{-rt-w(t)}$, with $w(t) > 0$. We noted that in the case where risk is exogenous, as in Section 12.1, the presence of risk has the same effect as an increase in the discount rate. Here, however, risk is endogenous. Higher current consumption entails higher emissions, which increase the future trajectory of pollution, decreasing the future survival probability, thereby decreasing future expected utility.

The optimization problem consists of the maximand, expression 12.17, and the equations of motion for the pollution stock and the variable w , equations 12.11 and 12.16, together with their boundary conditions. In the case of exogenous risk, we were able to convert the problem under uncertainty into a deterministic control problem, at the cost of making the objective somewhat more complicated. Here we achieve a deterministic control problem,

but the cost is the inclusion of an additional state variable, w . There is no free lunch.

The current value Hamiltonian for this control problem is

$$H = e^{-w}U(x, y) + \mu_1(y - g(x)) + \mu_2h(x),$$

where μ_1 and μ_2 are the costate variables associated with the pollution stock, x , and the risk variable, w . The necessary condition for maximizing the Hamiltonian, and the costate equations are

$$e^{-w}U_y(x, y) + \mu_1 = 0 \quad (12.18)$$

$$\dot{\mu}_1 = \mu_1(r + g') - \mu_2h' - e^{-w}U_x \quad (12.19)$$

$$\dot{\mu}_2 = r\mu_2 + e^{-w}U. \quad (12.20)$$

Equation 12.18 has the usual interpretation: at the optimum, the marginal utility of an additional unit of emissions must equal the negative of the shadow value of the stock of pollution. This shadow value is negative.

We obtain the differential equation for emissions by following the usual procedure: we differentiate equation 12.18, then use equations 12.19 and 12.20 to eliminate $\dot{\mu}_1$ and $\dot{\mu}_2$, and then use equation 12.18 to eliminate μ_1 . The necessary conditions contain a single algebraic equation, because there is only one control variable. We can use this algebraic equation to eliminate only one of the costate variables. Thus, the differential equation for emissions still contains one costate variable. Following the steps listed above, this differential equation is

$$\dot{y} = a(y, x) + b(y, x, \rho, h(x)) \quad (12.21)$$

where

$$a(y, x) \equiv \frac{(r + g')U_y + U_x - (y - g)U_{xy}}{U_{yy}}$$

$$b(y, x, \rho, h(x)) = \frac{h(x)U_y + h'(x)\rho}{U_{yy}} \quad (12.22)$$

and

$$\rho(t) \equiv e^{w(t)}\mu_2(t). \quad (12.23)$$

We decomposed \dot{y} into two functions. The function a is independent of the hazard function, a fact that becomes useful when we consider the comparative

statics of the steady state with respect to the risk. Note also that we eliminate the costate variable μ_2 using the definition in equation 12.23. The reason for this definition will be apparent shortly.

At this level of generality, we can learn little about the optimal trajectory or the effect of risk outside the steady state. Therefore, we concentrate on steady state analysis. We assume that the trajectories in both the problems with and without risk (conditionally) converge to steady states. We want to know how risk affects the level of the steady state stock of pollution. Two points are important in considering this analysis.

First, recall that the trajectory of the control variable and the associated stock of pollution are conditioned on the event not yet having occurred. Therefore, the steady state that we examine is the level of pollution to which the stock would asymptotically converge provided that the catastrophe does not occur. This conditional convergence occurs only as $t \rightarrow \infty$. However, we adopted assumptions above that imply that $\lim_{t \rightarrow \infty} F(t) = 1$, i.e. the catastrophe occurs with certainty as t becomes infinitely large. (We used that assumption in evaluating a limit when integrating by parts to obtain the maximand in expression 12.14.) Given continuity, we know that for any $1 > \varepsilon > 0$, there exists a $\tau(\varepsilon)$ such that the probability that the disaster has occurred is greater than $1 - \varepsilon$ for $t > \tau(\varepsilon)$. If we look far enough into the future, we can be arbitrarily sure that the catastrophe has occurred. Nevertheless, we are interested in the effect of risk on a steady state, even though we know that the catastrophe is virtually certain to occur before the stock of pollution gets extremely close to that steady state. Hereafter, when we speak of the steady state, the reader should keep in mind that this is a conditional steady state; moreover, we know that the catastrophe will occur, and the problem end, before the actual pollution stock gets “very close” to the steady state, unless the initial condition happens to be very close to the steady state.

The second point is that we do not obtain the algebraic conditions that determine the steady state simply by setting the equations of motion for the state variables and the costate variables equal to 0, and using these together with the first order condition 12.18 to obtain five algebraic equations to solve for the five unknowns: the two state variables and corresponding costate variables and the control variable. Such a procedure would be nonsensical because the state variable w does not approach a steady state. Our assumptions above which imply that $h(x)$ is bounded away from 0, and equation

12.16, imply that $w(t)$ diverges to infinity.

We therefore have to be a bit careful in determining the algebraic equations that determine the steady state. We are interested in a steady state for the pollution stock, x . From equation 12.11 we see that $\dot{x} = 0$ requires that y is a constant in the steady state, i.e. $\dot{y} = 0$. Equation 12.21 implies that in order for y to be constant, it must be the case that ρ is constant. Taking derivatives of equation 12.23 and using equations 12.16 and 12.20 to eliminate the costate variable, we obtain an expression for $\dot{\rho}$. Setting this expression equal to 0 we obtain the steady state value of ρ , which we denote (with obvious abuse of notation) as

$$\rho = -\frac{U(x, y)}{r + h(x)} = -\int_0^{\infty} e^{-ht} e^{-rt} U dt. \quad (12.24)$$

Hereafter it is understood that ρ , x and y refer to steady state values unless we indicate otherwise by using a time argument. Equation 12.24 states that (the steady state) ρ equals the negative of the expected value of the program at the steady state. We include the second equality in equation 12.24 in order to emphasize that ρ equals the negative of the integral of the constant flow U , discounted at rate r and multiplied by the probability that the catastrophe occurs by time t . In the steady state, where the hazard is constant, the random time of catastrophe is exponentially distributed.

We now have the ingredients to obtain the steady state values. Clearly a steady state for x requires

$$y = g(x). \quad (12.25)$$

We set $\dot{y} = a(\cdot) + b(\cdot) = 0$, using equations 12.24 and 12.25 to write the steady state condition for y :

$$(r + g' + h)U_y + U_x = \frac{Uh'}{r + h}. \quad (12.26)$$

Equations 12.25 and 12.26 are two equations in two unknowns, the steady state value of the pollution stock and emissions.

Before we use these equations to determine the effect of risk on the steady state, we digress to explain the meaning of the variable ρ . In particular, we want to explain why it must be constant in a steady state, even though neither w nor μ_2 converge to constants. Define the expected present value

of the program at time t , given that the catastrophe has not yet occurred and given that the system begins in a steady state, as W :

$$W = E_\tau \int_t^\tau e^{-r(s-t)} U ds = \frac{U}{r+h}. \quad (12.27)$$

Due to the assumption that we begin in a steady state, W is independent of calendar time. Define J as the expectation at time 0 of the current value of the program at time t :

$$J = S(t)W.$$

That is, J equals the value of the program given that the system survives until time t (W) times the probability that it survives until time t , ($S(t)$). Now recall the interpretation of a shadow value. The shadow value of w , the costate variable μ_2 , equals the marginal change in the value of the program due to a change in the value of the state variable, here w . Using this interpretation of the costate variable and equation 12.15, we have

$$\mu_2(t) = \frac{\partial J}{\partial w} = \frac{\partial (S(t)W)}{\partial w} = \frac{\partial (e^{-w(t)}W)}{\partial w} = -e^{-w(t)}W$$

Rearranging, and using equation 12.23 we have

$$e^{w(t)}\mu_2(t) = \rho(t) = -W.$$

This equation implies that in the steady state $\rho(t)$ is constant, even though neither w nor μ_2 converge to constants. The derivation also provides another way of seeing that the steady state value of ρ equals the negative of the expectation of the value of the program, beginning with a pollution stock in the steady state.

We now return to consider the effect of risk on the steady state pollution stock. First note that in the absence of risk, the steady state condition 12.26 reduces to the familiar (from Chapter xx) requirement that

$$U_y = \frac{-U_x}{r+g'} = - \int_0^\infty e^{-(r+g')t} U_x dt. \quad (12.28)$$

In the steady state the marginal value of emissions is U_y . An extra unit of emissions creates an extra unit of stock which creates additional damages U_x . However, this additional unit decays at the constant rate g' , so the additional

damage t units of time in the future is $e^{-g't}U_x$. The present value of this future loss is $e^{-rt}(e^{-g't}U_x)$. The accumulation of all of these losses is the integral in equation 12.28. The steady condition requires that the marginal benefit of an additional unit of emissions equals the marginal cost of that additional unit.

In the absence of risk, the equations of motion for the stock of pollution and the flow of emission are equations 12.11 and 12.21 with $b \equiv 0$. By linearizing this system at the deterministic steady state, we obtain

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} \approx \begin{bmatrix} -g' & 1 \\ a_x & a_y \end{bmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = A \begin{pmatrix} x \\ y \end{pmatrix},$$

where the matrix A is defined implicitly. We noted in Chapter xx that if there is a steady state in a deterministic control problem, then it must be a saddle point. We have assumed that the trajectory converges to a steady state. Therefore, this steady is a saddle point; this conclusion implies (using the results from Chapter xx) that $|A| < 0$.

To determine the effect of risk on the steady state, we proceed in two steps. First, we determine how any non-zero value of the function $b(y, x, \rho, h(x))$ affects the level of the steady state; then we determine the effect of risk on the value of $b(\cdot)$. To achieve the first step, consider the pair of equations $y - g(x) = 0 = a(y, x) - b$. In order to determine the effect of b on the (or a) value of x that is part of the solution to this system, we totally differentiate the system and use Cramer's Rule to obtain

$$\frac{dx}{db} = \frac{1}{|A|} < 0, \quad (12.29)$$

where the inequality uses the fact that the steady state is a saddle point. Thus, we have accomplished the first step. We know that a positive value of b decreases the steady state stock of pollution, and a negative value increases the steady state stock. To accomplish the second step, we use equalities 12.1 and 12.24 to write

$$b > 0 \iff \frac{U_y}{U} < \frac{h'}{(r+h)h}. \quad (12.30)$$

Relation 12.30 tells us the qualitative effect of risk on the steady state. At one extreme, suppose that the risk is exogenous. In this case, h is independent of x , i.e. $h' \equiv 0$, so the second inequality in relation 12.30 is not satisfied;

here, $b < 0$, and inequality 12.29 implies that risk increases the steady state stock of pollution. This conclusion is consistent with the intuition developed in Chapter 12.1, where we saw that the inclusion of exogenous risk is qualitatively the same as increasing the discount rate. In Chapter xx we noted that a higher discount rate increases the steady state stock of pollution. At the other extreme, suppose that the risk is sensitive to the stock of pollution. In this case, h' is large; if it is sufficiently large that the second inequality in relation 12.30 is satisfied, then risk reduces the steady state stock. More generally, risk decreases the steady state stock of pollution if the ratio $\frac{h'}{h}$ is large, and increases the steady state stock if this ratio is small.

The analysis shows that the distinction between avoidable and unavoidable risk is crucial. If the risk is unavoidable, it increases the incentive to eat, drink and be merry (for tomorrow we die), thereby driving up the stock of pollution. However, if the risk is avoidable, we should be more ascetic, keeping the stock of pollution small in the hope that we will survive another day.

12.3 A jump process

The stochastic control of jump processes provides an alternative way of analyzing the problem with event uncertainty. Using a jump processes rather than converting the stochastic problem to a deterministic one (as above) is especially convenient if there is more than a single kind of event, or if the magnitude of the event might take different values. The jump process model is also useful if the state variable being controlled could have discontinuities.

For example, emissions leads to the accumulation of GHGs, and the higher GHG stock increases the probability of melting permafrost; this melting leads to rapid release of additional GHGs. If this rapid increase transpires on a faster time scale than the ordinary accumulation of GHG stocks, it may be reasonable to approximate it as a jump. The melting of the permafrost might not be a catastrophic event, and there may be various levels of melting, corresponding to which there are various levels of the jump in GHGs. For example, in the absence of melting the GHG stock does not jump, and different levels of melting lead to a jump of, for example, 1, 2, or 3 additional units of GHG stock. These events increase the state variable by a discrete amount, but need not lead to catastrophes. After such an event occurs,

the decision maker continues with the problem, only with a higher stock of GHGs.

In the interest of simplicity, we emphasize the case in which the jump can take a single value. We then apply the procedure to the problem of catastrophic change considered above.

12.3.1 The DPE for a jump process

Here we provide the necessary conditions for optimal control of a jump process. In this section, we define the state variable as v , rather than x as above, because subsequently we treat v as a vector. We want to reserve the symbol x to be the pollution stock, as above.

The equation of motion for the state variable v , with control variable y is

$$dv = f(v, y)dt + d\pi, \quad v(0) = v_0, \text{ given} \quad (12.31)$$

where

$$\Pr \{d\pi = c\} = C(v, y) dt \quad (12.32)$$

$$\Pr \{d\pi = 0\} = 1 - C(v, y) dt \quad (12.33)$$

We write the equation of motion in differential form, because if a jump at time t the change in v is discontinuous; therefore, the time derivative of v does not exist at time t . In this setting, $\pi(t)$ is a stochastic process; at each point in time the change in π , $d\pi$, takes one of two values: c with probability $C(v, y) dt$ and 0 with probability $1 - C(v, y) dt$. We could generalize to a multivariate distribution, by allowing the change in the random variable to take a third value, say g , with probability $G(v, y)dt$. In that case, the probability of no jump equals $1 - [C(v, y) + G(v, y)] dt$. This model implies that the probability of an event over a small interval of time, dt , is proportional to the length of that interval, and may depend on the value of the state variable and control variable at the beginning of the interval.

Suppose that the flow of welfare is $L(v, y, t)$ and the decision maker wants to maximize the expectation of the integral of this flow over $[0, T]$. We obtain the continuous time dynamic programming equation using almost the same approach taken in Chapter xx. The only difference is that here we have an expectations operator because the equation of motion is random. Unlike the

continuous time stochastic control problem in Chapter yy, we do not have to use the Ito calculus. As a consequence, we use a first order Taylor expansion (as in the deterministic setting) rather than a second order Taylor expansion (as in the continuous time stochastic control setting).

First, we define the value function as the expectation of the maximized value of the program, given an arbitrary starting time t and value of the state variable v . Second, we break the integral into two parts, “the next short interval” and “everything after the next short interval”. Third, we recognize that the expectation of the maximized value of “everything after the next short interval” equals the value function at the later date and the new value of the state variable. The following three equations show these three steps:

$$\begin{aligned}
J(v, t) &= \max_{\{y\}_t^T} E_{\{\pi\}_t^T} \int_t^T L(v, y, s) ds \\
&= \max_{\{y\}_t^{t+dt}} E_{\{\pi\}_t^{t+dt}} \left[\int_t^{t+dt} L(v, y, s) ds + \max_{\{y\}_{t+dt}^T} E_{\{\pi\}_{t+dt}^T} \int_{t+dt}^T L(v, y, s) ds \right] \\
&= \max_{\{y\}_t^{t+dt}} E_{\{\pi\}_t^{t+dt}} \left[\int_t^{t+dt} L(v, y, s) ds + J(v + dv, t + dt) \right]. \quad (12.34)
\end{aligned}$$

The first order Taylor expansion of the expectation of the integral is

$$E_{\{\pi\}_t^{t+dt}} \left[\int_t^{t+dt} L(v, y, s) ds \right] = L(v, y, t) dt + o(dt). \quad (12.35)$$

The first order Taylor expansion depends on the current values of v, y, t , which are known; thus the expectations operator is vacuous here.

We use equations 12.31 - 12.33 to write

$$\begin{aligned}
&E_{\{\pi\}_t^{t+dt}} [J(v + dv, t + dt)] \\
&= (C(v, y) dt) \times J(v + f(v, y)dt + c, t + dt) \\
&+ (1 - C(v, y) dt) \times J(v + f(v, y)dt, t + dt).
\end{aligned}$$

This equation uses the fact that the value of dv depends on the stochastic value of $d\pi$. The first order Taylor expansion of the last expression equals

$$\begin{aligned}
&E_{\{\pi\}_t^{t+dt}} [J(v + dv, t + dt)] \\
&= J(v, t) + [J(v + c, t) - J(v, t)] \times (C(v, y) dt) \\
&+ J_v(v, t) f(v, y)dt + J_t(v, t) + o(dt). \quad (12.36)
\end{aligned}$$

Using equations 12.35 and 12.36 in the last line of equation 12.34, we obtain

$$J(v, t) = \max_y [L(v, y, t)dt + J(v, t) + [J(v + c, t) - J(v, t)] \times (C(v, y) dt) \\ + J_v(v, t) f(v, y)dt + J_t(v, t) + o(dt)].$$

We cancel $J(v, t)$ from both sides, divide by dt and take the limit as $dt \rightarrow 0$ to obtain the DPE for the control problem with a jump process:

$$-J_t(v, t) = \max_y [L(v, y, t) + J_v(v, t) f(v, y) \\ + [J(v + c, t) - J(v, t)] \times (C(v, y))]. \quad (12.37)$$

If the problem is autonomous, i.e. if $T = \infty$ and $L(v, y, t) = e^{-rt}U(v, y)$, then the value function is multiplicatively separable, with the form $J(v, t) = e^{-rt}V(v)$. Substituting this form into the DPE, we have the DPE for the current value function $V(v)$:

$$rV(v) = \max_y [U(v, y) + V_v(v) f(v, y) + [V(v + c) - V(v)] \times (C(v, y))]. \quad (12.38)$$

We obtain the DPE for the deterministic control problem if either $c = 0$ or $C(v, y) \equiv 0$. The inclusion of risk adds the term $[J(v + c, t) - J(v, t)] \times (C(v, y))$ in the case of the non-autonomous problem and the term $[V(v + c) - V(v)] \times (C(v, y))$ in the case of the autonomous problem. Both of these terms equal the loss in the value of the program due to the occurrence of the jump, times the probability of this occurrence.

12.3.2 The jump process and catastrophic risk

In order to apply the DPE above we need to treat the state v as a vector. The first element of v is the pollution stock, x , with the equation of motion 12.11. The second element of v is the stochastic process, π . We interpret π as an indicator function: $\pi = 0$ implies that the disaster has not yet occurred, and $\pi = 1$ implies that the disaster has occurred. Once the disaster occurs, it cannot be reversed; subsequent utility is 0, and the problem is over. Using the definition of the hazard rate, we have the equation of motion for π .

$$\Pr \{d\pi = 1 \mid \pi = 0\} = h(x) dt \\ \Pr \{d\pi = 0 \mid \pi = 0\} = 1 - h(x) dt \\ \Pr \{d\pi = 1 \mid \pi = 1\} = 1.$$

The “current value” value function is $V(x, \pi)$. We have $V(x, 1) = 0$, reflecting the fact that once the catastrophe occurs, the continuation value is 0; $V(x, 0)$ equals the maximized value of the expectation of the discounted stream of pre-catastrophe utility, $U(x, y)$. Using these definitions in the current value DPE 12.38 yields the DPE for this problem, given the initial condition $\pi = 0$:

$$\begin{aligned} rV(x, 0) &= \max_y [U(v, y) + V_x(x, 0) f(v, y) + [V(x, 1) - V(x, 0)] \times h(x)] \\ &= \max_y [U(v, y) + V_x(x, 0) (y - g(x))] - V(x, 0) h(x) \implies \\ &(r + h(x)) V(x, 0) \\ &= \max_y [U(v, y) + V_x(x, 0) (y - g(x))]. \end{aligned}$$

The formulation of the DPE in the last two lines shows the manner in which the risk of catastrophe plays a role analogous to an increase in the discount rate. However, the risk, unlike the discount rate, is endogenous.

The first order condition for an interior maximization of the right side of the DPE is

$$U_y(v, y) + V_x(x, 0) = 0. \quad (12.39)$$

Applying the envelope theorem to the DPE (i.e. differentiating both sides of the DPE with respect to x and evaluating at the maximum) gives

$$h'V(x, 0) + (r + h) V_x(x, 0) = U_x(v, y) - V_x(x, 0) g' + V_{xx}(x, 0) (y - g(x)).$$

Evaluating this expression at a steady state, where $y - g(x) = 0$, gives

$$h'V(x, 0) + (r + h) V_x(x, 0) = U_x(v, y) - V_x(x, 0) g'. \quad (12.40)$$

We use the first order condition 12.39 and evaluate the DPE in the steady state (where $(r + h(x)) V(x, 0) = U(v, y)$) to eliminate V and V_x from equation 12.40 to obtain the steady state condition

$$h' \frac{U}{r + h} - (r + h + g') U_y = U_x. \quad (12.41)$$

Equations 12.26 and 12.41 are equivalent, as of course they must be, since they are the result of analyzing the same problem, using different methods. The analysis using jump processes is (arguably) simpler, because it does not require using the definition of ρ in equation 12.23.

12.4 Recent applications

This section [under construction] discusses recent applications of models of event uncertainty to problems in environmental and resource economics.

13 Stochastic Control

This section introduces continuous time stochastic methods. Continuous time methods can help us to find closed form solutions and to render tedious discrete time calculations more elegant. They also expand the set of numeric solution techniques. After reviewing the basics of continuous time stochastic processes we introduce the standard Brownian motion (Wiener process). It serves as the point of departure for a large set of continuous diffusion processes that economists employ to model stochastic prices or resource stocks. These diffusion processes defy the usual definition of integration and differentiation. We introduce the stochastic integral heuristically and explain the resulting rules of Ito's differential calculus. We discuss basic stochastic differential equations and their relation to partial differential equations. Then, we extend dynamic programming to the stochastic continuous time setting and explore a resource extraction problem. The first section of this chapter reviews the basics of probability theory required to introduce stochastic processes.

13.1 Probability Theory and Stochastic Processes - A Brief Review

We invite the reader familiar with probability spaces and measure theory to skip to the next section. Here, we review some basic concepts of probability theory that prove useful or even essential for continuous time stochastics. In probability theory, a probability space defines what we need to talk meaningfully about a stochastic problem at hand. First, the *sample space* Ω defines the set of possible outcomes $\omega \in \Omega$. An outcome ω can be the realization of a coin toss, or a sequence of coin tosses, or the realization of a resource price or its entire path. For example, let our space of interest be the outcome of rolling a die. The sample space Ω has the six elements $\omega_1 = 1, \omega_2 = 2, \dots, \omega_6 = 6$. Generically, we write that $\omega = \omega_i$ for a particular realization. In our example, we would write more naturally $\omega = 5$ if the die shows the number 5.

Second, a probability space requires a σ -algebra \mathcal{F} (*sigma-algebra*) on Ω . It defines the sets of outcomes (events) that the observer can distinguish. For example, an event can be the set $\{\omega_1, \omega_2\}$, which we interpreted as either ω_1 or ω_2 having occurred. If we were to bet on the die showing either of

1 or 2 this information is all we require to know whether we won. Note that whenever we know whether $\{\omega_1, \omega_2\}$ occurred, we also know whether its complement $\{\omega_3, \omega_4, \omega_5, \omega_6\}$ occurred: if the die shows either 1 or 2 it cannot show 3, 4, 5, or 6; if it does not show 1 or 2, then it must show one of 3, 4, 5, or 6. A σ -algebra summarizes a consistent information set on Ω by requiring that if some event is part of \mathcal{F} then also the complement of the event has to be part of \mathcal{F} . More precisely, a σ -algebra on Ω satisfies the following axioms

1. It contains the sample space Ω : $\Omega \in \mathcal{F}$.
2. With any element it also contains its complement: $E \in \mathcal{F} \Rightarrow E^c \in \mathcal{F}$.
3. It is closed under infinite unions: $E_i \in \mathcal{F} \forall i \in \mathbb{N} \Rightarrow \cup_{i=1}^{\infty} E_i \in \mathcal{F}$.

Why do we include unions of different events? If we know whether the die shows one of 1 or 2 ($\{\omega_1, \omega_2\}$) and we know whether the die shows one of 3 or 4 ($\{\omega_3, \omega_4\}$) then we also know whether the die shows one of 1 through 4 ($\{\omega_1, \omega_2, \omega_3, \omega_4\}$). If we merely bet on our die showing either of 1 or 2, we are content with the σ -algebra $\mathcal{F} = \{\emptyset, \{\omega_1, \omega_2\}, \{\omega_3, \omega_4, \omega_5, \omega_6\}, \Omega\}$. If we want to distinguish all possible outcomes of rolling the die we need a richer σ -algebra, the power set 2^Ω . The power set of Ω is the set of all subsets. For finite outcome spaces, its cardinality is two to the power of the cardinality of the sample space Ω , which explains the common notation “ 2^Ω ”.

There are two reasons for not simply choosing the power set as the underlying σ -algebra \mathcal{F} . First, we often want to model partial information by choosing a coarser set than the power set. In particular, if a realization ω characterizes a time path of a resource price, the σ -algebra capturing our present information should distinguish only those realizations that differ in the past, but not those realizations that only differ in the future. This idea will lead us to the definition of a filtration (a time indexed sequence of σ -algebras) further below. Second, if the sample space Ω is not finite, the power set is so large that it contains (somewhat pathological) events whose probabilities we cannot reasonably define (measure). In this case the “analogue” of the power set for finite sample spaces is the Borel set (or Borel σ -algebra), which contains all the events (subsets) whose probabilities we possibly want to measure, but excludes the pathological sets.

Third, a probability space contains a *probability measure*. In general, a map $\mathbb{P} : \mathcal{F} \rightarrow \mathbb{R}$ is a probability measure if

1. The probability of the full set is unity: $\mathbb{P}(\Omega) = 1$.
2. Any event has a non-negative probability: $\mathbb{P}(E) \geq 0 \forall E \in \mathcal{F}$.
3. \mathbb{P} is σ -additive: $\mathbb{P}(\sum_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} \mathbb{P}(E_i)$, where $\sum_{i=1}^{\infty} E_i$ denotes the union of pairwise disjoint events $E_i \in \mathcal{F}$.

The conceptual difference between a probability measure and a probability distribution is that the measure defines the probability directly for all events, e.g., including the event $\{\omega_1, \omega_2\}$. Note that probability measures deal with discrete and continuous distributions jointly.⁸⁷ The triple (Ω, \mathcal{F}, P) (sample space, σ -algebra, probability measure) defines a *probability space*.

A *random variable* maps a stochastic event into a some measurable outcome, e.g., your payoff. Formally, it is a map $X : (\Omega, \mathcal{F}) \rightarrow (\hat{\Omega}, \hat{\mathcal{F}})$, where Ω is our sample space and \mathcal{F} the σ algebra on Ω . The set $\hat{\Omega}$ describes the possible outcomes of the random variable (e.g. payoffs) and is usually a subset of the real numbers. The set $\hat{\mathcal{F}}$ is a σ -algebra on $\hat{\Omega}$ and characterizes the corresponding payoff events to which we want to assign probabilities. Continuing our example from above, assume that we are betting \$100 on the event $\{\omega_1, \omega_2\}$ that a die shows either of the numbers 1 or 2. If we loose, we owe \$50 to the bank. Then, the random variable X maps the event $\{\omega_1, \omega_2\}$ into the value 100 and the event $\{\omega_3, \omega_4, \omega_5, \omega_6\}$ into the value -50. A reasonable way to define the

⁸⁷There is a close relation between probability measures and cumulative distribution functions (cdfs) for probabilities that are defined on a number space. Given a probability measure on a sufficiently rich σ -algebra, we obtain the cumulative distribution function by defining $F(x) \equiv \mathbb{P}(\omega \leq x)$. Perhaps more interestingly, we also have a converse. Any function $F : \mathbb{R} \rightarrow [0, 1]$ that satisfies

1. F is monotonically increasing,
2. F is right-continuous,
3. $\lim_{x \rightarrow -\infty} F(x) = 0$ and $\lim_{x \rightarrow \infty} F(x) = 1$,

defines a probability measure on the Borel set (think of it as the set of all possibly interesting events) that is uniquely pinned down by requiring $\mathbb{P}((a, b]) \equiv F(b) - F(a)$ for all $a \leq b$ (which is a requirement on only a small subset of the sets in the Borel σ -algebra). We invite the reader to test that this definition also works for our example of a die, where the cdf has discontinuous jumps (of height $\frac{1}{6}$) to the left of 1, 2, 3, 4, 5, 6 (it is right continuous). In the discrete case it is somewhat trivial that defining the probabilities for all individual outcomes contains the necessary information to define the probability of all possible events. In the continuous case, one has to proof that the probability measure extends uniquely from some subset to the σ -algebra generated by these sets.

sample space of our numeric values is $\hat{\Omega} = \{-50\}, \{100\}$ and a reasonable way to define the σ -algebra $\hat{\mathcal{F}}$ is as the set $\{\emptyset, \{-50\}, \{100\}, \{-50, 100\}\}$.

We say that our random variable X is $(\mathcal{F}, \hat{\mathcal{F}})$ -measurable, or simply *measurable*, if we can meaningfully assign probabilities to each element of $\hat{\mathcal{F}}$ given a probability measure on \mathcal{F} . In our example, we want to be able to assign probabilities to the payoff events \$-50 and \$100 given a the probability space (Ω, \mathcal{F}, P) that characterizes our roll of the die. If our σ -algebra of the underlying probability space is $\mathcal{F} = \{\emptyset, \{\omega_1, \omega_2\}, \{\omega_3, \omega_4, \omega_5, \omega_6\}, \Omega\}$ we are good. Say \mathbb{P} assigns the probability $\frac{1}{3}$ to the event $\{\omega_1, \omega_2\}$ and the probability $\frac{2}{3}$ to its complement we can assign the probability $\frac{1}{3}$ to the payoff \$100 of our random variable and the probability $\frac{2}{3}$ to our payoff \$-50 (loss). If instead our σ -algebra of the underlying probability space is the power set $\mathcal{F} = 2^\Omega$ then we are just as good. The power set is much richer than the earlier σ -algebra and we obtain the same probabilities on the payoff space.

Now assume that the σ -algebra of the underlying probability space is $\mathcal{F} = \{\emptyset, \{\omega_1, \omega_3, \omega_5\}, \{\omega_2, \omega_4, \omega_6\}, \Omega\}$. This σ -algebra only contains the information whether the die shows an odd or an even number. In this case, our underlying probability space does not allow us to assign probabilities to the payoffs of our random variable, which maps the outcome of seeing either of 1 or 2 into the payoff \$100 and the complement to \$-50. We say that our random variable is not measurable. Based on the information available in our probability space we cannot say whether we won or whether we lost and we cannot measure our payoff. Formally, the random variable X is measurable if $X^{-1}(\hat{\mathcal{F}}) \equiv \{X^{-1}(E) | E \in \hat{\mathcal{F}}\} \subset \mathcal{F}$, i.e., if the inverse image of any event in the payoff space is measurable in the original probability space. If the random variable X is measurable, we can define the implied probability measure \mathbb{P}^X on the payoff space for all events $E \in \hat{\mathcal{F}}$ by defining $\mathbb{P}^X(E) = \mathbb{P}(X^{-1}(E))$.

Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, we define a continuous time *stochastic process* as a measurable function $X : \Omega \times [0, \infty) \rightarrow \mathbb{R}$. We write $X_t(\omega)$ to denote the realized value at time t for state of the world ω . Note that a realization of ω pins down the full path, not just one period's realization. We would not want to define the probabilistic realizations period by period because we have to be able to characterize the joint distribution of the realizations at different points in time, e.g., to introduce serial correlation or even to define independence. Constructing the underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$ for general stochastic processes is quite technical and achieved by Kolmogorov's extension theorem. The theorem provides the existence of the

underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$ if we specify the joint distribution of our stochastic processes in consistent manner. As a consequence, we usually don't even spell out the probability space, but work with the joint distributions instead. The next section will rely on Kolmogorov's extension theorem when we introduce the Brownian motion by characterizing its joint distribution.

A *filtration* captures an observer's information at a given point in time. Formally, it is an increasing sequence of σ -algebras $\{\mathcal{F}_t\}_{t \geq 0}$ with $\mathcal{F}_t \subset \mathcal{F}$ for all t . By an increasing sequence of sets we mean that $t_1 < t_2 \Leftrightarrow \mathcal{F}_{t_1} \subset \mathcal{F}_{t_2}$. By definition, the filtration \mathcal{F}_{t_2} at the later time t_2 is finer than the filtration \mathcal{F}_{t_1} at time t_1 and the filtration at time t_2 can distinguish more events (events that realized between t_1 and t_2). A stochastic process $X : \Omega \times [0, \infty) \rightarrow \mathbb{R}$ is adapted to the filtration $\{\mathcal{F}_t\}_{t \geq 0}$ if its realization in period t given by $X_t : \Omega \rightarrow \mathbb{R}$ is $(\mathcal{F}_t, \hat{\mathcal{F}})$ -measurable for all t , where $\hat{\mathcal{F}}$ is the σ -algebra on our payoff space, usually the Borel set.

Martingales are a particularly convenient class of stochastic processes. Let X be a stochastic process that is adapted to the filtration $\{\mathcal{F}_t\}_{t \geq 0}$. It is a martingale if $E(X_{t_2} | \mathcal{F}_{t_1}) = X_{t_1}$. Here, $E(X_{t_2} | \mathcal{F}_{t_1})$ denotes the expectation of X_{t_2} conditional on the information in period $t_1 < t_2$. Thus, a martingale is a stochastic process whose best forecast for a future period is its present value. Note that a conditional expectation of a random variable is itself a random variable: the expectation depends on the information at time t_1 , which (usually) only realizes in period t_1 . Another way to observe this stochastic nature of the conditional expectation is by realizing that the conditional expectation above equals X_{t_1} , which is a random variable. A convenient property of conditional expectations is the *law of iterated expectations*, or tower law. It holds for all random variables Z , not just for martingales, and states: if $E(|Z|) < \infty$ and $\mathcal{F}_{t_1} \subset \mathcal{F}_{t_2}$ then $E[E(Z | \mathcal{F}_{t_2}) | \mathcal{F}_{t_1}] = E(Z | \mathcal{F}_{t_1})$. A forecast of a random variable Z cannot be improved by conditioning it on the finer information set \mathcal{F}_{t_2} if that information set itself has to be forecasted based on the coarser information set \mathcal{F}_{t_1} .

13.2 Brownian Motion

The Brownian Motion is a stochastic process that underlies virtually all diffusion processes, i.e., stochastic processes without jumps. The standard Brow-

nian motion is known as the Wiener process. It formalizes the concept of a random walk in continuous time. A symmetric random walk in discrete time denotes a model in which a particle (or a price) move up or down by an equal amount with equal probability in every period. The best guess for the particles future location in the random walk model is it's present location (martingale property). The variance of its position increases linearly in time.⁸⁸ The Brownian motion captures these characteristics in continuous time.

For a formal definition, we denote the underlying probability space by (Ω, \mathcal{F}, P) and define the stochastic process X or X_t as a measurable function $X : \Omega \times [0, \infty) \rightarrow \mathbb{R}$. We write $X_t(\omega)$ to denote the realized value at time t for state of the world ω . A Wiener process or *standard Brownian motion* is a stochastic process B defined by the properties⁸⁹

1. $B_0 = 0$;
2. For any time t and $s > t$ the difference $B_s - B_t$ is normally distributed with mean zero and variance $t - s$, i.e., $B_s - B_t = \epsilon_{t,s}\sqrt{s - t}$ with $\epsilon_{t,s} \sim N(0, 1)$;
3. The increments for non-overlapping time intervals are independent, i.e.

⁸⁸We can model the particle's location by a binomial distribution identifying the steps up with the number of "successes". The variance of a binomial distribution is proportional to the number of jumps (here: periods). In detail, assume that a particle starts moving from location zero and moves up or down one unit in every period $t \in \mathbb{N}$. We focus on its location after an even number of periods t , implying an even location k . If it jumped up half of the times then the particle is back at zero. For given t , every jump up in addition to $\frac{t}{2}$ moves the particle 2 positions up because it jumped one more time up instead of down. In general, the particle will be at location l if we observed $\frac{l+t}{2}$ jumps up (successes). So the particle is at l with probability $\binom{t}{\frac{l+t}{2}} \frac{1}{2^t}$. The variance of the number of successes is $\frac{t}{4}$, which translates in a variance of the particle's location of t because every success moves the particle by two units (one move up instead of down, $l = 2k - t$, where k is the number of successes). Note that the binomial distribution B converges to a normal distribution, which describes the increments of the continuous time Brownian motion that we are about to introduce.

⁸⁹Note that continuity of the sample paths is not always required in the definition of a standard Brownian motion. However, one can show that for any Brownian motion (defined without the continuity assumption) there exists a continuous modification, i.e., a stochastic process that coincides with the original Brownian motion almost surely and whose realized paths are continuous (as we require directly in the above definition).

for times t_0, t_1, \dots, t_n such that $0 \leq t_0 < t_1 < \dots < t_n < \infty$ it holds $E \epsilon_{t_i, t_{i+1}} \epsilon_{t_j, t_{j+1}} = 0$ for $j > i$.

4. Each realized path is continuous, i.e., for each $\omega \in \Omega$ the path $B(\omega, \cdot) \in \mathbb{R}^{[0, \infty)}$ is continuous.

In summary, a Brownian motion is a continuous stochastic process with independent normally distributed increments whose variance increases linearly in time.

Zooming in on a Brownian displays a peculiar behavior. For any finite time difference $\Delta = t - s$ we can rearrange the expression for the change of the Brownian motion $\Delta B = B_t - B_s$ in point 2 of the definition above into

$$\frac{\Delta B}{\Delta t} = \frac{\epsilon}{\sqrt{\Delta t}}.$$

This (stochastic) difference quotient goes to ∞ when we perform the limit $\Delta t \rightarrow 0$, implying that the “slope” of the Brownian approaches infinity as we zoom in far enough. Appendix 13.9.1 formalizes this reasoning. We show that the absolute variation of a Brownian motion is infinite on any finite time interval. The absolute variation adds the ups and downs of the Brownian motion in absolutes. If you were to walk along the path of a Brownian motion, you had to walk an infinite infinite distance in every ever so small fraction of a second. Most of these ups and downs cancel each other so that the net movement is finite. However, the extreme up and down movement of a Brownian motion on small times scales make it (almost surely) non-differentiable.⁹⁰

Let $\bar{\mathcal{F}}$ denote the filtration generated by the Brownian motion B . The σ -algebra $\bar{\mathcal{F}}_t$ captures the information about the past realizations of the Brownian motion B at time t . For technical reasons, we also have to include all sets of measure zero in the filtration, defining the filtration \mathcal{F} as the sequence of sigma algebras from $\bar{\mathcal{F}}$ and all measure zero sets. For practical purpose, the filtration \mathcal{F}_t captures the information made available by the Brownian

⁹⁰A Brownian motion is Hölder continuous of degree $\alpha = .5$. A function f is Hölder continuous of degree $\alpha \in [0, 1]$ if there exists a finite constant C such that $\text{frac}|f(y) - f(x)||y - x|^\alpha < C$ for all x and y in the domain of f . Note that $\alpha = 0$ merely implies boundedness, $\alpha > 0$ implies continuity, and $\alpha = 1$ implies continuous differentiability.

motion at time t . The standard Brownian motion is a *martingale*: at any time the best guess for its future value is its current value: $E[B_s | \mathcal{F}_t] = B_t$ for all $t \leq s$ (see point 2 of the definition).

13.3 Stochastic Integration

Assume that the price of a resource stock or an asset follows a Brownian motion B_t . If we hold a constant amount of the stock $\bar{f} \in \mathbb{R}$, we can define our gains (or losses) due to the price change over a time interval $[t_0, t_1]$ by

$$\bar{f} (B_{t_1} - B_{t_0}) . \quad (13.1)$$

More generally, the level of the resource stock we hold, or the composition of an optimal asset portfolio, changes continuously over time. In order to capture such changes we have to replace expression (13.1) by a stochastic integral. The stochastic integral extends the classical notion of an integral or a measure integral by defining integration with respect to a Brownian motion.

The usual procedure of defining an integral is to define it via summation for a set of simple functions. Then, approximate a more general set of functions by help of the simple functions, and then to extend the definition by limit to the set of general functions. The initial step in defining the stochastic integral is similar to the familiar procedure. We define the stochastic integral for step functions, i.e., functions $f : [a, b] \rightarrow \mathbb{R}$ of the form

$$f(t) = \sum_{i=0}^{n-1} \bar{f}_i \mathbf{1}_{[t_i, t_{i+1})}(t) \quad \forall t \in [a, b], \quad (13.2)$$

with $a = t_0 < t_1 < \dots < t_n = b$, $\bar{f}_i \in \mathbb{R} \forall i$, and $\mathbf{1}_{[t_i, t_{i+1})}$ denoting the unit function that takes the value one on $[t_i, t_{i+1})$ and zero otherwise. We define the integral of f with respect to the Brownian motion B as

$$\int_a^b f(t) dB_t \equiv \sum_{i=0}^{n-1} \bar{f}_i (B_{t_{i+1}} - B_{t_i}). \quad (13.3)$$

The second step approximates a set of general functions by step functions. However, we cannot simply extend the integral defined in (13.1) to a continuously changing function by taking a point-wise limit of the integral. The

previous section, and in more detail Appendix 13.9.1, discusses that the absolute variation of the Brownian motion is infinite and that a classical time limit would imply an infinite slope of the Brownian motion at any given point in time. In contrast to the absolute variation, the quadratic variation of a Brownian motion is finite. That is, the squared version of the up and down movements can be summed resulting in a finite value. The definition of the stochastic integral uses that the quadratic variation is better behaved than the absolute variation. The crucial step in defining the stochastic integral is the Ito-Isometry.

Proposition 8 (Ito-isometry): The integral defined in equation (13.3) satisfies

$$\mathbb{E} \left(\int_a^b f_t dB_t \right)^2 = \int_a^b \mathbb{E} f_t^2 dt .$$

More generally, this results holds for general random processes f_t that are adapted to the filtration \mathcal{F}_t and square integrable: $f_t \in L^2 = \{\int f_t^2 dt < \infty \text{ almost surely}\}$.

The proposition relates the variance of the stochastic integral to a standard integral with respect to time. The second part of the statement gets ahead of the our present definition of the stochastic integral. We proof the proposition for integrals over step functions. By definition (13.3) we have

$$\begin{aligned} \mathbb{E} \left(\int_a^b f(t) dB_t \right)^2 &= \mathbb{E} \left(\sum_{i=0}^{n-1} \bar{f}_i (B_{t_{i+1}} - B_{t_i}) \right)^2 \\ &= \mathbb{E} \sum_i \bar{f}_i^2 (B_{t_{i+1}} - B_{t_i})^2 + \mathbb{E} \sum_{i \neq j} \bar{f}_i \bar{f}_j (B_{t_{i+1}} - B_{t_i})(B_{t_{j+1}} - B_{t_j}) \\ &= \sum_i \bar{f}_i^2 \mathbb{E} (B_{t_{i+1}} - B_{t_i})^2 = \sum_i \bar{f}_i^2 (t_{i+1} - t_i) = \int_a^b f(t)^2 dt . \end{aligned} \quad (13.4)$$

The double sum over i and j in the second line drops out because the increments of a Brownian motion on disjunct time intervals are independent (property 3 in the definition). The expected value of the square of the increment of the Brownian motion on the finite time interval $t_{i+1} - t_i$ returns the time interval by property 2 of the definition of a Brownian motion.

Intuitively, the general definition of the stochastic integral makes use of line (13.4) in the proof above and the fact that the quadratic variation of the Brownian motion is well-defined in the continuous time limit. Formally, the general definition recognizes the left hand side of the Ito isometry in Proposition 8 as the norm of the integral $\int_a^b f(t)^2 dB_t$ on our underlying probability space.⁹¹ Similarly, for a deterministic function f , the right hand side of the Ito isometry is the l^2 -norm of the function f (the norm defined by square integrating a function). Thus, the Ito isometry defines a (one to one) map between the integrals $\int_a^b f(t) dB_t$ in our probability space equipped with the l^2 norm and the space of all square integrable functions, denoted by L^2 . Such a distance preserving one to one map between two metric spaces is called an isometry (hence the name Ito isometry). In particular, if a sequence converges in one of the isometric spaces, then its image converges in the other. Now we can use the fact that the subspace of all step functions lies dense in L^2 to extend our definition of the stochastic integral to all square integrable functions. To define the integral of an arbitrary $f \in L^2$, we approximate it by a sequence of step functions. Then, the Ito isometry implies that the integral defined in equation (13.3) converges to a well defined integral of f . So far, we assumed that f was deterministic. A stochastic integral with respect to deterministic functions is called the Wiener integral and predates the Ito stochastic integral.

In general, we want to integrate adapted stochastic processes with respect to a Brownian motion. For example, an optimal decision rule implies that we sell or buy a stock as a result of the (past) realizations of a Brownian motion determining its market price. Then, not only the price moves stochastically, but also the quantity of the stock that we are holding. Because it responds to the realizations of the price movement, the control of the stock holding is stochastic process that is adapted to the (filtration of) the Brownian motion that characterizes the price movement. The extension of the stochastic integral to general adapted stochastic processes is analogous to the definition of the Wiener integral that we introduced above. It replaces the deterministic step functions by stochastic step functions, where the coefficients \tilde{f}_i in Definition (13.2) are random variables that are measurable with respect to \mathcal{F}_t . Then, the Ito isometry is mapping the general integral to square integrable

⁹¹The norm is defined as $\|I\| = \int_{\Omega} I(\omega)^2 d\mathbb{P}$ and the case of the stochastic integral we have $E \left(\int_a^b f(t) dB_t \right)^2 = \int_{\Omega} \left(\int_a^b f(t) dB_t \right)^2 d\mathbb{P} = \left\| \int_a^b f(t) dB_t \right\|^2$.

adapted processes that are defined under a combination of the Lebesgue measure and the underlying probability measure \mathbb{P} .⁹² As a consequence, Proposition 8 takes the expected value of the adapted stochastic process on the right hand side of the Ito isometry.

The resulting integral

$$\int_0^T f_t dB_t$$

is itself an adapted stochastic process with continuous sample paths. Moreover, the integral is linear, i.e., for $a, b \in \mathbb{R}$ we have

$$\int_0^T af_t + bg_t dB_t = a \int_0^T f_t dB_t + b \int_0^T g_t dB_t .$$

If f_t is bounded,⁹³ then $\int_0^T f_t dB_t$ is itself a martingale. Given that $B_0 = 0$, we then have

$$\mathbb{E} \int_0^T f_t dB_t = 0$$

and, by the Ito-isometry,

$$\text{Var} \left(\int_0^T f_t dB_t \right) = \int_0^T \mathbb{E} f_t^2 dt .$$

⁹²In more detail, the adapted process $f(t)$ is a function of $\omega \in \Omega$ and $\int_a^b \mathbb{E} f(t, \omega)^2 dt = \mathbb{E} \int_a^b (f(t, \omega))^2 dt = \int_{\Omega} \left(\int_a^b f(t, \omega)^2 dt \right) d\mathbb{P} = \|f\|_{\lambda \otimes \mathbb{P}}$ where λ is the Lebesgue measure accounting for the standard integration $\int_a^b f(t, \omega)^2 dt$ with respect to time. The norm $\|\cdot\|_{\lambda \otimes \mathbb{P}}$ is the l^2 norm on the measure space $([a, b] \times \Omega, \mathcal{B}[a, b] \otimes \mathcal{F}, \lambda \otimes \mathbb{P})$, and $\mathcal{B}[a, b]$ denotes the Borel σ -algebra on the interval $[a, b]$. Here, the Ito isometry maps the stochastic integral from the underlying probability space (equipped with the square norm) into the space of all square integrable function under the measure $\lambda \otimes \mathbb{P}$ living on $([a, b] \times \Omega, \mathcal{B}[a, b] \otimes \mathcal{F}, \lambda \otimes \mathbb{P})$ equipped with the square norm.

⁹³That is, there exists k such that $f_t(\omega) < k$ for all t and $\omega \in \Omega$.

13.4 Ito calculus

We can construct more general diffusion processes as stochastic integrals of the form

$$S_t = \alpha + \int_0^t \mu_\tau d\tau + \int_0^t \sigma_\tau dB_\tau, \quad (13.5)$$

where $\alpha \in \mathbb{R}$, $\sigma_\tau \in L^2$, and $\mu_\tau \in L^1$, i.e. $\int_0^t |\mu_\tau| dt < \infty$ almost surely for all t . S_t is an example of an Ito-process. This diffusion process is often written in the differential form

$$dS_t = \mu_t dt + \sigma_t dB_t, \quad S_0 = \alpha. \quad (13.6)$$

However, we remind the reader that a Brownian motion is not differentiable. Equation (13.6) is only a short-hand notation for equation (13.5). The differential definition of a stochastic process is formally interpreted by adding the omitted integrals back to both sides of the equation.

Whereas we cannot differentiate a Brownian motion, we can differentiate the expected value of the stochastic process S_t , if μ_τ and σ_τ are continuous (adapted) processes. In particular, it holds

$$\begin{aligned} \left. \frac{d}{d\tau} \mathbb{E}[S_\tau | \mathcal{F}_t] \right|_{\tau=t} &= \mu_t \\ \left. \frac{d}{d\tau} \text{Var}_t(S_\tau) \right|_{\tau=t} &= \left. \frac{d}{d\tau} \mathbb{E}[S_\tau^2 | \mathcal{F}_t] - (\mathbb{E}[S_\tau | \mathcal{F}_t])^2 \right|_{\tau=t} = \sigma_t^2 \end{aligned}$$

where derivatives are taken from the right. Thus, we can interpret μ_t as the change of the expectation of S_t (conditional on the information available at time t) and σ_t as the change of the (conditional) standard deviation at time t . Commonly, we refer to the process μ_t as the *drift* and to the process σ_t as the *diffusion* of the stochastic process S_t . Let $dX_t = \tilde{\mu}_\tau d\tau + \tilde{\sigma}_\tau dB_\tau$. The *unique decomposition property* of Ito processes states that the two Ito processes S_t and X_t coincide for all t almost surely if and only if $\tilde{\mu}_\tau = \mu_\tau$ and $\tilde{\sigma}_\tau = \sigma_\tau$ almost everywhere.

We now adapt our differential calculus to evaluate functions of Ito processes. By definition of a Brownian motion in section 13.2, the standard deviation evolves with the root of the passing time. As a consequence, we have to

keep track of second order changes in the Brownian motion even when we are only interested in first order changes in time. Let $F : \mathbb{R} \times [0, \infty) \rightarrow \mathbb{R}$ be a twice continuous differentiable function of the stochastic process s_t defined in equation (13.6) and time. We are interested in the stochastic process $X_t = F(S_t, t)$. Expanding X_t first order in t and second order in dB_t delivers

$$dX_t = F_t dt + F_S dS_t + \frac{1}{2} F_{SS} (dS_t)^2 + o(dt) .$$

Inserting $\mu_t dt + \sigma_t dB_t$ for dS_t we obtain

$$dX_t = F_t dt + F_S \mu_t dt + F_S \sigma_t dB_t + \frac{1}{2} F_{SS} \sigma_t^2 (dB_t)^2 + o(dt) .$$

In the last equation, we omit terms of order dt^2 and of order $dt dB_t$. The intuition for conserving the term $(dB_t)^2$ builds on the second defining characteristic of the Brownian motion ($\Delta B_t = \epsilon \sqrt{\Delta t} \Rightarrow \mathbb{E} (\Delta B_t)^2 = \mathbb{E} \epsilon^2 \Delta t = \Delta t$) or, more formally, the argument of our calculation in the proof of Proposition 8. By the same argument, we replace $(dB_t)^2$ by dt to obtain the first order approximation of the total change of X_t over the infinitesimal time interval dt . A formal version of this reasoning yields Ito's lemma, a corner stone of stochastic calculus.

Proposition 9 (Ito's Lemma): Let S_t be an Ito process with $dS_t = \mu_t dt + \sigma_t dB_t$ and let $F : \mathbb{R} \times [0, \infty) \rightarrow \mathbb{R}$ be a twice continuously differentiable function. Then the process X_t defined by $X_t = F(S_t, t)$ is an Ito process with

$$dX_t = \left[F_S(S_t, t) \mu_t + F_t(S_t, t) + \frac{1}{2} F_{SS}(S_t, t) \sigma_t^2 \right] dt + F_S(S_t, t) \sigma_t dB_t .$$

The new or “non-standard” term $\frac{1}{2} F_{SS}(S_t, t) \sigma_t^2$ in the Ito formula represents a contribution from Jensen's inequality: if the function F is convex in S_t then a positive deviation of S_t from its expected value implies a higher gain than the loss implied by a negative deviation of the same magnitude. If the function is linear or if the variance is zero, then the additional term does not contribute. Thus, the new Ito term is a straight forward implication of our earlier analysis how Jensen's inequality affects payoffs under uncertainty, only that we have continuous fluctuations.

An application of Ito's formula analyzes the price process

$$X_t = c \exp[\alpha t + \sigma B_t], \quad t \geq 0, \quad (13.7)$$

$c, \alpha, \sigma \in \mathbb{R}$. Such a process is called lognormal because the logarithm of X_t is distributed normally.⁹⁴ To apply Ito's lemma, we choose $F(B_t, t) = c \exp[\alpha t + \sigma B_t]$ so that the underlying Ito process is the standard Brownian motion itself and $\mu_t = 0$ and $\sigma_t = 1$ and find

$$dX_t = \mu X_t dt + \sigma X_t dB_t, \quad X_0 = c, \quad (13.8)$$

where $\mu = \alpha + \frac{\sigma^2}{2}$. Alternatively, we could have chosen $F(S_t, t) = c \exp[S_t]$ with $S_t = \alpha t + \sigma B_t$ and obtained the same result. The process is known as *geometric Brownian motion* and frequently employed to model price behavior. The parameter μ specifies the "instantaneous" expected rate of return and σ the "instantaneous" standard deviation of the rate of return. The coefficient σ is often called the *volatility* of X_t . Equation (13.8) is a stochastic differential equation: the right hand side depends itself on the stochastic process X_t . Ito's lemma helped us to verify that this stochastic differential equation has the solution given by equation (13.7).

We close this section by stating Ito's formula for the case of a multidimensional Brownian motion. We assume that the Brownian motions B_t^1, \dots, B_t^N are independent and we collect them in (column) vector notation as \mathbf{B}_t . Then, an Ito process is of the form

$$S_t = \alpha + \int_0^t \mu_\tau d\tau + \int_0^t \boldsymbol{\sigma}_\tau d\mathbf{B}_\tau,$$

where $\alpha \in \mathbb{R}$, $\mu_\tau \in L^1$, and $\boldsymbol{\sigma}_\tau = (\sigma_\tau^1, \dots, \sigma_\tau^N)$ with $\sigma_\tau^1, \dots, \sigma_\tau^N \in L^2$. In differential form the process writes as

$$dS_t = \mu_t dt + \boldsymbol{\sigma}_t d\mathbf{B}_t, \quad S_0 = \alpha.$$

In general, we can have M such Ito processes, and our function can depend on all of those processes. Then, we define the vector of stochastic processes

$$d\mathbf{S}_t = \boldsymbol{\mu}_t dt + \boldsymbol{\underline{\sigma}}_t d\mathbf{B}_t, \quad S_0 = \alpha,$$

⁹⁴ $\log(X_t) = \log(c) + \alpha t + \sigma B_t$ and B_t is distributed normally.

where $\boldsymbol{\mu}_t$ and $\underline{\boldsymbol{\sigma}}_t$ are valued in \mathbb{R}^M and $\mathbb{R}^{M \times N}$ (underlining marks a matrix). We call such a process an Ito process in \mathbb{R}^M .

Proposition 10 (Ito's Lemma in \mathbb{R}^M): Let \mathbf{S}_t be an Ito process in \mathbb{R}^M as defined above and let $F : \mathbb{R}^M \times [0, \infty) \rightarrow \mathbb{R}$ be a twice continuously differentiable function. Then $F(\mathbf{S}_t, t)$ is an Ito process satisfying

$$\begin{aligned} F(\mathbf{S}_t, t) &= F(\mathbf{S}_0, 0) \\ &+ \int_0^t \left[\mathbf{F}_S(\mathbf{S}_\tau, \tau) \boldsymbol{\mu}_\tau + F_\tau(\mathbf{S}_\tau, \tau) + \frac{1}{2} \text{tr}(\underline{\boldsymbol{\sigma}}_\tau \underline{\boldsymbol{\sigma}}_\tau^\dagger \underline{\mathbf{F}}_{SS}(\mathbf{S}_\tau, \tau)) \right] d\tau \\ &+ \int_0^t \mathbf{F}_S(\mathbf{S}_\tau, \tau) \underline{\boldsymbol{\sigma}}_\tau d\mathbf{B}_\tau . \end{aligned}$$

Here, we stated Ito's lemma in integral form and emphasize that also Proposition 9 is formally interpreted by adding the stochastic integral to both sides of the equation. The combined derivative under the time integral is sometimes denoted by \mathcal{D} , i.e.,

$$\mathcal{D} F(\mathbf{S}_t, t) \equiv \mathbf{F}_S(\mathbf{S}_t, t) \boldsymbol{\mu}_t + F_t(\mathbf{S}_t, t) + \frac{1}{2} \text{tr}(\underline{\boldsymbol{\sigma}}_t \underline{\boldsymbol{\sigma}}_t^\dagger \underline{\mathbf{F}}_{SS}(\mathbf{S}_t, t)) , \quad (13.9)$$

which reduces the Ito formula to

$$F(\mathbf{S}_t, t) = F(\mathbf{S}_0, 0) + \int_0^t \mathcal{D} F(\mathbf{S}_\tau, \tau) d\tau + \int_0^t \mathbf{F}_S(\mathbf{S}_\tau, \tau) \underline{\boldsymbol{\sigma}}_\tau d\mathbf{B}_\tau . \quad (13.10)$$

Finally, we note several alternative formulation of the ‘‘Ito term’’ using $A = \underline{\boldsymbol{\sigma}}_\tau$ and the symmetry of the matrix $B = \underline{\mathbf{F}}_{SS}$: $\text{tr}(AA^\dagger B) = \sum_{i,j,k} A_{i,j} A_{j,k}^\dagger B_{k,i} = \sum_{i,j,k} A_{j,i}^\dagger B_{i,k} A_{k,j} = \text{tr}(A^\dagger B A) = \sum_{i,j} C_{i,j} B_{i,j}$ where $C = AA^\dagger = \underline{\boldsymbol{\sigma}}_\tau \underline{\boldsymbol{\sigma}}_\tau^\dagger$.

13.5 Stochastic Differential Equations

The previous section applied Ito's lemma to the geometric Brownian motion $X_t = c \exp[\alpha t + \sigma B_t]$ and arrived at the stochastic differential equation (13.8). In contrast, most economic question imply a stochastic differential equation (SDE) and our answer to the economic question requires us to solve the SDE. Equation (13.8) is a special case of an SDE of the form

$$dX_t = \mu(X_t, t) dt + \sigma(X_t, t) dB_t, \quad X_0 = c , \quad (13.11)$$

where $\mu : \mathbb{R} \times [0, \infty) \rightarrow \mathbb{R}$ and $\sigma : \mathbb{R} \times [0, \infty) \rightarrow \mathbb{R}$ are general functions.

One way of solving SDEs is by means of a sophisticated guess and Ito's lemma, just as we did in the previous section for the geometric Brownian motion. Sometimes, we can derive a successful candidate for guessing the solution from an analogous non-stochastic differential equation. A second approach is to simulate (a large number of sample paths of) the stochastic process and calculate its statistical properties numerically (see next paragraph). A third approach is to relate the solution of the SDE to the solution of a partial differential equation (PDE) and solve the corresponding PDE (or hope that someone else has solved it already).⁹⁵

A simple Monte Carlo simulation of the solution to the SDE (13.11) on a time interval $[0, T]$ follows the recipe:

1. Choose a time step Δt .
2. Start at X_0 .
3. Generate a stochastic time path:
 - draw $\epsilon_i \sim N(0, 1)$, i.i.d., for $i = 0, \dots, \frac{T}{\Delta t}$ and calculate
 - $X_{(i+1)\Delta t} = X_{i\Delta t} + \mu(X_{i\Delta t}, i\Delta t) \Delta t + \sigma(X_{i\Delta t}, i\Delta t) \sqrt{\Delta t} \epsilon_i$
for all $i = 0, \dots, \frac{T}{\Delta t}$.
4. Repeat steps 2-3 many times to estimate the distributional properties of X_t .

For most well-behaved stochastic processes such an Euler discretization will allow us to derive the desired statistical features of the solution. However, recall that Section 13.3 defined the general stochastic integral by means of the Ito isometry because the Brownian motion can behave rather ugly on small time intervals. For many stochastic processes these extreme fluctuations of the Brownian motion will “cancel out” in the numeric integration

⁹⁵The existence of solutions to general SDEs is itself a somewhat tricky topic. We will be looking for what are called strong solutions of an SDE. Strong solutions to the SDE are solutions that are stochastic processes on the original probability space, adapted to the filtration generated by the Brownian motion. There is a weaker solution concept referred to as “weak solutions”, where the processes solving the SDE live on a modified probability space.

process performed by the above algorithm. Then, decreasing the time step of the above algorithm, our solution, the estimated moments of our solution, converges. However, there are stochastic processes for which this simple approach will not work and the numeric algorithm above will blow up for small enough time steps, even if the stochastic integral implied by the SDE is well defined.

The third, and much employed solution approach relates the solution of an SDE to the solution of a corresponding PDE. At the heart of this solution method lies the *Feynman-Kac stochastic representation formula*.

Proposition 11 (Feynman-Kac): Assume that $F(t, x)$ is a solution to the boundary value problem

$$\begin{aligned} \frac{\partial F}{\partial t} + \mu(X_t, t) \frac{\partial F}{\partial x} + \frac{\sigma(X_t, t)^2}{2} \frac{\partial^2 F}{\partial x^2} &= 0 \\ F(T, x) &= \Phi(x). \end{aligned} \quad (13.12)$$

Assume furthermore that the process

$$\sigma(\tau, X_\tau) \frac{\partial F}{\partial x}(\tau, X_\tau) = 0$$

is in square integrable almost surely (where X is defined as below). Then F has the representation

$$F(t, x) = \mathbb{E}[\Phi(X_T) | \mathcal{F}_t],$$

where X satisfies the SDE

$$\begin{aligned} dX_t &= \mu(X_t, t) dt + \sigma(X_t, t) dB_t \\ X_t &= x. \end{aligned} \quad (13.13)$$

The result states that the solution of the partial differential equation, $F(t, x)$, returns the expected value of our last period value function at time t under the stochastic process (13.16). We note that the result has been derived originally by Feynman (for quantum mechanics) and Kac (for the heat diffusion equation) as a reinterpretation of the PDE opening up a way to simulate the solution to the PDE rather than vice versa. Yet, given that we know

the solution to several PDE's by now and that there are numeric methods available to solve PDE's it can be helpful to go the other way round.

The derivation of the simple Feynman-Kac formula is straight forward. We start by rewriting Ito's lemma in equation (13.10) for the time interval $[t, T]$ as

$$\begin{aligned} F(S_T, T) &= F(S_t, t) + \int_t^T \mathcal{D} F(S_\tau, \tau) d\tau + \int_0^t F_S(S_\tau, \tau) \sigma_\tau dB_\tau \\ \Leftrightarrow F(S_t, t) &= \Phi(S_T) - \int_t^T \mathcal{D} F(S_\tau, \tau) d\tau - \int_0^t F_S(S_\tau, \tau) \sigma_\tau dB_\tau . \end{aligned}$$

Assuming that the function F solves the PDE, $\mathcal{D} F(S_\tau, \tau)$ is everywhere zero. Moreover, if the process $F_S(S_\tau, \tau) \sigma_\tau$ is sufficiently well behaved, the stochastic integral vanished in expectation and we find

$$F(S_t, t) = \mathbb{E}[\Phi(S_T) | \mathcal{F}_t] .$$

The PDE (13.12) is homogenous and does not depend on F itself but only on its derivatives. The general form of the Feynman-Kac formula includes such dependencies.

Proposition 12 (Feynman-Kac, general): Assume that $F(t, x)$ is a solution to the boundary value problem

$$\begin{aligned} \frac{\partial F}{\partial t} + \mu(X_t, t) \frac{\partial F}{\partial x} + \frac{\sigma(X_t, t)^2}{2} \frac{\partial^2 F}{\partial x^2} - r(t, x)F + g(t, x) &= 0 \quad (13.14) \\ F(T, x) &= \Phi(x). \end{aligned}$$

Assume furthermore that F , g , and Φ satisfy some regularity and growth conditions. Then F has the representation

$$\begin{aligned} F(t, x) &= \mathbb{E} \left[e^{-\int_t^T r(\tau, X_\tau) d\tau} \Phi(X_T) \right. \\ &\quad \left. + \int_t^T g(\tau, X_\tau) e^{-\int_t^\tau r(s, X_s) ds} d\tau \middle| \mathcal{F}_t \right] , \end{aligned} \quad (13.15)$$

where X satisfies the SDE

$$\begin{aligned} dX_t &= \mu(X_t, t) dt + \sigma(X_t, t) dB_t \\ X_t &= x . \end{aligned} \quad (13.16)$$

In this general form, we interpret $r(t, x)$ as a potentially time and state dependent discount rate and $g(t, x)$ as a continuous payoff. The solution to the PDE (13.14) is the present discounted value of the final period payoff $\Phi(X_T)$ and the sum (integral) of the discounted payoff stream $g(t, x)$. Here, discounting takes place at a changing continuous time rate that can be stochastic itself. If the discount rate r was constant and deterministic, the representation in equation (13.15) simplifies to

$$F(t, x) = e^{-r(T-t)} \mathbb{E} \left[\Phi(X_T) \middle| \mathcal{F}_t \right] + \mathbb{E} \left[\int_t^T g(\tau, X_\tau) e^{-r(\tau-t)} d\tau \middle| \mathcal{F}_t \right]$$

fleshing out the discounted payoff interpretation of the expression.

13.6 Stochastic Dynamic Programming

We derive the Hamilton-Jacobi Bellman equation for the case of continuous stochastic processes governing the equations of motion or the objective function. The derivation is similar to the deterministic case. In difference to section 11, we have to acknowledge the second order term in Ito's formula when expanding the value function. We assume the technical underpinnings elaborated above, in particular, a probability space (Ω, \mathcal{F}, P) , a standard filtration \mathcal{F}_t capturing information available at time t , and the necessary integrability constraints on all processes. Let $x_t \in \mathbb{R}^L$ be a vector of states. A control in the stochastic framework is a, generally vector valued, adapted process c_t . We can have a controlled drift $g(c_t, x_t)$ as well as a controlled diffusion $h(c_t, x_t)$. We say that a control c_t is admissible given an initial state x_0 if there is a unique Ito process x^c such that

$$\begin{aligned} dx_t^c &= g(c_t, x_t^c) dt + h(c_t, x_t^c) dB_t, \\ x_0^c &= x_0. \end{aligned} \tag{13.17}$$

Equation (13.17) is our equation of motion. We use the controls to maximize an objective function of the form

$$\mathbb{E} \int_0^T U(c_t, x_t^c, t) dt + F(x_T^c),$$

where we assume that the expression is well defined for all admissible control paths.

We proceed as in the section on deterministic continuous time dynamic programming and motivate the Hamilton-Jacobi-Bellman equation for one state variable by taking the continuous time limit of the discrete time version. Making the discrete time step explicit, we have

$$J(x_t, t) = \max_{c_t} U(x_t, c_t, t) \Delta t + \mathbb{E} J(x_{t+\Delta t}, t + \Delta t) .$$

Expanding the value function on the right hand side in Δt yields

$$\begin{aligned} J(x_t, t) = \max_{c_t} U(x_t, c_t, t) \Delta t + \mathbb{E} \left[J(x_t, t) + J_x(x_t, t) \Delta x_t \right. \\ \left. + J_t(x_t, t) \Delta t + \frac{1}{2} J_{xx}(x_t, t) (\Delta x_t)^2 + o(\Delta t) \right] . \end{aligned}$$

Inserting $\Delta x_t = g(c_t, x_t) \Delta t + h(c_t, x_t) \Delta B_t$, and equating $\mathbb{E}(\Delta B_t)^2 = \Delta t$ under the expectation operator implies

$$\begin{aligned} 0 = \max_{c_t} U(x_t, c_t, t) \Delta t + \mathbb{E} \left[J_x(x_t, t) g(c_t, x_t) \Delta t + J_x(x_t, t) h(c_t, x_t) \Delta B_t \right. \\ \left. + J_t(x_t, t) \Delta t + \frac{1}{2} J_{xx}(x_t, t) h(c_t, x_t)^2 \Delta t \right] + o(\Delta t) . \end{aligned}$$

Using $\mathbb{E} J_x(x_t, t) h(c_t, x_t) dB_t = 0$, dividing by Δt , and letting $\Delta t \rightarrow 0$ results in the *Hamilton-Jacobi-Bellman equation* of the stochastic setting

$$\begin{aligned} 0 = \max_{c_t} U(x_t, c_t, t) + J_x(x_t, t) g(c_t, x_t) + J_t(x_t, t) \\ + \frac{1}{2} J_{xx}(x_t, t) h(c_t, x_t)^2 , \end{aligned} \tag{13.18}$$

where we eliminated the expected value operator as all functions are measurable at time t . Note that the continuous time formulation of our stochastic dynamic programming equation replaces the expectation operator with the “Ito or Jensen’s inequality term”.

We note that the derivatives in equation (13.18) correspond to $\mathcal{D} J(x, t)$ where \mathcal{D} is the differential operator defined in equation (13.9) and, in the general multidimensional case,

$$\mathcal{D} J(x_t, t) = J_x(x_t, t) g_t + J_t(x_t, t) + \frac{1}{2} \text{tr} \left(h_t h_t^\dagger J_{xx}(x_t, t) \right) .$$

Thus, we obtain the HJB in the stochastic setting by replacing the total derivative in the deterministic setting (equation 11.1) with the operator \mathcal{D}

$$0 = \max_{c_t} U(x_t, c_t, t) + \mathcal{D} J(x_t, t) . \quad (13.19)$$

In addition, we have the boundary condition

$$J(x_T, T) = F(x_T) . \quad (13.20)$$

Equation (13.19) is the general form of the HJB for multidimensional stochastic processes.

We can verify the optimality of a policy based on a value function solving the HJB in a similar way as we did in section 11. Let the function J solve the HJB (13.19) so that for any admissible (not necessarily optimal) control paths we have

$$0 \geq U(x_t, c_t, t) + \mathcal{D} J(x_t, t) , \quad (13.21)$$

with equality for the optimal control c^* . For any given admissible control path we find by Ito's formula

$$J(x_T^c, T) = J(x_0, 0) + \int_0^T \mathcal{D} J(x_\tau^c, \tau) dt + \int_0^T J_x(x_\tau^c, \tau) h(c_t, x_t^c) dB_t \quad (13.22)$$

Under the assumption that $J_x(x_\tau^c, \tau) h(c_t, x_t^c)$ is bounded, or some other assumption sufficient to make the integral with respect to the standard Brownian motion a martingale, the expectation of the last integral is zero. Adding $\int_0^T U(x_t^c, c_t, t) dt$ to both sides of equation (13.22), observing the boundary condition (13.20) and inequality (13.21), and taking expected values we find

$$\begin{aligned} \mathbb{E} \int_0^T U(x_t^c, c_t, t) dt + F(x_T^c) &= J(x_0^c, 0) \\ &+ \mathbb{E} \int_0^T \mathcal{D} J(x_\tau^c, \tau) + U(x_t^c, c_t, t) dt \\ &\leq J(x_0, 0) , \end{aligned}$$

with equality for the optimal control process c^* satisfying equation (13.21). Thus, c^* is indeed optimal and $J(x, 0)$ denotes the maximal value of the program.

Finally, as in the deterministic setting, for a discounted, autonomous, infinite horizon problem with $U(x_t, c_t, t) = u(x_t, c_t) \exp[-\delta t]$ we can transform the HJB to a stationary form where the value function $V(x) = J(x_t, t) \exp[\delta t]$ is independent of time. Substituting the new value function into equation (13.18) and dividing out the exponential yields

$$\delta V(x) = \max_{c_t} u(x_t, c_t) + V_x(x_t)g(c_t, x_t) + \frac{1}{2}V_{xx}(x_t)h(c_t, x_t) . \quad (13.23)$$

13.7 The Linear Quadratic Example under Uncertainty

As our first application of the continuous time stochastic dynamic programming equation (HJB) we revisit the (simplest non-trivial) linear quadratic example of section 11.4 and add a stochastic term to the equation of motion. The optimization problem is

$$\max E \int_0^\infty -\frac{ay^2 + bx^2}{2} \exp[-\delta t] dt \quad (13.24)$$

$$dx_t = ydt + \sigma x dB_t ,$$

with x_0 given, $a, b > 0$, and B_t being a standard Brownian motion. We note that the stochastic equation of motion has a controlled drift and an uncontrolled volatility that is proportional to the stock level. The HJB equation (13.18) is

$$-J_t(x, t) = \max_y -\frac{ay^2 + bx^2}{2} \exp[-\delta t] + J_x(x, t)y + \frac{1}{2} \underline{J_{xx}(x, t)\sigma^2 x^2} ,$$

where the underlined second order term is new to the stochastic setting.⁹⁶

In the stationary form of the HJB (equation 13.23), where we work with the stationary value function $V(x) = J(x_t, t) \exp[+\delta t]$, we find

$$\delta V(x) = \max_y -\frac{ay^2 + bx^2}{2} + V_x(x)y + \frac{1}{2} \underline{V_{xx}(x)\sigma^2 x^2} .$$

⁹⁶Instead of conveniently switching to the stationary HJB, we could as well solve equation (13.24). The first order condition results in $y = \frac{J_x(x, t)}{a} \exp(\delta t)$. Inserting this solution back into the HJB equation (13.24) yields the partial differential equation $-J_t(x, t) = \frac{J_x(x, t)^2}{2} \exp[\delta t] - \frac{bx^2}{2} + \frac{1}{2} J_{xx}(x, t)\sigma^2 x^2$. Separating out a multiplicative time dependence at this point will lead us to the ODE that we are about to derive from the stationary HJB.

Maximizing over y gives the same result as earlier in Section 11 because the second order term is independent of the control y . We find

$$-ay + V_x(x) = 0 \quad \Rightarrow \quad y = \frac{V_x(x)}{a} . \quad (13.26)$$

Reinserting the solution into the HJB equation (13.25) delivers the (inhomogenous) second order differential equation

$$\delta V(x) = \frac{V_x(x)^2}{2a} - \frac{bx^2}{2} + \frac{V_{xx}(x)\sigma^2 x^2}{2} . \quad (13.27)$$

We note that we are back to an ordinary deterministic differential equation without stochasticity. Had we continued with the non-stationary HJB equation (13.24) we would end up with a deterministic PDE instead (see footnote ??). We try the same form for the value function that solved the analogous problem in the deterministic setting: $V(x) = \alpha x^2$. The trial solution transforms equation (13.27) into the algebraic equation

$$\begin{aligned} \delta \alpha x^2 &= \frac{2\alpha^2 x^2}{a} - \frac{bx^2}{2} + \alpha \sigma^2 x^2 \\ \Leftrightarrow \left(\delta \alpha - \frac{2\alpha^2}{a} + \frac{b}{2} - \alpha \sigma^2 \right) x^2 &= 0 , \end{aligned}$$

which is satisfied for all x whenever $x = 0$ or

$$\begin{aligned} \alpha^2 - \frac{a(\delta - \sigma^2)}{2} \alpha &= \frac{ab}{4} \\ \Leftrightarrow \alpha &= \frac{a}{4} \left(\delta - \sigma^2 \pm \sqrt{(\delta - \sigma^2)^2 + \frac{4b}{a}} \right) . \end{aligned}$$

Having found this solution, we also verified that the validity of functional form of our trial solution. We can characterize the difference between the deterministic and the stochastic solution for the value function coefficient α by replacing the discount rate in the deterministic setting by $\delta \rightarrow \delta - \sigma^2$ in the stochastic setting. The value function is

$$\begin{aligned} J(x, t) &= V(x) \exp[-\delta t] = \alpha x^2 \exp[-\delta t] \\ &= \frac{a}{4} \left(\delta - \sigma^2 - \sqrt{(\delta - \sigma^2)^2 + \frac{4b}{a}} \right) x^2 \exp[-\delta t] , \end{aligned}$$

where we picked the negative root because we know that the value function is negative.⁹⁷ Observe that the future value $J(x, t)$ is a stochastic process driven by the realizations of the stochastic process x_t .

We obtain the optimal control in feedback form by inserting the value function derivative $V_x(x) = 2\alpha x$ into the first order condition for our control, equation (13.26)

$$y(x) = \frac{V_x(x)}{a} = \frac{2\alpha x}{a} = \frac{1}{2} \underbrace{\left(\delta - \sigma^2 - \sqrt{(\delta - \sigma^2)^2 + \frac{4b}{a}} \right)}_{\equiv \gamma} x .$$

We note that the control rule depends on the volatility σ . The principle of certainty equivalence does not hold because of the *multiplicative uncertainty* in the equation of motion captured by the term $\sigma x_t dB_t$. Keeping in mind that $\alpha < 0$ we find that $\frac{\partial \gamma}{\partial \sigma} < 0$. Because γ is negative, uncertainty implies that our control tries harder to drive the stock x_t to zero under uncertainty. The actual control path is a stochastic process because it depends on the realization of the Brownian motion driving the stock.

We observe from equation (13.28) that the optimally controlled state x_t follows the SDE

$$dx_t = \gamma x dt + \sigma x dB_t , \quad (13.28)$$

where $\gamma < 0$. From our applications of Ito's lemma to the geometric Brownian motion in Section 13.4 we know that the solution to the SDE (13.28) is the geometric Brownian motion

$$x_t = x_0 \exp \left[\left(\gamma - \frac{\sigma^2}{2} \right) t + \sigma B_t \right] .$$

The optimally controlled stock follows a geometric Brownian motion whose drift presses the stock's evolution to zero ($\gamma < 0$), but whose stochastic term with volatility σ will continuously bounce the stock around. Note that in the long run the damping through the drift will outweigh the bouncing Brownian motion. The linear term in the drift outgrows the volatility of the Brownian motion, which grows less than linearly in time.

⁹⁷Note, however, that for large σ^2 the expression can be negative also for the positive root.

13.8 Resource Extraction under Uncertainty

This subsection discusses (Pindyck 1980) classic paper on resource extraction under uncertainty. The model extends the standard resource extraction problem with stock dependent extraction costs to a setting where demand has a component following a geometric Brownian motion and the estimate of available resource deposits follows a standard Brownian motion.

The market demand function is given as

$$p = y_t f(q_t) ,$$

where $f(q)$ with $f'(q) < 0$ is the deterministic part of the demand function depending on the extraction flow of the resource q . The random demand component y_t follows a geometric Brownian motion⁹⁸

$$dy_t = \alpha y_t dt + \sigma_y y_t dB_t^{i=1} .$$

The reserves R , or rather the amount of the resource expected to exist and to be extractable, also follow a random process given by

$$dR_t = -q_t dt + \sigma_R dB_t^{i=2} .$$

The representative firm maximizes the objective

$$\max_q \mathbb{E} \int_0^T \underbrace{[y_t f(q_t) - C(R_t)] q_t \exp^{-rt}}_{\equiv \Pi(y_t, R_t, t)} dt ,$$

where $C(R)$ denote average production costs satisfying $C'(R) < 0$. For a competitive firm, we assume that it treats $f(q) = \bar{f}$ as an exogenous parameter in its optimization problem.

⁹⁸We write $dB_t^{i=1}$ for first Brownian motion in order to avoid taking the index of the second Brownian motion $dB_t^{i=2}$ mistakenly for a square.

The HJB equation for the problem is⁹⁹

$$\begin{aligned}
 0 &= \max_q \Pi(y_t, R_t, t) + \mathcal{D} J(y_t, R_t, t) & (13.29) \\
 &= \max_q \Pi(y_t, R_t, t) + J_t(y_t, R_t, t) + J_R(-q_t) + J_y \alpha y_t \\
 &\quad + \frac{1}{2} J_{yy} y_t^2 \sigma_y^2 + \frac{1}{2} J_{RR} \sigma_R^2 .
 \end{aligned}$$

Carrying out the maximization yields

$$\Pi_q(y_t, R_t, t) = J_R(y_t, R_t, t) . \quad (13.30)$$

Note that in a competitive equilibrium profits are linear in q and, thus, equation (13.30) is not a priori valid for solving the optimization problem. However, if $\Pi_q > J_R$ firms would produce at maximum capacity, and if $\Pi_q < J_R$ firms would not produce at all. Thus, market clearing will imply that in the equilibrium $J_R = \Pi_q = \frac{\Pi}{q}$.

Rather than plugging the latter equation into the HJB in order to obtain the corresponding SDE, we use these two equations to eliminate the value function and derive the Euler equation. For this purpose, we first differentiate equation (13.29) with respect to the resource stock yielding

$$\begin{aligned}
 0 &= \frac{d}{dR} \Pi(y_t, R_t, t) + \frac{d}{dR} \mathcal{D} J(y_t, R_t, t) \\
 &= \Pi_R(y_t, R_t, t) + \mathcal{D} J_R(y_t, R_t, t) . & (13.31)
 \end{aligned}$$

Two remarks apply to this calculation. First, we assumed $\frac{d}{dR} \mathcal{D} J(y_t, R_t, t) = \mathcal{D} \frac{d}{dR} J(y_t, R_t, t)$. You can carry out the differentiations on both sides to see that the assumption comes down to the exchangeability of the third

⁹⁹In applying Ito's formula equation (13.9) we have

$$dS_t = \begin{pmatrix} \alpha y_t \\ -q_t \end{pmatrix} dt + \begin{pmatrix} y_t \sigma_y & 0 \\ 0 & \sigma_R \end{pmatrix} dB_t$$

leading in particular to the term

$$\frac{1}{2} \text{tr} \left[\sigma_t \sigma_t^\dagger F_{SS}(S_t, t) \right] = \frac{1}{2} \text{tr} \left[\begin{pmatrix} y_t^2 \sigma_y^2 & 0 \\ 0 & \sigma_R^2 \end{pmatrix} \begin{pmatrix} J_{yy} & J_{yR} \\ J_{Ry} & J_{RR} \end{pmatrix} \right] = \frac{1}{2} y_t^2 \sigma_y^2 J_{yy} + \frac{1}{2} \sigma_R^2 J_{RR} .$$

order derivatives of the function J and, thus, an assumption of sufficient smoothness of the value function. Second, we took a derivative with respect to a stochastic process. On what foundation can we do so? Assume we have two twice continuously differentiable functions $F^1(x_1, x_2) = F^2(x_1, x_2)$, depending on the same set of stochastic processes $dx_1 = \mu_{x_1}dt + \sigma_{x_1}dB^1$ and $dx_2 = \mu_{x_2}dt + \sigma_{x_2}dB^2$. Expressing the resulting processes in differential form implies

$$\begin{aligned} & \mathcal{D} F^1(x_1, x_2)dt + F_{x_1}^1 \sigma_{x_1} dB^1 + F_{x_2}^1 \sigma_{x_2} dB^2 \\ &= \mathcal{D} F^2(x_1, x_2)dt + F_{x_1}^2 \sigma_{x_1} dB^1 + F_{x_2}^2 \sigma_{x_2} dB^2 \end{aligned}$$

so that the unique decomposition property of Ito processes implies $F_{x_1}^1 = F_{x_1}^2$ and $F_{x_2}^1 = F_{x_2}^2$.

The second step on our way to the Euler equation proceeds in differentiating equation (13.30) by \mathcal{D} , again relying on the unique decomposition property and comparing the deterministic parts, yielding

$$\mathcal{D} \Pi_q(y_t, R_t, t) = \mathcal{D} J_R(y_t, R_t, t) . \quad (13.32)$$

Combining equations (13.31) and (13.32) we obtain

$$\begin{aligned} & \mathcal{D} \Pi_q(y_t, R_t, t) = -\Pi_R(y_t, R_t, t) \quad (13.33) \\ \Rightarrow & \int_t^T \mathcal{D} \Pi_q(y_t, R_t, t) dt = - \int_t^T \Pi_R(y_t, R_t, t) dt \\ \Rightarrow & \underbrace{\Pi_q(y_T, R_T, T)}_{=0(\text{terminal condition})} - \Pi_q(y_t, R_t, t) - \int_t^T \Pi_{qy} y_t \sigma_y dB_t^{i=1} \\ & - \int_t^T \Pi_{qR} \sigma_R dB_t^{i=2} = - \int_t^T \Pi_R(y_t, R_t, t) dt \\ \Rightarrow & \Pi_q(y_t, R_t, t) = \int_t^T \Pi_R(y_t, R_t, t) dt \quad (13.34) \\ & - \int_t^T \Pi_{qy} y_t \sigma_y dB_t^{i=1} - \int_t^T \Pi_{qR} \sigma_R dB_t^{i=2} \end{aligned}$$

Taking expectations in equation (13.34) returns

$$\Pi_q(y_t, R_t, t) = E \int_t^T \Pi_R(y_t, R_t, t) dt . \quad (13.35)$$

Equation (13.35) tells us that the increase in profits from selling one more unit should equal the expected sum of the discounted future increase in profit that would result from holding that unit in the ground (thereby reducing extraction cost).¹⁰⁰

Returning to the stochastic equation (13.34), and evaluating it at time t , the marginal profits on the left hand side of the equation are known. But what about the stochastic terms on the right hand side? The Brownian motions determine the stochastic adapted process Π_R in the first integrand on the right hand side. Indeed, the profits from keeping one more unit in the ground depend on how the reserves and the demand evolve over time. Writing the equation in differential form yields¹⁰¹

$$d\Pi_q(y_t, R_t, t) = -\Pi_R(y_t, R_t, t)dt + \Pi_{qy}y_t\sigma_y dB_t^{i=1} + \Pi_{qR}\sigma_R dB_t^{i=2} \quad (13.36)$$

While in equation (13.34) the point in time t at which we evaluated marginal profits was fix, in equation (13.36) we have evaluation time evolve. Then we see that the marginal profits from selling a unit of the resource fluctuate with the volatility of demand and the updates of the reserve estimates.

We can further calculate the precise form of the Euler equation by further evaluating the profit function. Let us do this for the example of a competitive equilibrium. Here profits are linear in q and $J_R = \Pi_q = \frac{\Pi}{q}$. Then, dividing by the exponential, equation (13.33) becomes

$$-r[f(q_t)y_t - C(R_t)] - C'(R_t)(-q) - \frac{1}{2}C''(R_t)\sigma_R^2 + f(q_t)\alpha y_t = qC'(R_t)$$

Setting $p_t = y_t f(q_t)$ implying $\frac{d}{dt}E p_t = f(q_t)\alpha y_t$ delivers

$$\frac{d}{dt}E p_t = r[p - C(R_t)] + \frac{1}{2}C''(R_t)\sigma_R^2 . \quad (13.37)$$

¹⁰⁰And selling it in the last period where by assumption the marginal profits are zero (terminal condition).

¹⁰¹Write $dP_i q$ in differential form using the Ito formula and equation (13.33).

Equation (13.37) is the Hotelling rule for the price of the resource extended to the stochastic setting. It holds for the expected price change. The new term $\frac{1}{2}C''(R_t)\sigma_R^2$ stems from the fact that high realizations of the random variable increase the production costs more than low realizations reduce the costs. The term accelerates the price increase for convex costs. Interestingly, for extraction costs that only depend linearly on the resource stock, the expected price change (conditional on the current reserve estimate) is the same as in a certain world. Moreover, demand uncertainty has no influence on the expected price path. We refer to Pindyck (1980) for exploring how uncertainty has an ambiguous effect on the expected extraction rate, which does depend on second order terms of the demand function.

13.9 Appendix

The appendix provides more insight into the technicalities surrounding Brownian motion and stochastic integration. Our approach is heuristic.

13.9.1 More on Brownian Motion

We briefly point out what it means that the quadratic variation of any continuously differentiable function vanishes, whereas the quadratic variation of a Brownian motion is positive. In consequence, the absolute variation of a Brownian motion diverges and a Brownian motion is not continuously differentiable on any finite time interval. Define a sequence of partitions of the interval $[a, b]$ by $a = t_0^n < t_1^n < t_2^n < \dots < t_n^n = b$ with $\max\{|t_{k+1}^n - t_k^n|, k = 1, 2, \dots, n-1\} \rightarrow 0$ for $n \rightarrow \infty$. Let $f \in C^1([a, b])$ be a continuously differentiable function on the interval $[a, b]$ and let B be a Brownian motion on the same interval.

1. The quadratic variation of f vanishes:

$$\sum_{k=0}^{n-1} [f(t_{k+1}^n) - f(t_k^n)]^2 \rightarrow 0 \text{ for } n \rightarrow \infty$$

2. The quadratic variation of B is $b - a$:

$$\sum_{k=0}^{n-1} [B_{t_{k+1}^n} - B_{t_k^n}]^2 \rightarrow b - a \text{ for } n \rightarrow \infty$$

3. The total variation of f is finite:

$$\sup_{\text{all partitions}} \sum_{k=0}^{n-1} |f(t_{k+1}^n) - f(t_k^n)| = \int_a^b |f'(x)| dx$$

4. The total variation of B diverges:

$$\sup_{\text{all partitions}} \sum_{k=0}^{n-1} |B_{t_{k+1}^n} - B_{t_k^n}| = \infty \quad \text{almost surely} \quad (13.38)$$

Given results 3 and 4 we know that a Brownian motion is not continuously differentiable on any interval $[a, b] \subset \mathbb{R}_+$. In fact, it can be shown that every path of a Brownian is almost surely nowhere differentiable.

13.9.2 Stochastic Integration

Recall how we usually define an integral. We start with a definition for a class of basic functions like step function (Riemann integral) or primitive functions (Lebesgue integral, measure integral). Then, we show that we can approximate a more general class of functions by sequences of these primitive functions, and we define the limit of the corresponding sequence of integrals of the approximating functions to be the integral of the function itself. The same procedure is used to define the stochastic integral. It is straight forward to define the integral for simple functions of the form

$$f(t) = \sum_{k=0}^{n-1} c_k 1_{[t_k, t_{k+1})}(t), \quad (13.39)$$

where $1_{[s,t)}(t)$ is an indicator function that is unity for $t \in [s, t)$ and zero everywhere else. Then we define the so called Wiener-integral of f by

$$\int_a^b f(t) dB_t \equiv \sum_{k=0}^{n-1} c_k (B_{t_{k+1}} - B_{t_k}). \quad (13.40)$$

For a standard integral we could approximate general function by a sequence of the primitive functions and define the integral by limit. However, equation (13.38), i.e. the infinite total variation of a Brownian motion, should make us cautious with respect to simply taking a limit in an equation of the form (13.40). Instead, the general definition of a Wiener and a stochastic integral makes use of the Ito-isometry, which states that for any function f of the form (13.39) holds

$$\mathbb{E} \left(\int_a^b f(t) dB_t \right)^2 = \int_a^b |f(t)|^2 dt.$$

This isometry is used to extend the definition of the integral to the space of square-integrable functions $f \in L^2$. In a further step, the Wiener integral is extended to the stochastic integral by allowing the integrand f to be a

stochastic process itself. For this purpose, let F define the filtration generated by the Brownian motion and let f be an F -adapted square integrable process. The same isometry as above is used to extend the definition of the integral from the simple functions of the form (13.39) with random coefficients c_j to the set of all stochastic processes that are square integrable in expectation.

13.10 Problems

Exercise 7 *Solve the following linear quadratic optimal control problem.*

$$\max \int_0^\infty -\frac{ay^2 + bx^2}{2} \exp[-\delta t] dt \quad s.t. \quad dx_t = ydt + \sigma dB_t ,$$

x_0 given, $a, b > 0$, and B_t a standard Brownian motion.

- a) *Write down the regular as well as the stationary HJB equation. How do they differ from the analogous problem under certainty? How do they differ from the problem discussed in the chapter on stochastic control?*
- b) *Solve the problem, deriving the value function, the optimal control rule, and the equation for the optimally controlled stock. Explain your steps and discuss the resulting equations. Does the principle of certainty equivalence hold? Compare this result to the findings in the chapter on discrete time stochastic control and explain possible differences.*
- c) *Solve the stochastic differential equation describing the equation of motion of the optimally controlled stock.*
Procedure: "Guess" that the solution is of the form

$$X_t = \alpha(t) \left(X_0 + \int_0^t \beta(\tau) dB_\tau \right) .$$

Take the total differential. You do not need much of Ito's formula because the Brownian motion is already under the integral. The total change of the integral at time t is merely the $\beta(\tau)dB_\tau$ evaluated at time t . Bring the result to a form that you can compare to the actual controlled equation of motion for the stock S_t . Match coefficients, which might involve solving a simple differential equation.

d) Calculate future expected values of the stock and its variance.

14 Discounting

Consumption today is not directly comparable with consumption at a different point in time. The discount factor for consumption enables us to compare consumption at different points in time. Discounting is an especially important element of environmental problems that involve trade-offs in consumption across widely different times. Climate policy is the leading example of this kind of trade-off, because decisions taken in the near future may have major effects on welfare in the distant future.

14.1 The social discount rate

Discount rates are defined as the rate of decrease (the negative of the rate of change) of the discount factor. It is important at the outset to distinguish between discount rates and factors for utility and for consumption. We define β_t as the number of units of utility (utils) that we would give up today in order to obtain one more util at time t . It is the discount factor for utility. By definition, $\beta_0 = 1$. We define the discount rate for utility at time t as¹⁰²

$$\rho_t = -\frac{\dot{\beta}_t}{\beta_t},$$

where the dot denotes the (total) time derivative. The utility discount rate ρ_t is also known as the rate of pure time preference (RPTP). The RPTP is a measure of impatience, with larger values implying greater impatience. If $\rho_t = \rho$ is a constant, utility discounting is exponential: $\beta_t = e^{-\rho t}$.

We begin by defining the discount factor and the corresponding discount rate for consumption in the simplest case: there is a single consumption good, c ; there is no uncertainty; and welfare, W , equals the present value of the infinite stream of utility, $u(c)$. In this case, $W = \int_0^\infty \beta_t u(c_t) dt$. The consumption discount factor for time t equals the number of units of consumption we would surrender during a small interval, ε , beginning today in order to obtain one more unit of consumption during a small interval, ε , beginning at time t . If, prior to the transfer, consumption today is c_0 and consumption at time t is c_t , the welfare loss due to giving up Γ units of

¹⁰²Here we define the instantaneous discount rate. Another frequent definition of discount rates is as an average rate defined by $\frac{1}{t} \ln \beta_t$.

consumption today is approximately $u'(c_0)\Gamma\varepsilon$ and the welfare gain of one unit of consumption at time t is $\beta_t u'(c_t)\varepsilon$. We are willing to make this sacrifice if these two quantities are equal, i.e. if

$$\Gamma_t = \beta_t \frac{u'(c_t)}{u'(c_0)}. \quad (14.1)$$

The rate of decrease of Γ is the discount rate for consumption. This rate is more conveniently expressed in terms of the growth rate of consumption g and the consumption elasticity of marginal utility η , which is equal to the inverse of the elasticity of intertemporal substitution. These are defined as

$$g_t = \frac{\dot{c}_t}{c_t} \quad \text{and} \quad \eta(c_t) = -\frac{u''(c_t)}{u'(c_t)} c_t.$$

Then, equation (14.1) gives rise to the consumption discount rate

$$r_t = -\frac{\dot{\Gamma}_t}{\Gamma_t} = \rho_t + \eta(c_t) g_t. \quad (14.2)$$

Equation (14.2) is usually referred to as the Ramsey equation. More precisely, the actual Ramsey equation is an equilibrium condition in the Ramsey-Cass-Koopmans growth model stating that the right hand side of equation (14.2) has to equal the interest rate (or capital productivity) in the economy. In contrast, the derivation of the consumption discount rate r_t in equation (14.2) is independent of the market equilibrium. In the context of public project evaluation, the consumption discount rate r_t is referred to as the social discount rate (SDR).

A larger SDR means that we are willing to sacrifice fewer units of consumption today in order to obtain an additional unit of consumption t periods in the future. In the context of climate policy, a larger SDR means that we are willing to spend less today, e.g. through increased abatement or investment in low-carbon technology, to prevent future climate damage. A larger value of the RPTP, ρ , means that we place less value on future utility. A higher growth rate means that future consumption is higher; under the assumption of decreasing marginal utility of consumption the higher g lowers future marginal utility. A larger elasticity of marginal utility implies a faster decrease of marginal utility with consumption growth. Therefore, under positive growth, larger values of ρ , g , or η all imply a higher SDR, and less concern for the future.

Some applications assume: (i) isoelastic utility $u(c) = \frac{c^{1-\eta}}{1-\eta}$, so that η is constant; (ii) a constant growth rate, so that g is constant; and (iii) exponential discounting of utility, so ρ is constant. In this case, the SDR is also constant. More generally, one or more of the three components of r_t might depend on time. While g_t or $\eta(c_t)$ will quite commonly depend on time because of the dynamics in the economy, a time dependence of pure time preference would be exogenously imposed as a direct preference specification. As we discuss in section 14.6.2 such a time dependence of pure time preference can lead to time inconsistent choices.

14.2 The SDR and environmental policy

The social discount rate is used to evaluate legislation and public projects. In application, the employed values vary widely over countries and agencies. While the majority adopts a constant rate, the U.K. and France adopt time dependent discounting schemes. The social discount rate is important in evaluating environmental policy when the timing of costs and benefits differ, as with climate change policy where current decisions have long-lasting effects. We use the latter as an example to illustrate the policy relevance of the social discount rate. The Stern (2007) Review of Climate Policy uses a negligible RPTP of $\rho = 0.1\%$, a growth rate of $g = 1.3\%$, and the value $\eta = 1$, implying $r = 1.4\%$. In contrast, Nordhaus (2008) employs a RPTP of $\rho = 1.5\%$ and the value $\eta = 2$ in a model with an approximate consumption growth rate of $g = 1.5\%$, implying $r = 5.5\%$. The ratio of what we are willing to spend today, to avoid a dollar of damage 100 years from now, under these two sets of parameters is

$$\frac{\Gamma_{100}^{Stern}}{\Gamma_{100}^{Nord}} = \frac{e^{-.014 \cdot 100}}{e^{-.055 \cdot 100}} \approx 60.$$

For this example, the higher SDR decreases our willingness to spend today to avoid damages a century from now by a factor of 60. Nordhaus (2007) shows that this difference in social discounting can explain almost entirely the differences in policy recommendation between his integrated assessment of climate change based on DICE-2007 and that of the Stern Review: running DICE with a 1.4% SDR instead of 5.5% increases the near term social cost of carbon by a factor of 10 and almost quadruples the near term optimal abatement rate with respect to business as usual.

14.3 The positive and the normative perspective

The different choices of the social discount rate in (Nordhaus 2008) and (Stern 2007) assessment of climate change stem from different perspectives on the application of social discounting in policy evaluation. Nordhaus takes a positive approach to social discounting, while Stern takes a normative approach. The positive approach relies on measurements of the constituents of the social discount rate, while the normative approach chooses these parameters on ethical grounds. The measurement of the social discount rate is complicated by the fact that markets exhibit externalities, are incomplete, and, therefore, do not necessarily reflect the agents' intertemporal preferences correctly.

In principle, there are two different approaches to determine the social discount rate as reflected on the market. First, we can measure ρ and η based on a sufficiently large set of observations. We then combine these estimates with an exogenous consumption growth scenario, or use them in an evaluation model where growth is determined endogenously, as in the integrated assessment of climate change. Second, we can measure the interest rate in the economy. Then, the original derivation of the Ramsey (1928) equation (14.2) states that in equilibrium this interest rate is equivalent to the consumption discount rate. This second method is particularly prone to picking up market imperfections like transaction costs or distortions in the intertemporal consumption-investment trade-off. These market imperfections also result in a wide spectrum of different interest rates observable on the market. Usually, interest rate based approaches to measuring the social discount rate rely on the interest paid on government bonds. These provide an opportunity cost measure for a dollar spent on public projects. The maturity of government bonds limits how far into the future we can measure this opportunity cost; in the U.S. it is currently given by the 30-year treasury bond.

The Stern (2007) Review argues that intergenerational trade-offs encompassing many generations cannot be judged merely on the basis of market observations. Society has to employ ethical reasoning in order to represent those generations that are not currently alive and, hence, not reflected on the market. The sequence of economists who argued that ethical reasoning imposes a zero RPTP is long and includes Ramsey (1928), Pigou (1932), Harrod (1948), Koopmans (1963), Solow (1974), Broome (1992). While the Stern Review's choice of a close to zero RPTP is based on intergenerational

equity concern, it simultaneously adopts a comparatively low value for the propensity to smooth consumption over time η , implying a rather low preferences for intergenerational consumption smoothing. Traeger (2007) presents a different normative argument for a zero RPTP based entirely on rationality constraints for decision making under uncertainty, rather than on ethical arguments. Schneider, Traeger & Winkler (2012) extend equation (14.2) to account for overlapping generations. They reveal strong normative assumptions underlying the positive approach and show that the usual arguments of the normative approach are complicated by an equity trade-off between generations alive at the same point in time versus equity over generations across time.

14.4 Discounting under Uncertainty

The social discount rate relies on future consumption growth, which is uncertain. Within the standard model, only strongly serially correlated or catastrophic risk have a serious impact on the discount rate. We briefly discuss two extensions that incorporate general risk aversion and ambiguity aversion into the social discount rate, showing that these can have a major impact on the discount rate. We close the section commenting on (Weitzman 2009) Dismal Theorem and the Weitzman-Gollier puzzle.

14.5 Stochastic Ramsey equation

The social discount rate under uncertainty is generally defined using a certain consumption trade-off as in section 14.1 shifting consumption into an uncertain future. Then, the resulting consumption discount factor Γ_t captures the present value of a certain consumption unit in an uncertain future. As a consequence, the right hand side of equation (14.1), defining Γ_t , gains an expected value operator expressing that marginal utility gained from an additional future consumption unit is uncertain. For the subsequent analysis, we assume two periods, isoelastic preferences $u(c) = \frac{c^{1-\eta}}{1-\eta}$, and a normal distribution of the growth rate $\tilde{g} = \tilde{c}_1/c_0$ with expected growth rate μ and standard deviation σ . Then the consumption discount rate is

$$r = \delta + \eta\mu - \eta^2 \frac{\sigma^2}{2} . \quad (14.3)$$

The contributions of time preference and of expected growth coincide with the corresponding terms under certainty in equation (14.2). The third term $-\eta^2 \frac{\sigma^2}{2}$ results from consumption growth risk and decreases the social discount rate, increasing the present value of a certain consumption unit in the future. It is proportional to the growth variance σ^2 and the square of the consumption elasticity of marginal utility η . In the current context, η is frequently interpreted as a measure of risk aversion. However, it still is the measure of aversion to intertemporal substitution and section 14.5.1 explores a model incorporating general risk attitude.

We can interpret the timing in our setting in two different ways. First, the time between the first and the second period can be one year. Then, δ , μ , and σ will generally be in the order of percent. For example, Kocherlakota (1996)

estimates $\mu = 1.8$ and $\sigma = 3.6$ for almost a century of U.S. consumption data. Then, the risk term in equation (14.3) will be an order of magnitude smaller than the other terms: for $\eta = 2$ ($\eta = 1$) the growth contribution is $\eta\mu = 3.6\%$ (1.8%) while the risk contribution is 0.5% (0.1%). Under the assumption of an iid growth process, equation (14.3) captures the constant, annual social discount rate.

Second, we can interpret period 0 as the investment time of a project, and period 1 as the payoff period. Assuming a constant annual expected growth rate, the two first terms on the right hand side of equation (14.3) increase linearly in time. Dividing the equation by the time span t between investment and payoff yields the average annual consumption discount rate. The first two contributions to this average rate are just as in (14.3), while the risk term becomes $-\eta^2 \frac{\sigma^2}{2t}$. For a random walk of the growth rate, the variance grows linearly in time, confirming the result that an iid growth process implies a constant annual discount rate. However, under serial correlation the variance increases more than linearly in time and the risk term increases the further the payoff lies in the future. Then, long term payoffs are discounted at a lower discount rate than short term payoffs: the term structure of the social discount rate decreases. Due to this finding, France and the U.K. adopted falling discounting schemes for project evaluation. We discuss the case of perfect serial correlation in more detail in section 14.5.4

14.5.1 General risk attitude

Equation (14.3) is based on the intertemporally additive expected utility model. In this model, the consumption elasticity of marginal utility has to capture two distinct preference characteristics: the propensity to smooth consumption over time and risk aversion. Positively, these two attitudes differ. Also normatively, there is no reason that the two should coincide.¹⁰³ In general, risk affects economic evaluation in two different ways. First, a stochastic process generates wiggles in the consumption paths. Agents dis-

¹⁰³It is a widespread misunderstanding that (von Neumann & Morgenstern 1944) axioms, underlying the expected utility model, together with additive separability on certain consumption paths, underlying the discounted utility model, would imply the discounted expected utility model. Merging these two assumptions results in a more general evaluation framework that distinguishes risk attitude from the propensity to smooth consumption over time (Traeger 2010a).

like these wiggles if they have a propensity to smooth consumption over time. Second, agents dislike risk merely because it makes them uncertain about the future. This is an intrinsic aversion to risk that is not captured in the intertemporally expected utility standard model. The finance literature has successfully exploited general risk attitude to explain various asset pricing puzzles. In the context of determining the social discount rate, the most important puzzles solved are the equity premium and the risk free rate puzzles. Resolving these puzzles requires a model that gets the risk free rate right and explains the premium paid on risky equity. In a positive approach, where we measure preferences or interest rates based on market observations, it is important to use a model that gets these rates right. In a normative approach, the model forces the decision maker to think about both risk aversion and intertemporal (or intergenerational) consumption smoothing.

We keep the assumptions of a normal distribution of the growth rate and of isoelastic preferences, now with respect to both: consumption smoothing over risk and over time. Calling the measure of intrinsic risk aversion RIRA for relative intertemporal risk aversion, Traeger (2008) derives the social discount rate

$$r = \delta + \eta\mu - \eta^2 \frac{\sigma^2}{2} - \text{RIRA} |1 - \eta^2| \frac{\sigma^2}{2}. \quad (14.4)$$

Here, the parameter η only expresses the propensity to smooth consumption over time. The second term on the right hand side captures the growth effect, while the third term captures the dislike of the agent for the wiggles in the consumption path generated by a stochastic process. The new term is proportional to the measure of intrinsic risk aversion, which is not captured in the standard model, and further reduces the discount rate. Increasing risk aversion (in the Arrow-Pratt as well as in the intrinsic sense) reduces the discount rate. In contrast, increasing η generally increases the discount rate. Disentangled estimates and calibrations in the asset pricing context result commonly in $\eta = 2/3$ and $\text{RRA} \in [8, 10]$ (Vissing-Jørgensen & Attanasio 2003, Basal & Yaron 2004, Basal, Kiku & Yaron 2010). Picking $\text{RRA} = 9$ yields a coefficient of relative intertemporal risk aversion of $\text{RIRA} = 25$ and a discounting effect of intrinsic risk aversion that is $\text{RIRA} |1 - \eta^2| \frac{\sigma^2}{2} / \eta^2 \frac{\sigma^2}{2} \approx 31$ times larger than the effect created by aversion to consumption wiggles. In our numerical example with $\mu = 1.8\%$ and a standard deviation of $\sigma = 3.8\%$ the growth effect reduces to $\eta\mu = 1.2\%$, the standard risk to 0.03% , and the

effect of intrinsic risk aversion equals 0.9%. Then, the social discount rate becomes $\rho+0.3\%$ and reduces almost to pure time preference, which the cited calibrations generally find to be close to zero as well. See Traeger (2008) for a sensitivity analysis and Gollier (2002) for a different treatment of discounting in the case of general risk attitude. Note that equations (14.3) and (14.4) hold not just for certain project payoffs, but also in the case where the project payoff is independent of baseline uncertainty. Traeger (2008) discusses the case of correlation between project payoff and baseline uncertainty.

14.5.2 General uncertainty attitude

In general, decision makers do not know the probability distributions governing the future with certainty. Situations whether the decision maker does not know the underlying probabilities are known as situations of ambiguity, hard uncertainty, or deep uncertainty (as opposed to risk). Klibanoff, Marinacci & Mukerji (2005) and Klibanoff, Marinacci & Mukerji (2009) develop a convenient model of decision making under ambiguity known as the smooth ambiguity model. It is similar to a standard Bayesian model, but distinguishes the attitude with respect to known probabilities (risk) from the attitude with respect to unknown probabilities (ambiguity), which are identified with the Bayesian prior. Traeger (2010*b*) generalizes the model and establishes its normative foundation: acknowledging the existence of different types of uncertainty, risk aversion measures depend on the type of uncertainty a decision maker faces, even within a framework based on the von Neumann & Morgenstern (1944) axioms. The smooth ambiguity model corresponds to the special case where risk attitude coincides with the attitude for consumption smoothing, but differs from the attitude to ambiguity. The measure of ambiguity aversion is similar to that of intertemporal risk aversion; we denote the coefficient of relative ambiguity aversion by RAA. We assume once more isoelastic preferences and normal growth. However, this time the expected value μ^* of the normal distribution is unknown. Given a particular value μ^* , the standard deviation is once more denoted σ . The expected growth rate μ^* is governed by a normal distribution with expected value μ and standard deviation τ . Traeger (2008) calculates the resulting extension of the Ramsey equation as

$$r = \delta + \eta\mu - \eta^2 \frac{\sigma^2 + \tau^2}{2} - \text{RAA} \left| 1 - \eta^2 \right| \frac{\tau^2}{2}. \quad (14.5)$$

The formula resembles equation (14.4) for intertemporal risk aversion. The difference are, first, that in the Bayesian model the overall uncertainty creating consumption wiggles is captured by the sum of the variance of both normal distributions. Second, intrinsic uncertainty aversion only affects second order uncertainty captured by τ . Extending the model to disentangle risk aversion from both ambiguity aversion and the propensity to smooth consumption over time, implies that the Ramsey equation collects both terms, those proportional to intertemporal risk aversion in equation (14.4) and that proportional to ambiguity aversion in equation (14.5) (Traeger 2008). In the case of isoelastic preferences, intrinsic uncertainty aversion in the sense of intertemporal risk aversion and smooth ambiguity aversion always reduces the social discount rate. Gierlinger & Gollier (2008) and Traeger (2011a) establish general conditions under which general uncertainty lowers the social discount rate. The latter paper also shows how a decrease in confidence in the futurity of the growth forecast can lead to a falling term structure.

14.5.3 Catastrophic risk

Weitzman (2009) develops an argument that catastrophes would swamp the importance of discounting. In a Bayesian decision model with isoelastic preferences he assumes that the stochastic process governing growth is unknown. In contrast to the analysis in sections 14.5.1 and 14.5.2, Weitzman does not incorporate general risk or uncertainty attitude. Instead of assuming a normal prior on expected growth, Weitzman puts an uninformative prior on the variance of the growth process. He shows that the resulting overall uncertainty is sufficiently fat tailed to imply an infinite consumption discount factor, implying an infinite weight on future consumption. Weitzman calls this result a Dismal Theorem. A simplified perspective on his result, neglecting the precise model of uncertainty and learning in Weitzman (2009), is that inserting enough uncertainty into the model implies that as $\tau \rightarrow \infty$ in equation (14.5) the discount rate goes to minus infinity. In utility terms, the intuition for Weitzman's result is that his model exhibits a sufficiently slow decline of the probability mass characterizing that future consumption approaches zero and marginal utility infinity.¹⁰⁴ Weitzman makes the point

¹⁰⁴It might be useful to step back from elaborate economic welfare representations. In terms of preferences Weitzman's model contains a zero consumption state and, probabilistically, a lottery that is to be avoided by all means. Weitzman shows that the willingness to

that, even if we bound marginal utility away from minus infinity, the discount factor would be highly sensitive to the precise bound.

The social discount rate here and in Weitzman's calculation gives the value of a certain marginal consumption unit shifted into the future. Weitzman constructed an immensely uncertain future and then calculates the value of handing the future agents the first certain unit. If such a certain transfer mechanism would be available, this transfer should happen. With the first unit transferred infinity goes away and we can calculate the optimal amount that should be transferred into the future. The discount rate is like a price. If we offer an agent dying of thirst in the desert the first sip of water, he would likely give up all his worldly belongings in exchange. However, this measurement would not generally give us the market value of water. If, in contrast, uncertainty is insuperable, then we cannot simply calculate the social discount rate based on a certain consumption transfer, but have to account for uncertainty in the transfer and its correlation to baseline uncertainty (Traeger 2008). The empirical plausibility of the magnitude of uncertainty that Weitzman (2009) assumes is also questionable in the climate context in which it is motivated. See Millner (2011) for a discussion and extension of Weitzman's model.

14.5.4 The Weitzman-Gollier puzzle

Weitzman (1998, 2001) and Gollier (2004) analyze the social discount rate in the presence of uncertainty about future economic productivity. Both authors assume perfectly serially correlated interest rates. Weitzman derives a falling term structure and Gollier derives an increasing term structure from assumptions that are apparently the same. This finding is known as the Weitzman-Gollier puzzle. Two insights help to resolve the puzzle. First, the original papers on the puzzle did not take into consideration the change of marginal utility over time and risk states (Gollier 2009, Gollier & Weitzman 2010, Freeman 2010). Second, Gollier's reasoning is concerned with the uncertain payoff of an investment project, while Weitzman's reasoning gets at growth uncertainty changing baseline consumption in the future. Gollier asks for the following certainty equivalent discount rate: what average

pay to get rid of this state are 'all means'. Note that the expected utility model with isoelastic utility does not satisfy the usual choice axioms when including the zero consumption level.

annual productivity must a certain project have in order to match expected annual productivity of the uncertain project? The term structure of this rate generally increases: the payoffs of the uncertain project grow exponentially over time under full serial correlation, and the highest interest rate scenario dominates the (linear) expected value. In contrast, Weitzman's suggested rate is in the spirit of equation (14.3), which implies a falling term structure under serial correlation.¹⁰⁵ If the payoff uncertainty of the project under evaluation is independent of the market interest, then the value of the uncertain project increases over time with respect to a certain project as Gollier's discount rate implies. Both the certain and the uncertain project increase in value over time in a world of serially correlated uncertainty, relative to a world of certainty, as reflected by Weitzman's decreasing discount rate. In the case where project payoff and market interest are perfectly correlated the effect pointed out by Gollier vanishes. Then, the exponential payoff growth emphasizing the high payoff states is offset by the simultaneous decrease of the marginal utility obtained from an additional consumption unit, because the realization of the project payoff simultaneously determines the total consumption growth in the economy.

¹⁰⁵Weitzman (1998, 2001) argues only by means of productivity in the economy. However, a close examination of his argument shows that the relation between consumption growth and productivity growth makes his formula almost correct Gollier & Weitzman (2010). It is only almost correct because it overlooks that the consumption share generally responds to the resolution of uncertainty over the market interest (Freeman 2010, Traeger 2012).

14.6 Discounting: Extensions

This section surveys some important extensions of the Ramsey formula. We start by relaxing the assumption of an aggregate consumption good, and analyze how limited substitutability between environmental services and produced consumption affects the discount rate. We then discuss the case of hyperbolic discounting as triggered by a non-constant RPTR. Finally, we explain that the explicit treatment of overlapping generations generally leads to a model equivalent to that of non-constant RPTP.

14.6.1 Environmental versus produced consumption

Above we assumed the existence of an aggregate consumption commodity. This assumption becomes crucial if different classes of goods are not perfect substitutes. In particular, produced consumption is likely to be an imperfect substitute for environmental goods and services. Moreover, the provision and consumption of environmental goods and services does not generally grow at the rate of technological progress. Then, as our economy grows, environmental goods and services become relatively more valuable over time. We can incorporate this effect into a cost benefit analysis by introducing a time dependent conversion factor that translates the costs and benefits in terms of environmental good into current value produced consumption units. Alternatively, we can price environmental and produced consumption by present value prices and apply good-specific discount rates for cost benefit analysis. In both approaches, the underlying discount rate is affected by imperfect substitutability.

We assume that a representative agent derives utility $u(c_t, e_t)e^{-\rho t}$ from consuming produced goods c_t and environmental consumption and services e_t . We define the discount factor of the consumption good as above and the discount factor for the environmental good as the amount of the environmental good that an agent is willing to give up in the present in order to receive an additional unit of the environmental good in the future. This rate is known as the ecological discount rate.¹⁰⁶ The discount rate characterizing the rate

¹⁰⁶This rate indicates how the value of a unit of the environmental good changes over time. If we are concerned how much of a consumption unit in the present an agent should give up for a future unit of environmental services, then we simply have to multiply the corresponding ecological discount factor with the relative price of the two goods in the

of change of the discount factor for consumption becomes

$$\delta_c(t) = \rho + \eta_{cc}(t) \hat{c}(t) + \eta_{ce}(t) \hat{e}(t) \quad (14.6)$$

with $\hat{c}(t) = \dot{c}_t/c_t$, $\eta_{cc}(t) = \eta_{cc}(c_t, e_t) = -\frac{\partial^2 u}{\partial c^2} c / \frac{\partial u}{\partial c}$ and $\eta_{ce} = \eta_{ce}(c_t, e_t) = -\frac{\partial^2 u}{\partial c \partial e} e / \frac{\partial u}{\partial c}$. Unless both goods are perfect substitutes ($\eta_{ce} = 0$), the consumption discount rate for produced consumption depends both, the growth of produced consumption and on environmental growth (or decline).

Assuming Cobb-Douglas utility $u(c_t, e_t) = c_t^{a_c} e_t^{a_e}$ (where $a_c + a_e = 1$) eliminates the overall growth effect because Cobb-Douglas utility is linear homogeneous. We use this function form to focus on the effect of growth differences between produced and environmental consumption. Then, the consumption discount rate for the produced good (14.6) simplifies to

$$\delta_c(t) = \rho + a_e(\hat{c}_t - \hat{e}_t) .$$

Relatively faster growth in produced consumption increases the produced consumption discount rate. Similarly, this faster growth of produced consumption reduces the discount rate for the environmental goods and services:

$$\delta_e(t) = \rho - a_c(\hat{c}_t - \hat{e}_t) .$$

Thus, if produced consumption grows more rapidly than consumption of environmental goods, the discount rate to be applied in a cost benefit analysis for environmental good preservation is lower than the discount rate for produced consumption. This adjustment of the social discount rate for the environmental good reflects an increase in the relative scarcity of the environmental good causing its (relative) price to increase. For constant growth rates, both social discount rates are constant. However, this is a consequence of the unit elasticity of substitution between environmental and produced consumption. In general, these good-specific discount rates change over time. Both the discount rate for produced consumption as well as the discount rate for environmental goods and services can fall over time as a consequence of limited substitutability (Hoel & Sterner 2007, Traeger 2011b).

present.

14.6.2 Hyperbolic discounting

Many models of dynamic public policy involve non-constant social discount rates. The nature of the resulting policy problem depends on whether this non-constancy causes time inconsistency. Time inconsistent policies can imply an ongoing revision of the formerly optimal policy, even in the absence of new information. In contrast, a declining term structure caused by falling growth rates, serially correlated uncertainty, or limited between-good substitutability leads to time consistent plans. Here we analyze the most common model giving rise to non-constant discount rates that cause time inconsistent plans: models employing a non-constant RPTP.

Ramsey (1928) noted “. . . My picture of the world is drawn in perspective. . . I apply my perspective not merely to space but also to time.” The obvious meaning of “perspective applied to time” is that events in the more distant future carry less weight today, just as objects in the distance appear smaller. Any positive discount rate, including constant discounting, creates this type of perspective applied to time. However, perspective means more than the apparent shrinking of distant objects. The simplest model of perspective applied to space, known as “one point perspective”, can be visualized as the appearance of railroad tracks viewed straight on, so that the two rails appear to converge at the horizon. The distance between adjacent railroad ties appears to grow smaller the more distant are the ties, but the rate of change appears to fall (Karp 2009). This kind of perspective means that not only do distant objects appear smaller, but also that we are less able to distinguish between the relative size of two objects, the further they are from us. Hyperbolic discounting, which assumes that the discount rate falls over time, is the time analog of this spatial perspective.

Hyperbolic discounting arises in both behavioral models of individual decision problems (Laibson 1997) and in long-lived environmental problems (Cropper, Ayded & Portney 1994). In the former setting, individuals’ tendency to procrastinate is a prominent rationale for hyperbolic discounting. The rationale in the environmental setting is more closely tied to the fact that the problem of interest (e.g. climate change) occurs on a multi-generation scale. If we care less about our grandchildren than we do about our children, and care still less about generations that are more distant from us, our preferences are consistent with a positive discount rate on the generational time scale. If, in addition, we make less of a distinction between two contiguous

generations in the distant future compared to two generations close to us, our pure rate of time preference is hyperbolic. We might have a preference for our children relative to our grandchildren but scarcely distinguish between those born in the 20'th and the 21'st generation from ours. If individuals have this kind of time perspective and if the social planner aggregates the preferences of agents currently alive, then the social planner has hyperbolic discounting.

Non-constant discounting arising from preferences, as described above, causes optimal programs to be time inconsistent. That is, at any point in time the current social planner would like to deviate from the plan that was optimal for an earlier social planner. The time inconsistency is easiest to see using a discrete time example of the " β, δ model", where the sequence of discount factors used at t to weigh payoffs at times $\tau \geq t$ is $1, \beta, \beta\delta, \beta\delta^2, \beta\delta^3 \dots$. If $\beta = \delta$ the discount factor is constant, and discounting is exponential. If $\beta < \delta$, discounting is "quasi-hyperbolic". Consider a project that reduces time $t+1$ utility by $\frac{\beta+\delta}{2}$ units and increases $t+2$ utility by 1 unit, and suppose $\beta < \delta$. A planner at time t would accept this project, because the present value of the utility loss, $\beta\frac{\beta+\delta}{2}$, is less than the present value of the utility gain, $\beta\delta$. However, the planner at time $t+1$ rejects the project, because for that person the present value of the utility loss is $\frac{\beta+\delta}{2}$, which is greater than the present value of the utility gain, β . The case $\beta < \delta$ is associated with procrastination: a tradeoff that looks attractive when viewed from a distance becomes less attractive when viewed from close up. If a unit of time is taken to be the span of a generation, quasi-hyperbolic discounting implies that we are willing to make smaller sacrifices for our children than we would like them (and all subsequent generations) to make for their children.

One resolution to the time-inconsistency problem assumes that the initial planner chooses the current action under the belief that her entire sequence of preferred actions will be carried out. This resolution is dubious in a multi-generation context, where a current decision maker is unlikely to believe that she can set policy for future generations. A second resolution is to treat the policy problem as a sequential game amongst policymakers (Harris & Laibson 2001, Karp 2007). The optimal action for a social planner at time t depends on her belief about how policymakers will behave in the future. In a Markov perfect equilibrium, actions, and therefore beliefs about future actions, are conditioned only on directly payoff-relevant state variables. Often those variables have a physical interpretation, e.g. an environmental stock.

14.6.3 Overlapping generations

A closely related explanation for non-constant discounting rests on a model of overlapping generations. Suppose that agents discount their own future flow of utility at a constant pure rate of time preference, ρ , and that in addition they discount the welfare of the not-yet born at rate λ . Agents with both “paternalistic” and “pure” altruism care about the utility flows of future generations; for these agents, $\lambda < \infty$, and for agents without altruism, $\lambda = \infty$. Agents with pure altruism – unlike agents with paternalistic altruism – consider the effect on intermediate generations of the welfare of more distant generations (Ray 1987, Andreoni 1989).

If agents’ lifetimes are exponentially distributed, with no aggregate uncertainty, all agents currently alive have the same expected lifetime (Yaari 1965, Blanchard 1985). Absent other considerations (e.g. different levels of wealth, because older agents have had more opportunity to accumulate) agents currently alive are identical, so there is a representative agent in the usual sense. If instead, agents’ have random lifetimes with finite support (Calvo & Obstfeld 1988) or finite deterministic lifetimes (Schneider et al. 2012), the older agents have shorter remaining (expected) lifetimes. In this case, a social planner, perhaps a utilitarian, aggregates the preferences of agents alive at a point in time.

For the case of exponentially distributed lifetimes and paternalistic altruism, the discount factor of the representative agent is the weighted sum of two exponentials (Ekeland & Lazrak 2010). (Models with pure and paternalistic altruism are observationally equivalent.) The associated discount rate is non-constant, except for the two limiting cases, $\lambda = \rho$ or $\lambda = \infty$; in the first limiting case, the social discount rate is constant at ρ and in the second it is constant at $\rho +$ the death rate. If $\infty > \lambda > \rho$ the social discount rate increases over time asymptotically to $\rho +$ the death rate. If $\lambda < \rho$ the social discount rate decreases over time asymptotically to λ . For both $\lambda < \rho$ and $\rho < \lambda < \infty$ agents have a degree of altruism and non-constant discounting, but only $\lambda < \rho$ corresponds to hyperbolic discounting. We noted above that many have argued that the only ethical choice is to treat all generations symmetrically, regardless of their date of birth. In this context, that requires $\lambda = 0$, so that the social planner’s evaluation of the stream of an agent’s utility does not depend on when she was born.

The previous section explains why a time inconsistency problem arises when

discount rates are non-constant. As noted above, one resolution is to consider a Markov perfect equilibrium in the game amongst generations. A second resolution eliminates the time inconsistency problem by assuming that the social planner at any point in time discounts the utility of those currently alive back to their time of birth, rather than to the current time.

Table 1 Parameters of the model

$\beta = 0.985$	utility discount factor (rate of pure time preference $\delta_u = 0.015$)
$\rho = -1$	characterizes elasticity of intertemporal substitution as $\frac{1}{1-\rho}$
$\delta_K = 0.1$	depreciation rate of capital (in code it is persistence of capital 0.9)
$\delta_M = 0.01$	decay rate of CO_2
$\sigma = 0.1$	emission intensity of GDP, i.e. (CO_2 in GtC/GDP in trillion USD, called 'emint' in code)
$AL = 37.842$	effective labor in million, composed of <ul style="list-style-type: none"> · $A = .0058$ (level of labor productivity) · $L = 6514$ in million, (Population=Labor)
$\kappa = .3$	capital elasticity in production
$B = 1.1$	in GtC/year, CO_2 from land use change and forestry (B1 in code)
$EF = 0$	External forcing
$\Psi = 0.0561$	abatement cost, multiplicative constant
$a_2 = 2.8$	abatement cost, exponent
$b_1 = 0$	damage, intercept
$b_2 = 0.00284$	damage, quadratic constant (% of GDP lost at 1°C above preindustrial)
$b_3 = 2$	damage, quadratic exponent
$s = 3.08$	Climate sensitivity, i.e. equilibrium response to doubling of atmospheric CO_2 concentration with respect to preindustrial concentrations (value differs from DICE)
$M_{preind} = 596$	in GtC, preindustrial stock of CO_2 in the atmosphere
$K(0) = 137$	in trillion 2005-USD, initial value for global capital stock
$M(0) = 808.9$	in GtC, Initial stock of atmospheric CO_2

15 Time Inconsistency

Many economic problems involve situations where different types of agents, all of whom solve dynamic problems and who have rational expectations, differ in their strategic power. For example, a policymaker such as the government recognizes that its decisions affect the behavior of a group of agents who behave non-strategically. To discuss the generic case, we refer to the strategic agent as the leader, and the non-strategic agent(s) as the follower(s). This terminology emphasizes the parallel between the dynamic models that we consider here and static models with a Stackelberg equilibrium. In the Stackelberg equilibrium, the leader moves first, and chooses his action with a view to affecting the subsequent action of the follower.

Exogenous uncertainty is peripheral to the issues we are interested in here, so for simplicity we assume that there is no exogenous uncertainty. The statement that agents have rational expectations means that they have rational “point” expectations. That is, they are able to understand how all agents behave and therefore predict the outcome of the variables that are of interest to them. In the absence of exogenous uncertainty, the rational expectations assumption means the predictions are completely accurate. There are no surprises in this setting.

We are interested in situations where the decision trajectory that is optimal for the leader at an initial time, say time 0, is not optimal at a subsequent time. In this case, the leader has an incentive, at time $t > 0$, to change the trajectory of decisions that she announced at an earlier time. The policy that the leader announced at time $t = 0$ is said to be time-inconsistent. Time inconsistency typically arises when the leader has a second best policy, i.e. one that generates a “secondary distortion”, as defined below.

In the absence of an ability on the part of the leader to commit to the policy trajectory she announced at time 0, rational agents would not believe that this policy trajectory would actually be implemented. The time inconsistent equilibrium is therefore generally considered implausible.

This chapter uses examples to illustrate circumstances where the optimal policy (from the standpoint of the leader) is time inconsistent, and it shows how to determine an equilibrium that is time inconsistent. We discuss the distinction between time consistency and subgame perfection or (Markov perfection), and we explain a source of multiplicity of Markov perfect equilibria

(MPE).

15.1 An optimal tax example

This example illustrates why time inconsistency typically arises when the leader is restricted to use a second best policy, i.e. one that generates a secondary distortion. The example here is based on a simple dynamic general equilibrium problem in which all private agents (consumers, firms, workers, etc.) are price takers. Individuals allocate their time between work and leisure, a decision that depends on the endogenous wage. Income consists of labor income and the return on investment; individuals allocate their income between consumption and investment. Firms rent capital and hire workers. There are no distortions, so a competitive equilibrium without taxes is efficient. The government wants to maximize the present discounted welfare of the representative agent, so it chooses zero taxes.

There is an obvious asymmetry between the government and private agents. The latter understand how prices evolve, but they take prices as given. All of these agents are small relative to the aggregate economy, so no private agent is able to influence aggregate outcomes. In contrast, the government is large enough to influence the economy, by means of its taxes. Here it is clear why we refer to the government as the leader, and all other agents as followers.

Now alter the model so that the government has an exogenously given budget constraint: the government needs to raise a certain amount of revenue in each period. If there were taxes that leave unchanged private decisions regarding labor supply and demand and investment, then those taxes are first best. However, many taxes change individuals' decisions; we refer to such a change as a "secondary distortion". For example, a tax on capital changes investment decisions and a tax on labor changes the supply of labor.

For concreteness, suppose that the government has at its disposal only taxes on the returns to capital (profits) and on the return to labor (wages). The theory of optimal taxation tells us that the level of an optimal tax is inversely related to the elasticity of supply of the factor being taxed. For example, if the supply of labor is fixed, i.e. the elasticity of supply of labor is zero, then a tax on wage income changes the after-tax wage without altering the amount of labor in the market; in this limiting case, the labor tax raises revenue

without creating a “distortion”. In the model described above, workers have an alternative use of their time (leisure), so the elasticity of supply of labor is positive, a fact that tends to decrease the optimal tax on wages. In contrast, at a point in time, the stock of capital is fixed, i.e. its supply is perfectly inelastic. Therefore, in a static setting it is optimal to raise all of the needed revenue by taxing capital, not labor.

The government in our example faces a dynamic rather than a static problem. Although the current tax does not affect the current supply of capital, the anticipation of future capital taxes does affect the incentive to invest, thereby altering future supply of capital. There is a difference between the long run and the short run elasticity of supply of capital. Consider the situation of the government at time $t = 0$ deciding on the tax portfolio that will be implemented at $t > 0$. The capital tax at t , τ_t , serves two functions. Most obviously, it raises revenue at time t ; this is the reason for the tax. Conditional on the stock of capital at time t , the incentives described in the previous paragraph militate in favor of a relatively large capital tax. However, τ_t also influences investment decisions over $[0, t)$, thereby affecting the stock of capital at time t . A larger tax discourages investment and leads to a lower time- t capital stock. This fact militates in favor of a smaller value of τ_t .

The optimal value of τ_t , from the standpoint of the government at time 0, must balance the offsetting incentives. The time inconsistency arises because the balance of incentives changes over time. For example, at time $t - \varepsilon > 0$, everything that has happened in the past, including past investment, is taken as given; conditional on past events, τ_t can affect investment only over the interval $[t - \varepsilon, t)$. As ε becomes small, the incentive to use a low τ_t in order to encourage investment diminishes, but the incentive to use a high τ_t to raise revenue at time t is unchanged. The fact that the incentives change over time means that the optimal value of τ_t changes as we get closer to time t . Therefore the value of τ_t that was optimal at time $t = 0$ ceases to be optimal at $t > 0$.

The policy trajectory that is optimal at time $t = 0$ is often called the optimal open loop policy, because it changes only as a function of time. Judd (cite) shows that in an economy of the type described above, the optimal open loop capital tax starts at a high level, and the optimal open loop labor tax begins at zero. However, in a steady state, the optimal open loop capital tax converges to 0, while the optimal labor tax rises over time. Of course,

this open loop trajectory is (in most circumstances) time inconsistent.

15.2 Two definitions

The terms “time consistent” and “subgame perfect” are sometimes used interchangeably. The difference between the two is important in games where agents are symmetric, in the sense that no agent (or group of agents) is identified as the leader. The difference is less important in the current context, where agents are asymmetric, simply because the time consistent equilibria that interest us are (in general) also subgame perfect. However, it is worth explaining what the two terms mean, and in particular the manner in which their meanings differ. In order to keep the discussion grounded, we frequently use the tax example of the previous subsection.

The state variable is the collection of objects that determine the evolution of the economy. The value of a state variable at a point in time is predetermined, but in many cases it is endogenous to the model. The modeler has some latitude in defining the state variable, but usually the context of the model suggests a natural choice. A refinement includes in the set of state variables only those are “directly payoff relevant”, i.e. those objects whose value *directly* affects current and future payoffs.

In the tax example above, we could define the state variable to be the current stock of capital and all past values of taxes. This definition is sensible if we want to build a model in which agents base their beliefs about future taxes on past taxes, and if in addition the government actually does condition the future taxes on past taxes. However, conditional on the current stock of capital, the past taxes have no *direct* influence on payoffs. Why should the future evolution of the economy depend on whether it reached the current stock of capital by one trajectory or another? One might reply that past taxes have an *indirect* influence on payoffs, to the extent that they affect agents’ beliefs about what taxes will be used in the future. In this example, the only directly payoff relevant state variable is the stock of capital.

A feasible (for the leader) open loop trajectory is a set of time-indexed capital stock, capital and labor tax, investment and labor decisions. Denote a feasible open loop trajectory as the vector $X_{t,0}$. The first subscript t is the time at which the variables of interest (capital, taxes, ...) are evaluated, and the second subscript 0 denotes the time at which all policies are

chosen. Feasibility means that the trajectory is consistent with the physical constraints, e.g. technology, and with the behavioral constraints arising from private agents' optimizing behavior. An open loop equilibrium, $X_{t,0}^{ol}$ is a feasible open loop trajectory that maximizes the value of the leader's payoff functional, evaluated at the initial time.

Now suppose that we reconsider the problem at time $t > 0$, assuming that all agents have behaved over $(0, t)$ in the manner prescribed by the equilibrium, so that the state variable at t , $X_{t,0}$, is "on the equilibrium path". This statement means that the state variable at t equals the value predicted for it at time 0. Allow the leader to re-optimize at time t . If the ensuing optimal path equals (at every point in time) the continuation of the path that was announced at time 0, i.e. if for all $s \geq t$ and for all $t > 0$, $X_{s,t}^{ol} = X_{s,0}^{ol}$, then the open loop trajectory announced at time 0 is time consistent. In the tax example, we saw that the open loop trajectory is (typically) not time consistent.

Subgame perfection implies time consistency, but time consistency does not imply subgame perfection. Time consistency requires that if the leader at time t is allowed to re-optimize *on the equilibrium trajectory*, she does not want to change the policies announced at time 0. Subgame perfection requires that if the leader at time t is allowed to re-optimize given *any feasible value of the state variable* (not only the value that was predicted at time 0 to arise at time t), she does not want to change the policies announced at time 0. Thus, subgame perfection subjects the candidate path to a stricter test, compared to time consistency.

If the state variable contains all directly payoff relevant variables, and only payoff relevant variables, a subgame perfect equilibrium is said to be Markov perfect. In the tax example, a Markov perfect equilibrium is a tax rule that is conditioned on the (only) directly payoff relevant state variable, the stock of capital; this tax rule must be optimal from the standpoint of the government at all times t and for all feasible levels of capital, given that the government knows that its successors will use this tax rule.

15.3 The durable goods monopoly

This section illustrates the time inconsistency of optimal policy, and the construction of Markov perfect (and therefore, time consistent) policies. Let

Q_t denote the stock of a durable good at time t , e.g. a machine. The (implicit) rental rate at time t of a unit of good, for an interval of time $\varepsilon > 0$ equals $F(Q_t)\varepsilon$. (We refer to this as an implicit rental rate, because in this model firms buy rather than renting the durable good. However, $F(Q)$ equals the amount that would make them indifferent between buying and renting, on an equilibrium price trajectory.) We take ε sufficiently small so that we can ignore discounting within a period, thus simplifying the exposition. Think of $F(Q)$ as a downward sloping demand for rental services. For example, Q is the stock of machines, and $F(Q)$ is the value of marginal product of the Q 'th machine. The addition of one machine, for ε units of time, increases profits by $F(Q)\varepsilon$, a magnitude that depends on Q , because the marginal productivity decreases with the number of machines.

At time t an agent can buy or sell a machine for P_t . That is, there is a second hand market for the durable good, and in this setting an old and a new unit of the durable good are identical and would sell for the same price. The existence of the second hand market means that the monopoly creates its own future competition by selling the durable good in the current period. Another version of this model treats buyers as a continuum of agents, each having a different reservation value for the durable good. The two models are equivalent, only the interpretation differs slightly.

The no-arbitrage condition for agents buying the machine is

$$P_t = F(Q_t)\varepsilon + e^{-r\varepsilon}P_{t+\varepsilon}. \quad (15.1)$$

This condition says that the amount an agent is willing to pay for a unit of the durable good equals the implicit rental rate during a period plus the present value of the next period resale price. A monopoly that produces machines at rate q incurs production costs at the rate $cq + \frac{\gamma}{2}q^2$, $\gamma \geq 0$. We assume that for Q sufficiently large, $F(Q) \leq c$, in order to avoid a technical complication. As above, we ignore discounting within a period. Therefore, if the monopoly produces and sells $q\varepsilon$ machines at the beginning of the period, its single period profits are

$$\left(P_t q - cq - \frac{\gamma}{2}q^2\right)\varepsilon.$$

The monopoly is able to sell, but not to rent machines. One explanation for this assumption is that renting gives rise to a moral hazard problem; renters do not take proper care of the machine, making renting unattractive.

To help fix ideas, consider a two period problem. In this case, agents buy the machine in the first period recognizing that they will use it in both periods; agents who buy in the second period understand that they obtain only one period's use of the machine. The initial stock of machines that agents own is Q_0 . If the monopoly can commit, at the beginning of the first period, to sales in both periods, its problem is

$$\begin{aligned} \max_{q_1, q_2} & \left([F(Q_0 + q_1\varepsilon) + e^{-r\varepsilon} F(Q_0 + q_1\varepsilon + q_2\varepsilon)] q_1 \right) \varepsilon \\ & + [e^{-r\varepsilon} F(Q_0 + q_1\varepsilon + q_2\varepsilon) q_2] \varepsilon \\ & - (cq_1 + ce^{-r\varepsilon} q_2 + \frac{\gamma}{2} [q_1^2 + e^{-r\varepsilon} q_2^2]) \varepsilon \end{aligned} \quad (15.2)$$

The first line of the maximand in expression 15.2 is the value of revenue from sales in the first period, and the second line is the present value of revenue in the second period. The third line equals the present value of costs. In writing the revenue terms we used the fact that equilibrium sales price in the second period equals the rental price (multiplied by ε), and the equilibrium sales price in the first period equals the discounted sum of rental prices in the two periods.

The parameter γ is important for subsequent results, because it makes production costs convex in the rate of production and therefore creates an incentive for the monopoly to shift some production from period 1 to period 2. In the interest of simplicity we temporarily set $\gamma = 0$. With this simplification we rewrite the optimization problem as

$$\max_{q_1, q_2} \left([F(Q_0 + q_1\varepsilon)q_1 - cq_1 + e^{-r\varepsilon} (F(Q_0 + q_1\varepsilon + q_2\varepsilon) (q_1 + q_2) - cq_2)] \right) \varepsilon.$$

The sales rates must be non-negative. At an interior equilibrium the first order conditions for q_1 and q_2 are, respectively,

$$\begin{aligned} q_1 > 0 & \implies F(1) + F'(1)q_1\varepsilon - c + e^{-r\varepsilon} [F(2) + F'(2)(q_1 + q_2)\varepsilon] = 0 \\ q_2 > 0 & \implies e^{-r\varepsilon} [F(2) + F'(2)(q_1 + q_2)\varepsilon - c] = 0, \end{aligned} \quad (15.3)$$

where the notation $F(i)$ and $F'(i)$ for $i = 1, 2$ indicates the function F and its derivative evaluated at the first and second period optimal stock levels. Note that if it is optimal to set $q_i = 0$ then the left side of the corresponding first order condition must be non-positive. Given concavity, there is a unique

local and global maximum. The optimal solution is to set $q_2 = 0$ and to choose $q_1 > 0$ to solve

$$(1 + e^{-r\varepsilon}) (F(1) + F'(1)q_1\varepsilon) - c = 0. \quad (15.4)$$

satisfies the non-negativity constraints and first first order conditions, and . Equation 15.4 is the standard condition for maximizing profits: marginal revenue equals marginal cost. Here marginal revenue is taken with respect to the revenue from sales, where the price equals the discounted sum of implicit rental rates. To confirm optimality, note that for $q_2 = 0$, we have $F(1) = F(2)$ and $F'(1) = F'(2)$; now the left side of equation 15.3 is

$$e^{-r\varepsilon} [F(2) + F'(2)(q_1 + q_2)\varepsilon - c] < 0,$$

where the inequality follows from equation 15.4. Thus, at $q_2 = 0$, a marginal increase in q_2 strictly decreases profits. Therefore it is optimal to set $q_2 = 0$.

The conclusion is that when production costs are linear, the optimal policy for the monopoly that can commit is to sell in the first period the amount that equates marginal revenue and marginal cost, and to sell nothing subsequently. The explanation is that with linear production costs, the monopoly has nothing to gain by delaying sales; doing so merely delays its receipt of profits, reducing the present value of profits. It is easy to see that this conclusion holds regardless of the number of remaining periods. With linear production costs, the monopoly that can commit wants to sell the amount that maximizes revenue from first period sales, $[(1 + e^{-r\varepsilon}) F(Q_0 + q\varepsilon) - c] q$, and nothing thereafter.

In the absence of the ability to commit in period 1 to period 2 sales level, the open loop equilibrium described above is not time consistent. In period 2, conditional on period 1 sales, the monopoly's optimization problem is

$$\max_{q_2} ([F(Q_0 + q_1\varepsilon + q_2\varepsilon) - c] q_2) \varepsilon$$

with first order condition (assuming an interior solution)

$$F(2) + F'(2)q_2\varepsilon - c = 0, \quad (15.5)$$

which is the standard condition from the static problem: set marginal revenue equal to marginal cost. The first order condition 15.3, under commitment,

and 15.5, without commitment, are different. Under commitment, in choosing second period sales the monopoly takes into account the effect of those sales on period 1 revenue. In the absence of commitment, the monopoly in the second period take first period profits as given, and therefore ignores the effect of second period sales on first period profits.

The solution to the first order condition 15.5 gives a second period equilibrium decision rule, which we denote $q_2 = \kappa(Q_1)$, for any $Q_1 = Q_0 + q_1\varepsilon$. The equilibrium second period price is therefore

$$P_2(Q_1 + \kappa(Q_1)) \equiv F(Q_1 + \kappa(Q_1))\varepsilon$$

and the equilibrium second period profits are

$$\pi_2(Q_1) \equiv [F(Q_1 + \kappa(Q_1))\varepsilon - c] \kappa(Q_1)\varepsilon.$$

Using this function and the no arbitrage condition 15.1 we can write the first period optimization problem for the monopoly who cannot commit as

$$\pi_1(Q_0) = \max_{q_1} (P_1(Q_1) - c) q_1\varepsilon + e^{-r\varepsilon} \pi_2(Q_1)$$

subject to $Q_1 = Q_0 + q_1\varepsilon$ and $P_1(Q_1) = F(Q_1)\varepsilon + e^{-r\varepsilon} P_2(Q_1 + \kappa(Q_1))\varepsilon$.

The first constraint follows from the definition of the stock. The second constraint uses the no arbitrage condition 15.1, and is a consequence of buyers' optimizing behavior.

We emphasize the relation between this problem and the kind of dynamic programming problems that we have seen earlier. The similarity is that in both this problem and the more familiar problem, we have to find the value function in order to solve the problem in the current period. The difference is that here we have one more endogenous function, the equilibrium price function. We obtain both of the endogenous functions, π_2 and P_2 by solving the problem backwards.

We now consider the limit of this problem as the number of remaining periods goes to infinity. We assume that this limit exists and seek a stationary solution, i.e. a decision rule, value function, and price function that depend on the current state variable, but not on calendar time. To this end, we merely remove the time subscripts in the previous two-period problem, to

write the dynamic programming equation for the infinite horizon problem as

$$\pi(Q) = \max_q (P(Q') - c)q\varepsilon + e^{-r\varepsilon}\pi(Q') \quad (15.6)$$

$$\text{subject to } Q' = Q + q\varepsilon \text{ and } P(Q') = F(Q')\varepsilon + e^{-r\varepsilon}P(Q' + \kappa(Q')\varepsilon). \quad (15.7)$$

The variable Q is the stock at the beginning of the period, and Q' is the stock after sales in the current period, $q\varepsilon$. The first order condition to this problem, for all $\varepsilon > 0$, is

$$\left(\left[\frac{dP(Q')}{dQ}q + e^{-r\varepsilon}\frac{d\pi(Q')}{dQ} \right] \varepsilon + P(Q') - c \right) = 0. \quad (15.8)$$

The term in square brackets is multiplied by ε . For small ε , the first order condition therefore is approximately the condition that price equals marginal cost, which is also the condition for a competitive firm. The Coase Conjecture states that as ε approaches 0, the behavior of the durable goods monopoly that is not able to commit to a sales trajectory approaches the behavior of a competitive industry. Coase described this result by saying that monopoly power “vanishes in the twinkling of an eye”.

We confirm this conjecture by examining the limiting equilibrium, as $\varepsilon \rightarrow 0$. The functions P and π depend on ε . We assume that these functions change smoothly with ε , so that the following Taylor expansion is valid. Using a Taylor approximation of the dynamic programming equation and the constraints, we write the limiting form of equations 15.6 and 15.7 as

$$r\pi(Q) = \max_q \left[P(Q) - c + \frac{d\pi(Q)}{dQ} \right] q \quad (15.9)$$

$$\text{subject to } \frac{dQ}{dt} = q \text{ and } \frac{dP}{dt} = rP(Q) - F(Q).$$

The equation of motion $\frac{dQ}{dt} = q$ is data, i.e. it is given by the statement of the problem. The constraint $\dot{P} = rP(Q) - F(Q)$, in contrast, is given by the optimizing behavior of private agents. The function $P(Q)$ is *not* data; it is endogenous, and must be determined in order to solve the maximization problem; in much the same way as the function $\pi(Q)$ must be determined. To distinguish the two types of constraints we refer to the second, $\dot{P} =$

$rP(Q) - F(Q)$, as a “side condition”. Although $P(Q)$ is endogenous to the model, the monopoly takes it as given. The function equals

$$P(Q_t) = \int_t^\infty e^{-r(s-t)} F(Q_s) ds,$$

i.e. it is actually a functional of $\{Q_s\}_{s=t}^\infty$. However, in a Markov perfect equilibrium, $q_t = \kappa(Q_t)$, for some function $\kappa(\cdot)$. Substituting this function in the equation of motion $\dot{Q} = q$, we can write the integral as a function of the current value of Q . Because the monopoly at time t takes the decision rule of future incarnations of itself as given, it takes the function $P(Q)$ as given. The function $P(Q)$ must be such that when the monopoly solves the DPE 15.9 subject to $\dot{Q} = q$, taking $P(Q)$ as given, the induced function $P(Q)$ solves the side condition $\dot{P} = rP(Q) - F(Q)$.

The maximand in the DPE 15.9 is linear in the control. The problem is analogous to that of a competitive firm that faces a constant price and has constant marginal production costs. The functions π and P are determined by the monopoly’s future behavior. Given the inability to commit, the monopoly at a point in time is not able to choose this behavior. At most, the monopoly can influence this behavior by affecting the level of the durable goods stock that it bequeaths to its successors (its future reincarnations). In the case at hand (but not more generally), that limited influence is useless, because agents – here, buyers – anticipate that future sales will be either $q = 0$ or $q = \infty$. In the latter case, the stock jumps to its steady state value. The steady state value is the level at which $\frac{dP}{dt} = 0$, i.e. where $P(Q) = \frac{F(Q)}{r}$. The steady state stock equals the level that satisfies $\frac{F(Q)}{r} = c$, which is the competitive level. In the limit as $\varepsilon \rightarrow 0$, the durable goods monopoly has no market power. Its “future selves” compete away all potential market power.

One further detail is worth mentioning. The previous paragraph shows that one equilibrium involves an instantaneous jump in the stock to the competitive level. Are there other equilibria? To see why the answer is “no”, note that if it is optimal to set q to a positive finite level, it must be the case that $P(Q) - c + \frac{d\pi(Q)}{dQ} = 0$. In this case, equation 15.9 implies that the value of the monopoly’s program is $r\pi(Q) = 0$, i.e. the monopoly makes 0 profits on any equilibrium trajectory.

15.3.1 An extension to convex costs

Kahn (198?) shows that when production costs are convex, i.e. for $\gamma > 0$, the durable goods monopoly retains some market power, i.e. earns rent even in the limit as $\varepsilon \rightarrow 0$. Proceeding formally, we can replace the dynamic programming equation 15.9 with

$$r\pi(Q) = \max_q \left[P(Q) - c - \frac{\gamma}{2}q + \frac{d\pi(Q)}{dQ} \right] q$$

subject to the same constraints as above. The convexity of costs means that setting $q = \infty$ is never an optimal decision. The positive value of γ does not change the steady state, because in the steady state sales equal zero.

The first order condition outside a steady state, i.e. when $q > 0$, is

$$P - (c + \gamma q) = -\frac{d\pi(Q)}{dQ} > 0, \quad (15.10)$$

where the inequality follows from the fact that a higher existing stock of the durable good reduces future sales, reducing the present value of the stream of future profits. The competitive firms sets price equal to marginal cost. Outside the steady state, the monopoly sets price strictly above marginal costs, and earns profits even though it is unable to commit. The monopoly slows sales relative to the competitive level, in order to increase the stream of profits. Monopoly power vanishes, but only asymptotically, not in the twinkling of an eye.

Kahn uses a linear rental function, $F = a - bQ$ and shows that the equilibrium price function is also linear, with $P = A - BQ$, and the value function (π) is quadratic. The parameters of these functions depend on the parameters of the problem, a, b, c, γ and r . (See the problem set.)

15.3.2 An extension to depreciation

Karp (1995) shows that if the durable good depreciates, there is a family of Markov perfect equilibria. Only one of these results in zero profits. In all other equilibria the monopoly earns profits, so “depreciation erodes the Coase Conjecture”. The multiplicity of equilibria arises for the same reason as in many dynamic strategic settings: the lack of a “natural boundary condition”.

In the case above, where the durable good does not depreciate, it must be the case that eventually the stock of the good converges to the competitive level, where $\frac{F(Q)}{r} = c$. It cannot be the case that in a Markov perfect equilibrium the monopoly stops producing when there are still opportunities for profit. In contrast, when the good depreciates, in a steady state the monopoly continues to produce enough to maintain the steady state. Here there is nothing in the description of the problem that tells us the level to which the stock must converge. There is not a natural boundary condition.

Setting $\gamma = 0$ and now including a positive depreciation rate, $\delta > 0$, we obtain the DPE and the constraints

$$r\pi(Q) = \max_q \left[P(Q) - c + \frac{d\pi(Q)}{dQ} \right] q - \frac{d\pi(Q)}{dQ} \delta Q \quad (15.11)$$

subject to $\frac{dQ}{dt} = q - \delta Q$ and $\frac{dP}{dt} = (r + \delta)P(Q) - F(Q)$.

As was the case before, the maximand is linear in the control; an interior equilibrium requires that the term that multiplies q equals 0:

$$P(Q) = c - \frac{d\pi(Q)}{dQ}. \quad (15.12)$$

This equilibrium condition implies, using the DPE 15.11,

$$r\pi(Q) = -\frac{d\pi(Q)}{dQ} \delta Q.$$

The solution to this differential equation is

$$\pi(Q) = kQ^{-r/\delta}, \quad (15.13)$$

where k is a constant of integration.

For $k = 0$ we obtain the solution in which the monopoly behaves exactly like a competitive firm, and earns zero profits. For positive k we can differentiate the expression for π in equation 15.13 and substitute the result into equation 15.12 to obtain the price function, indexed by k

$$P^k(Q) = c + \frac{rk}{\delta} Q^{-(r+\delta)/\delta}.$$

The superscript k on the price function is an index, indicating that the function depends on the still undetermined constant k . Further analysis

shows that there is an interval of k ranging from 0 to a positive level, which depends on the parameters of the problem. For values of k in this interval, the price function $P^k(Q)$ is an equilibrium price function, except for a set of initial levels of Q close to 0. For initial values of Q close to 0, the monopoly produces at an infinite rate, causing Q to jump to a larger level, after which P^k is the equilibrium price function, and sales are positive and finite. The stock converges asymptotically to a steady state, a monotonic function of k . Thus, we can characterize a particular equilibrium using the index k , or equivalently by using the corresponding steady state.

The conclusion is that the lack of a natural boundary condition leads to an infinite multiplicity of equilibria. To each of these equilibria there is a different steady state. All equilibria except for one, indexed by $k = 0$, result in positive profits for the monopoly along the entire trajectory path. Therefore, with depreciation, monopoly power need not vanish either in the twinkling of an eye or asymptotically.

15.4 A nonrenewable resource

We now consider a problem in which an importer of a non-renewable resource uses a tariff to extract rents from the competitive owner of the resource. The analysis for the case where the resource owner is a monopoly is similar. Here we assume that the importer is the only consumer of the resource. Section 15.5 discusses a possible outcome when agents other than the tariff-wielding importer also consume the resource.

The seller receives the price p_t and the importer uses a unit tariff m_t , so the domestic price in the importing country is $p_t + m_t$. Let x be the stock of the resource, $c(x)$ the unit cost of extraction, and r the discount rate. The competitive seller, who takes the price path as given, chooses an extraction profile $\{y\}_{t=0}^{\infty}$ to maximize

$$\int_0^{\infty} e^{-rt} (p_t - c(x_t)) y_t dt$$

subject to

$$\dot{x} = -y, x_t \geq 0.$$

The optimality conditions yield the Hotelling condition

$$\dot{p} = r(p - c(x)). \quad (15.14)$$

The importing country obtains a flow of welfare equal to $U(y) - py$, where $U(y)$ is the utility of consuming at rate y and py is the cost of imports. Consumers actually pay $(p + m)y$, so the consumer surplus is $U(y) - (p + m)y$, but the tariff revenues my , are returned in a lump sum, so the net flow of consumer welfare is $U(y) - py$. The tariff affects welfare via its influence on the producer price, p .

We first consider the open loop solution, in which the importer is assumed to be able to commit at time 0 to an arbitrarily long tariff profile. We show that in general this trajectory is time inconsistent. We then consider a parametric example for which we can find the Markov perfect equilibrium.

The importer actually chooses the sequence of tariffs. However, it is convenient to write the problem as if the importer chooses the sequence of imports. Given this sequence, we can find the sequence of tariffs. If imports are y , the domestic price is $U'(y)$; the tariff equals the difference between the domestic price and the producer price, $m = U'(y) - p$.

The importer is constrained by the equation of motion for the stock, a physical constraint, and the Hotelling condition, a behavioral constraint (one that arises from sellers' optimizing behavior). The importer's current value Hamiltonian in the open loop equilibrium is

$$H = U(y) - py - \mu y + \lambda (r(p - c(x))),$$

where μ is the costate variable for the state variable x and λ is the costate variable for the "state variable" p .

Of course, p is not a typical state variable. The importer takes its law of motion as given – because this law is determined by the seller's optimizing behavior – but the initial value of the price is not given. When we introduced the control problem (Chapter xx) we noted that for the case of a finite horizon, there were two types of boundary condition. If the terminal value of the state variable is given (i.e. exogenous) the terminal value of the corresponding costate variable is endogenous. However, if the terminal value of the costate variable is free (and if there is no scrap function), then the terminal value of the costate variable must be 0. The intuition for this transversality condition is that the costate variable equals the shadow value of the state variable. If the state variable at the terminal time is free, under optimal behavior its shadow value must be 0.

A similar boundary condition holds here, but it applies at the initial rather than the terminal time. The initial price is determined by the entire sequence of future tariffs. By altering this sequence of tariffs, the importer alters the initial price. In an open loop trajectory the importer chooses the sequence of tariffs optimally, and thus chooses it so that the shadow value of the price, at time 0, equals 0. The fact that the initial value of p is endogenous means that, relative to the standard optimal control problem, we are missing one piece of information. However, the boundary condition $\lambda_0 = 0$ replaces that missing information, so we still have the right number of optimality conditions needed to obtain a unique solution to the maximization problem.

The other necessary conditions for optimality are

$$\begin{aligned} U'(y) - p - \mu &= 0 \implies m \equiv U'(y) - p = \mu \\ \dot{\lambda} - r\lambda &= y - r\lambda \implies \dot{\lambda} = y \implies \lambda_t = \int_0^t y_t dt = x_0 - x_t \\ \dot{\mu} - r\mu &= \lambda c'(x). \end{aligned}$$

The middle line uses the boundary condition $\lambda_0 = 0$. We now follow the usual procedure to obtain the Euler equation: we differentiate the algebraic equation (the first line above) and use the costate equations. The result is

$$U''(y)\dot{y} = r[U'(y) - c(x)] + (x_0 - x_t) c'(x_t). \quad (15.15)$$

If the importer were allowed to re-optimize at a later time, $s > 0$, it would face the same problem as above, except with initial stock x_s rather than x_0 . The Euler equation would have the same form, except that $(x_s - x_t) c'(x_t)$ would replace the term $(x_0 - x_t) c'(x_t)$. Therefore, the open loop path announced at time 0 is not time consistent, except for the special case where costs are stock independent. In that case, the importer can use the tariff to extract all of the rent from the seller, i.e. the importer can essentially expropriate the resource. When the importer obtains all of the rent from the resource, it has no incentive to re-optimize. Here the importer achieves the “first best”: there is no “distortion” from the importer’s standpoint, and no incentive to re-optimize.

We now consider the Markov perfect equilibrium. The importer has an infinitesimal period of commitment. The payoff-relevant state variable is the stock of the resource. The tariff is conditioned on the current stock of the resource, and the exporter understands this fact. The equilibrium

producer price is therefore a function of the stock of resource remaining, $p_t = P(x_t)$. The function P is endogenous; we need to find this function in order to calculate the equilibrium. An equilibrium price function must satisfy the Hotelling condition, equation 15.14, at every value of x between 0 and the initial condition, i.e. for all feasible values of the state variable.

Let $J(x)$ be the importer's value function. Using the same methods as we employed for the durable goods monopoly problem, the dynamic programming equation is

$$rJ(x) = \max_y U(y) - P(x)y - J_x(x)y.$$

Again, we need to obtain two functions, J and P , to obtain the equilibrium. In a standard control problem, we would need to obtain only the value function.

This problem resembles that of the durable goods monopoly with convex production costs. The optimal level of imports is bounded due to the concavity of $U(y)$. The first order condition implies

$$m(x) \equiv U'(y) - P(x) = J_x(x) > 0. \quad (15.16)$$

The inequality follows because the importer obtains surplus from imports, and the present value of the stream of this surplus increases with the remaining stock. The monopsonistic importer of the non-renewable resource exercises market power (uses a positive tariff), just as the monopolistic supplier of the durable good does, even though neither is able to commit to the actions of its successors. Equations 15.10 and 15.16 show that the monopoly markup (the difference between price and marginal cost) and the monopsonistic tariff are both positive along the equilibrium trajectory. In a competitive setting, the markup and the tariff are both 0.

If the domestic demand function is linear, with $U'(y) = \alpha - y$ and the unit extraction cost function is also linear, with $c(x) = \alpha - bx$, it is possible to obtain an explicit solution to the problem (see the problem set).

15.5 Disadvantageous power

The economic literature is replete with examples where the ability to exercise market power, or the ability to cooperate, or the possession of some other

feature that would seem to benefit the agents with that feature, actually leaves them worse off. This apparently paradoxical result might occur if agents are unable to make commitments about future behavior. These examples are all illustrations of the “theory of the second best”, which says (roughly) “Unless you begin at a local optimum, a small change, such as an increase in market power, leads to a first order welfare change that could be positive or negative.”

One of the earliest examples of this situation arises when a proper subset of oligopolists form a cartel (cite Salant, Switzer and Reynolds) and firms choose quantities. The equilibrium is Nash-Cournot both with and without the cartel. When the cartel exists, members set joint output to maximize the cartel’s joint profits. In the absence of the cartel, each firm sets output to maximize its own profits. Each firm creates an externality on other firms, because an increase in sales by firm i reduces the price, reducing other firms’ profits. The cartel internalizes that externality for its members, i.e. it takes into account how an increase in sales by any member affects the profits of all members. This internalization causes the cartel to reduce its cumulative production. The optimal response of non-members is then to increase their own output. Actions in this setting are “strategic substitutes”, meaning that an increase in agent i ’s action causes agent j to decrease its own action; here the “actions” are sales level.

The possible reduction in equilibrium profits, for cartel members, occurs only because of the cartel’s (assumed) inability to commit to a sales level. If the cartel could commit, at the time of formation, to its cumulative output, the increase in market power could not possibly be disadvantageous. With commitment ability, it is feasible for the cartel to commit to produce the non-cartel level, leaving it with the non-cartel profit level. Although feasible, this action is typically not optimal, so the cartel can strictly increase its joint profits by some other action. In the absence of an ability to commit, the cartel is forced by the logic of the game to sell at the Nash-Cournot production level, raising the possibility that its profits will be lower. In a game involving both investment and sales, industry profits may be lower and consumer welfare higher even if all firms in the industry form a cartel (Gatsios and Karp (cite)). This possibility arises if investment precedes the formation of the cartel, which then decides on industry supply.

Maskin and Newbery (cite) show that an importer’s market power can be disadvantageous in the non-renewable resource game. The key to this pos-

sibility is that the importer faces competing demand from other consumers. That is, the importer who uses the tariff is an oligopsonist, but not a monopsonist. In order to see the reason for this possibility, take the (admittedly extreme) case of a two-period model, in which the tariff-imposing importer has negligible demand for the resource in the first period, and high demand in the second period; competing importers have negligible demand in the second period and high demand in the first period. The stock of the resource is finite, costless to extract, and is supplied by a competitive owner. The resource cannot be stored once extracted (or can be stored only at high cost). Because the world ends after two periods, and extraction is costless, cumulative extraction and sales during the two periods must equal the initial stock.

The competitive owner's equilibrium behavior requires that if sales are positive in both periods, the ratio of second to first period prices is $1 + r$, where r is the discount rate. If the oligopsonistic importer is able to commit to a zero tariff in both periods, it buys a non-negligible quantity in the second period and obtains non-negligible surplus. However, a zero second period tariff is not subgame perfect, if the remaining stock in the second period is non-negligible. The importer is virtually a monopsonist in the second period, because of the assumption that competing demand in the second period is negligible. Therefore, the importer has an incentive to use a high second period tariff in order to essentially expropriate the resource. Under this tariff, the second period producer price is close to 0, for any non-negligible stock level. The owner with rational expectations, recognizing this outcome, decides to sell all (or almost all) of the resource in the first period. By assumption, the tariff-wielding importer has essentially no use for the resource in this period, and no practical way of storing it, so it receives almost no surplus in equilibrium.

A final example illustrates that cooperation can be disadvantageous (cite Rogoff). Consider a two period, two country competitive general equilibrium. In the first period, the stock of capital is predetermined. Investment decisions during this period determine the amount of capital in the second period. Capital is internationally mobile. The government in each country taxes labor income and capital income (profits, rent) in order to raise enough revenue to satisfy an exogenous budget constraint. We consider the two situations, first when the two governments are not able to cooperate with each other in setting their tax policies, and then when the governments are able

to cooperate.

Absent cooperation, the international mobility of capital means that countries face “tax competition” in both periods. If one country levies a tax on capital, the other country has the incentive to charge a slightly lower tax in order to induce capital to come that country. If the marginal productivity of capital is nearly constant, then a slightly lower tax induces almost all of the capital to flow to the country with a lower tax. Each country therefore has an incentive to undercut the other country’s tax on capital, leading to a near-zero equilibrium capital tax. In order to satisfy their budget constraints, both countries then have to use large labor taxes. Unless labor supply is perfectly inelastic, this tax creates a distortion in the labor market. The positive side of this lack of cooperation is that investors know that the second period equilibrium capital tax will be low, so they have a substantial incentive to invest.

With cooperation, the two countries can jointly determine the taxes. In the second period, the existing stock of capital is fixed, so using the insight from the optimal tax literature described above, it is optimal to raise all (or at least most) of the needed revenue in the second period using a capital tax. Foreseeing this high capital tax, agents have a lower incentive (relative to the non-cooperative scenario) to invest in capital. Cooperation thus increases first period joint welfare, because it makes it possible to raise the revenue without distorting the labor market. But cooperation may decrease second period joint welfare, because it leads to a reduction in the equilibrium level of capital.

One could generate an endless supply of examples of disadvantageous market power, or disadvantageous cooperation. The important thing to recognize is that these examples all arise due to agents’ inability to make commitments today about actions that will be taken in the future. The consequences of this inability may increase with the degree of market power of the agent taking the action. All of these examples illustrate the theory of the second best.

15.6 Problems

1. Obtain the formulae for the coefficients of the endogenous price function $P = A - BQ$ resulting from the linear rental function $F = a - bQ$ in

Section 15.3.1, where $\gamma > 0$. (Use a machine!)

2. For the linear example in Section 15.4, with $U'(y) = \alpha - y$ and $c(x) = \alpha - bx$, obtain a closed form solution to the Markov perfect equilibrium and also to the competitive equilibrium. Using some numerical values for parameters, calculate the percentage gain in welfare for the importer, and the percentage loss in welfare for the exporter, in moving from a competitive to a (Markov perfect) monopsonistic equilibrium. (In passing, determine whether the setting the coefficient for y in the function $U'(y)$, and using the same intercept (α) in the two functions $U'(y)$ and $c(x)$ entails any loss in generality.

16 Disentangling Intertemporal Substitutability from Risk Aversion

Up to date the intertemporally additive expected utility specification is the the prevailing framework for dynamic analysis in economics. It is well known that the framework implicitly assumes that a decision maker's aversion to risk coincides with his aversion to intertemporal fluctuations. ? and ? derived an alternative setting in which these two a priori quite different characteristics of preference can be disentangled. This section presents the basic welfare framework and derives a related asset pricing formula. We introduce the concept of intertemporal risk aversion and of an intrinsic preference for the timing of risk resolution, both of which are closely related with the disentanglement of risk vs intertemporal consumption smoothing.

16.1 An Intertemporal Model of Risk Attitude

We will refer to the 'standard model' as the modeling framework where a decision maker evaluates utility separately for every period and for every state of the world and then sums it over time and over states. Formally let $\mathbf{x} = (x_0, x_1, x_2, \dots)$ denote a consumption path and $p(\mathbf{x})$ a probability distribution over the latter. Then, assuming stationary preferences, the decision maker's overall welfare is given by

$$U = E_p \sum_t \beta^t u(x_t)$$

where β is the utility discount factor. The curvature of u captures the decision maker's aversion to consumption fluctuations. Because the same utility function is used to aggregate over time and over risk, the decision maker's aversion to (certain) intertemporal fluctuations is the same as his aversion to risk fluctuations corresponding to different states of the world.

A priori, however, risk aversion and a decision maker's propensity to smooth consumption over time are two distinct concepts. Recall that (von Neumann & Morgenstern 1944) axioms are underlying the notion that in an atemporal setting we can evaluate risky scenarios by means of a von-Neumann Morgenstern utility index u^{vNM} so that

$$U = E u^{\text{vNM}}(x)$$

represents preferences. Here, the curvature of u^{vNM} captures risk aversion. In particular, the Arrow-Pratt measure of relative risk aversion $\text{RRA}(x) = \frac{u^{\text{vNM}''}(x)}{u^{\text{vNM}'}(x)}x$ attaches a numeric value to curvature and risk aversion that does not depend on affine transformations of u^{vNM} .¹⁰⁷ A similar set of axioms of additive separability¹⁰⁸ have been formulated for preferences evaluating consumption paths under certainty. These axioms give rise to a preference representation on certain consumption paths that takes the form

$$U = \sum_t u_t^{\text{int}}(x_t). \quad (16.1)$$

Assuming in addition a stationary evaluation of certain consumption paths makes the utility in the different periods coincide up to a common discount factor β :

$$U = \sum_t \beta^t u^{\text{int}}(x_t). \quad (16.2)$$

In equation (16.2) the concavity of the utility function u^{int} describes aversion to intertemporal consumption volatility. In a one commodity setting¹⁰⁹ this aversion to intertemporal volatility can be measured by means of the consumption elasticity of marginal utility $\eta = \frac{u^{\text{int}''}(x)}{u^{\text{int}'}(x)}x$.¹¹⁰ Note that this measure exactly corresponds to the Arrow Pratt measure of relative risk aversion, only in the context of periods rather than risk states. Instead of calling η an aversion measure to intertemporal volatility, we can also characterize it as a decision maker's propensity to smooth consumption over time.

¹⁰⁷The invariance under affine transformations is a desideratum born from the fact that a function $u^{\text{vNM}'} = au^{\text{vNM}} + b$ with $a, b \in \mathbb{R}, a > 0$ represents the same underlying preferences as does u^{vNM} .

¹⁰⁸See ?, ?, ?, ?, ?, ?, and ? for various axiomatizations of additive separability over time. Other than the von Neumann & Morgenstern (1944) axioms, these axioms allow for period specific utility functions (which would correspond to a state dependent utility model in the risk setting).

¹⁰⁹? generalized the one commodity measure by Arrow Pratt to a multi-commodity setting. Here risk aversion becomes good-specific and corresponds to the concavity of the utility function along a variation of the particular commodity keeping the others constant. The same concept of a multi-dimensional measure could be applied to intertemporal substitutability.

¹¹⁰Which is the inverse of the intertemporal elasticity of substitution.

A priori the utility functions u^{vNM} and u^{int} are two distinct objects that carry two different types of information. In Traeger (2010a) I derive this intuition more formally. The paper combines the von Neumann-Morgenstern axioms with the assumption that certain consumption path can be evaluated in the additively separable form (16.1) (or 16.2). The resulting preference representation features two independent functions that – in a one commodity setting – can be identified with the functions u^{vNM} and u^{int} above.

Let us derive the form of such an evaluation intuitively for a 2 period setting with certain consumption in the first and uncertain consumption in the second period ('certain \times uncertain' setting). Let

$$U_1(x_1, x_2) = u^{\text{int}}(x_1) + \beta u^{\text{int}}(x_2) \quad (16.3)$$

represent preferences over certain (two period) consumption paths. Let

$$U_2(p) = \text{E}u^{\text{vNM}}(x)$$

represent preferences over second period lotteries. Define x_2^p as the second period certainty equivalent to the lottery p by

$$\begin{aligned} U_2(x_2^p) &= u^{\text{vNM}}(x_2^p) \stackrel{!}{=} U_2(p) = \text{E}_p u^{\text{vNM}}(x) \\ \Rightarrow x_2^p &= u^{\text{vNM}^{-1}} \left[\text{E}_p u^{\text{vNM}}(x) \right] . \end{aligned} \quad (16.4)$$

Now use the certainty equivalent to extend the evaluation functional in equation (16.3) to a setting of uncertainty by defining

$$\begin{aligned} U(x_1, p) &= U_2(x_1, x_2^p) = u^{\text{int}}(x_1) + \beta u^{\text{int}}(x_2^p) \\ &= u^{\text{int}}(x_1) + \beta u^{\text{int}} \left(u^{\text{vNM}^{-1}} \left[\text{E}_p u^{\text{vNM}}(x) \right] \right) . \end{aligned}$$

In taking the inverse of u^{vNM} we have assumed a one dimensional setting. In general, we can show that u^{vNM} is always a strictly monotonic transformation of u^{int} (Traeger 2010a). Thus defining f by $u^{\text{vNM}} = f \circ u^{\text{int}}$ we can transform the equality (16.4) into

$$\begin{aligned} f \circ u^{\text{int}}(x_2^p) &\stackrel{!}{=} \text{E}_p f \circ u^{\text{int}}(x) \\ \Leftrightarrow u^{\text{int}}(x_2^p) &= f^{-1} \left[\text{E}_p f \circ u^{\text{int}}(x) \right] . \end{aligned}$$

and obtain in combination with equation (16.3) the representation

$$U(x_1, p) = u^{\text{int}}(x_1) + \beta f^{-1} [\mathbb{E}_p f \circ u^{\text{int}}(x)]$$

for the multi-commodity framework.

For a general time horizon, the corresponding preference representation evaluates consumption recursively. To write the evaluation formally we need some notation for the underlying choice objects. The reason is that in a recursive evaluation general objects of choice are no longer lotteries over consumption paths, but decision trees. While a lottery over a consumption path only contains information of the probability that a particular event takes place in some period t , the decision tree also contains information how we get to that event in period t , i.e. which lottery takes place in which period in order to give us the probability for the event in period t . As we will see in section 16.4, the particular way how we get to an event in period t might influence how we evaluate it. More precisely, it is possible that our evaluation of a lottery in the future not only depends on the probabilities attached to each of the outcomes, but also on the period in which the risk resolves (an information that is part of the decision tree, but not of the resulting probability distribution over the corresponding consumption paths).

Formally, a decision tree is a recursion of probability distributions over subtrees. General choice objects in the last period are lotteries p_T over consumption x_T . In the second last period certain consumption is denoted by x_{T-1} . However, before uncertainty resolves, the decision maker neither knows x_{T-1} with certainty, nor does he know the probability distribution that he faces at the beginning of period T with certainty. Instead, at the beginning of period $T-1$ he faces a lottery over the tuple (x_{T-1}, p_T) , which we will denote by p_{T-1} . Observe that each tuple (x_{T-1}, p_T) corresponds to a subtree and p_{T-1} gives us the weights attached to all possible subtrees. A recursive construction defining p_{t-1} as a lottery over tuples (x_{t-1}, p_t) for all $t \in \{1, \dots, T\}$ gives us the desired formal representation of a decision tree in the present p_0 .

The general recursive evaluation of such a decision tree is as follows. Consumption in the last period of a finite time horizon is simply $u_T(x_T)$ and then

$$\mathbf{u}_{t-1}(x_{t-1}, p_t) = u_t(x_{t-1}) + f_t^{-1} [\mathbb{E}_{p_t} f_t \circ \mathbf{u}_t(x_t, p_{t+1})] . \quad (16.5)$$

The function \mathbf{u}_{t-1} represents preferences in period $t-1$ after the uncertainty

with respect to $t-1$ has resolved. Note that each u_t evaluates choices between (sub-)trees that describe a unique future scenario with no further choices in later periods (open loop setting).¹¹¹ We will see in the next section how this evaluation corresponds to a value function depending on states rather than probability trees over the future.

Alternatively we can formulate our setting in a way that choice in every period takes place before uncertainty with respect to that period has resolved. Then, the functional $M_t(p_t) = f_t^{-1} [E_{p_t} f_t \circ u_t(x_t, p_{t+1})]$ is applied to evaluate the lottery p_t over tuples (x_t, p_{t+1}) . The corresponding recursion relation can also be written as

$$M_t(p_t) = f_t^{-1} [E_{p_t} f_t (u_t(x_t) + \beta M_{t+1}(p_{t+1}))],$$

specifying the value of a decision tree before uncertainty over a particular node (period) has resolved.

If we look at the choice between certain consumption paths the expected value operator drops out and the functions f_t^{-1} and f_t cancel. Then we recognize that – just like in the certain \times uncertain setting – the utility function u_t in equation (16.5) corresponds to our utility function u_t^{int} which describes the decision maker's propensity to smooth consumption over time. On the other hand, if we simply look at an uncertain evaluation in the last period, we recognize again that $f_t \circ u_t$ corresponds to u_t^{vNM} . Thus the composition $f_t \circ u_t$ expresses risk aversion in the Arrow Pratt sense.

Letting the time horizon approach infinity, we face an infinitely long and infinitely wide decision tree in every period. See ? for a formal definition of such an infinite decision tree. The advantage of the infinite time horizon is that our set of choice objects becomes the same in every period.¹¹² That

¹¹¹Let me point out a similarity between the (necessity of the) recursivity in the open loop evaluation in equation (16.5) and the (necessity of the) recursivity in a closed loop dynamic programming equation for the standard model. In a feedback optimization with standard preferences the value function in the next period depends on the realization of some random variable. That dependence is based on the fact that the realization of the random variable changes real outcomes and optimal choices in the future. In equation (16.5) the future value function depends on the realization of the random variables because the decision maker cares for the fact whether uncertainty has resolved earlier or later, even if this realization does not change real outcomes and choices. The underlying intrinsic (rather than instrumental) preference for the timing of uncertainty resolution is discussed in section 16.4.

¹¹²Of course, the set of *feasible* choices can vary between periods.

is, if P^∞ denotes the set of all infinite decision trees and $p_{t+1}^\infty \in P^\infty$, then also the lottery p_t^∞ over tuples (x_t, p_{t+1}^∞) is in P^∞ . Under the appropriate stationarity conditions we can rewrite equation (16.5) as¹¹³

$$\mathbf{u}(x_{t-1}, p_t) = (1 - \beta)u(x_{t-1}) + \beta f^{-1} [\mathbb{E}_{p_t} f \circ \mathbf{u}(x_t, p_{t+1})] . \quad (16.7)$$

16.2 Disentangling Risk Aversion and Intertemporal Substitutability

? famously developed a one commodity special case of the model in equation (16.7).¹¹⁴ Let $u(x) = x^\rho$ implying a CIES function over certain consumption paths with the inverse of the intertemporal elasticity of substitution given by $\eta = 1 - \rho$. Moreover assume constant relative Arrow Pratt risk aversion so that $f \circ u(x) = x^\alpha$ implying $f(z) = z^{\frac{\alpha}{\rho}}$. We assume $\alpha, \rho > 0$, an assumption that we adopt for simplicity only for the subsequent transformation. Then equation (16.7) becomes

$$\begin{aligned} \mathbf{u}(x_{t-1}, p_t) &= (1 - \beta)x_{t-1}^\rho + \beta \left[\mathbb{E}_{p_t} \mathbf{u}_t(x_t, p_{t+1})^{\frac{\alpha}{\rho}} \right]^{\frac{\rho}{\alpha}} \\ \Leftrightarrow V^*(x_{t-1}, p_t) &= \left((1 - \beta)x_{t-1}^\rho + \beta \left[\mathbb{E}_{p_t} (V^*(x_t, p_{t+1}))^\alpha \right]^{\frac{\rho}{\alpha}} \right)^{\frac{1}{\rho}} , \end{aligned} \quad (16.8)$$

where $V^* = \mathbf{u}^{\frac{1}{\rho}}$. Here, we admit $\alpha, \rho \in \mathbb{R}$.

A major contributions of ? is to derive a framework for estimating this type of preferences in the context of asset pricing. Let the probability measures $p_t \in P^\infty$ be generated by a stationary stochastic process $(\tilde{R}_t, \tilde{z}_t)_{-\infty}^\infty$, where $R_t \in [\underline{R}, \bar{R}]^K$ is a vector specifying asset returns and $z_t \in \mathbb{R}^Z$ captures other random information relevant to predicting future probabilities. We use a tilde for emphasizing the random nature of a variable. The state of the system is specified by wealth A_t and the history of realized asset returns and information variables, which is denoted by $I_t = (R_t, z_t)_{-\infty}^{t-1} \in \mathbf{I}_t = \times_{i=-\infty}^{t-1} [\underline{R}, \bar{R}]^K \times \mathbb{R}^Z$. The equation of motion for the agent's wealth is

$$A_{t+1} = (A_t - x_t)\omega_t \tilde{R}_t \quad (16.9)$$

¹¹³Note that the formulation is equivalent to $\mathbf{u}^*(x_{t-1}, p_t) = u(x_{t-1}) + \beta f^{*-1} [\mathbb{E}_{p_t} f^* \circ \mathbf{u}^*(x_t, p_{t+1})]$ with $\mathbf{u}(x_{t-1}, p_t) = (1 - \beta)\mathbf{u}^*(x_{t-1}, p_t)$ and $f(x) = f^*\left(\frac{x}{1-\beta}\right)$.

¹¹⁴In ? the authors also analyze a more general isoelastic model where risk aversion does not satisfy the von Neumann-Morgenstern axioms.

where ω_t is a vector whose component $\omega_{k,t}$ characterizes the share of assets invested in the k^{th} asset that has return $\tilde{R}_{k,t}$. In this setting, the probability distribution over future consumption derives from wealth, the stochastic process $(\tilde{R}_t, \tilde{z}_t)_{-\infty}^{\infty}$, and the (optimal) choice of the control variables x_t and ω_t .

Now we have an optimization problem with given equations of motions. We replace the recursive utility function (16.8) by the value function $V(A_t, I_t)$, which captures the maximal achievable utility as a function of wealth A_t and information I_t . The corresponding Bellman equation is

$$V(A_t, I_t) = \max_{x_t, \omega_t} \left((1 - \beta)x_t^\rho + \beta \mathbb{E} \left[V((A_t - x_t)\omega_t \tilde{R}_t, \tilde{I}_{t+1})^\alpha | I_t \right]^\frac{\rho}{\alpha} \right)^\frac{1}{\rho} \quad (16.10)$$

Note that now V is a ‘regular’ value function of a dynamic programming problem (16.10) and the choices we evaluate in period t are no longer degenerate for periods $t' > t$. Instead, we assume an optimal choice in every period as we recursively evaluate the future.

We will show that the value function is linear in wealth and multiplicatively separable in wealth and information. The *trial solution* $V(A_t, I_t) = A_t \Phi_t(I_t)$ yields the *Bellman equation*

$$\begin{aligned} A_t \Phi_t(I_t) &= \max_{x_t, \omega_t} \left((1 - \beta)x_t^\rho + \beta \mathbb{E} \left[\left((A_t - x_t)\omega_t \tilde{R}_t \Phi_{t+1}(\tilde{I}_{t+1}) \right)^\alpha | I_t \right]^\frac{\rho}{\alpha} \right)^\frac{1}{\rho} \\ \Leftrightarrow A_t \Phi_t(I_t) &= \max_{x_t, \omega_t} \left((1 - \beta)x_t^\rho + \beta (A_t - x_t)^\rho \mathbb{E} \left[\left(\omega_t \tilde{R}_t \Phi_{t+1}(\tilde{I}_{t+1}) \right)^\alpha | I_t \right]^\frac{\rho}{\alpha} \right)^\frac{1}{\rho} \end{aligned} \quad (16.11)$$

where the second line uses that A_t and x_t are known at time t .

The *first order condition for consumption* optimization, assuming an optimal portfolio composition ω_t^* , is

$$\begin{aligned} \rho(1 - \beta)x_t^{\rho-1} &= \beta \rho (A_t - x_t)^{\rho-1} \mathbb{E} \left[\left(\omega_t^* \tilde{R}_t \Phi_{t+1}(\tilde{I}_{t+1}) \right)^\alpha | I_t \right]^\frac{\rho}{\alpha} \\ \Leftrightarrow (1 - \beta)x_t^{\rho-1} (A_t - x_t) &= \beta (A_t - x_t)^\rho \mathbb{E} \left[\left(\omega_t^* \tilde{R}_t \Phi_{t+1}(\tilde{I}_{t+1}) \right)^\alpha | I_t \right]^\frac{\rho}{\alpha} \end{aligned} \quad (16.12)$$

Instead of solving this first order condition for x_t , we use equation (16.12)

and substitute it back into the (*now optimized*) Bellman equation (16.11)

$$\begin{aligned}
 A_t \Phi_t(I_t) &= \left((1 - \beta)x_t^\rho + (1 - \beta)x_t^{\rho-1}(A_t - x_t) \right)^{\frac{1}{\rho}} \\
 \Leftrightarrow A_t \Phi_t(I_t) &= \left((1 - \beta)x_t^\rho + (1 - \beta)x_t^{\rho-1}A_t - (1 - \beta)x_t^\rho \right)^{\frac{1}{\rho}} \\
 \Leftrightarrow \Phi_t(I_t) &= (1 - \beta)^{\frac{1}{\rho}} \left(\frac{x_t}{A_t} \right)^{\frac{\rho-1}{\rho}}. \tag{16.13}
 \end{aligned}$$

Equation (16.13) states that (also) the optimal consumption choice

$$x_t^* = \frac{\Phi_t(I_t)^{\frac{\rho}{\rho-1}}}{(1 - \beta)^{\frac{1}{\rho-1}}} A_t \equiv \Psi(I_t) A_t \tag{16.14}$$

is linear in wealth.

Note 1: Employing the linearity of consumption in wealth in equation (16.14) to eliminate the wealth from the Bellman equation (16.11) results in a reduced Bellman equation that implicitly characterizes the value of information $\Phi_t(I_t)$

$$\Phi_t(I_t) = \max_{\omega_t} \left(\left[(1 - \beta) \frac{\Phi_t(I_t)^{\frac{\rho}{\rho-1}}}{(1 - \beta)^{\frac{1}{\rho-1}}} \right]^\rho + \beta \left[1 - \frac{\Phi_t(I_t)^{\frac{\rho}{\rho-1}}}{(1 - \beta)^{\frac{1}{\rho-1}}} \right]^\rho \mathbb{E} \left[\left(\omega_t \tilde{R}_t \Phi_{t+1}(\tilde{I}_{t+1}) \right)^\alpha \mid I_t \right] \right)^{\frac{1}{\rho}} \tag{16.15}$$

Thus, the trial solution indeed solves the problem as long as the equation (16.15) has a solution.

Note2: From $V(A_t, I_t) = A_t \Phi_t(I_t)$ and equation (16.13) follows

$$\frac{V(A_t, I_t)}{x_t} = \frac{A_t}{x_t} \Phi_t(I_t) = \frac{A_t}{x_t} (1 - \beta)^{\frac{1}{\rho}} \left(\frac{x_t}{A_t} \right)^{\frac{\rho-1}{\rho}} = \left((1 - \beta) \frac{A_t}{x_t} \right)^{\frac{1}{\rho}}.$$

The value function normalized by current consumption is proportional to the wealth-consumption-ratio $\frac{A_t}{x_t}$ raised to the power of $\frac{1}{\rho} = \frac{1}{1-\eta} = \frac{IES}{IES-1}$, where *IES* abbreviates the intertemporal elasticity of substitution.

We now return to the *first order condition for consumption* and use equation (16.13) for $t + 1$ to replace the unknown information part of the value

function $\Phi_{t+1}(\tilde{I}_{t+1})$ in equation (16.12), and then we use the budget constraint (16.9) to replace A_{t+1} , yielding

$$\begin{aligned}
 (1 - \beta)x_t^{\rho-1}(A_t - x_t) &= \beta(A_t - x_t)^\rho \mathbb{E} \left[\left(\omega_t^* \tilde{R}_t (1 - \beta)^{\frac{1}{\rho}} \left(\frac{\tilde{x}_{t+1}}{\tilde{A}_{t+1}} \right)^{\frac{\rho-1}{\rho}} \right)^\alpha \middle| I_t \right]^{\frac{\rho}{\alpha}} \\
 \Leftrightarrow x_t^{\rho-1} &= \beta(A_t - x_t)^{\rho-1} \mathbb{E} \left[\left(\omega_t^* \tilde{R}_t \left(\frac{\tilde{x}_{t+1}}{(A_t - x_t)\omega_t^* \tilde{R}_t} \right)^{\frac{\rho-1}{\rho}} \right)^\alpha \middle| I_t \right]^{\frac{\rho}{\alpha}} \\
 \Leftrightarrow 1 &= \beta \mathbb{E} \left[\left(\omega_t^* \tilde{R}_t \left(\frac{A_t - x_t}{x_t} \frac{\tilde{x}_{t+1}}{(A_t - x_t)\omega_t^* \tilde{R}_t} \right)^{\frac{\rho-1}{\rho}} \right)^\alpha \middle| I_t \right]^{\frac{\rho}{\alpha}} \\
 \Leftrightarrow 1 &= \mathbb{E} \left[\left(\beta \omega_t^* \tilde{R}_t \left(\frac{\tilde{x}_{t+1}}{x_t} \right)^{\rho-1} \right)^\alpha \middle| I_t \right]^{\frac{\rho}{\alpha}} \tag{16.16}
 \end{aligned}$$

$$\Leftrightarrow \beta^{-1} = \mathbb{E} \left[\left(\omega_t^* \tilde{R}_t \left(\frac{\tilde{x}_{t+1}}{x_t} \right)^{\rho-1} \right)^\alpha \middle| I_t \right]^{\frac{\rho}{\alpha}} \tag{16.17}$$

That is our first Euler equation. However, we have K assets and, thus, are missing another $K - 1$ equations to determine our optimal asset dynamics. To find the missing equations we optimize the portfolio choice explicitly in equation (16.11). Using equations (16.13) and (16.9) find that this maximization problem is equivalent to

$$\begin{aligned}
 &\max_{\omega_t} \mathbb{E} \left[\left(\omega_t \tilde{R}_t \Phi_{t+1}(\tilde{I}_{t+1}) \right)^\alpha \middle| I_t \right] \\
 \Leftrightarrow &\max_{\omega_t} \mathbb{E} \left[\left(\omega_t \tilde{R}_t (1 - \beta)^{\frac{1}{\rho}} \left(\frac{\tilde{x}_{t+1}}{(A_t - x_t)\omega_t \tilde{R}_t} \right)^{\frac{\rho-1}{\rho}} \right)^\alpha \middle| I_t \right] \\
 \Leftrightarrow &\max_{\omega_t} \mathbb{E} \left[\left(\omega_t \tilde{R}_t \right)^\alpha \left(\frac{\tilde{x}_{t+1}}{x_t} \right)^{\alpha \frac{\rho-1}{\rho}} \middle| I_t \right],
 \end{aligned}$$

as $(A_t - x_t)$ (eliminated) and x_t (inserted) are known in t . Deriving the first order conditions we have to keep in mind that $\sum_k \omega_{k,t} = 1$. Letting

$d\omega_k = -d\omega_1$ we find

$$\mathbb{E} \left[\left(\omega_t \tilde{R}_t \right)^{\frac{\alpha}{\rho}-1} \left(\frac{\tilde{x}_{t+1}}{x_t} \right)^{\frac{\alpha}{\rho}(\rho-1)} \left(\tilde{R}_{k,t} - \tilde{R}_{1,t} \right) | I_t \right] = 0. \quad (16.18)$$

Equations (16.16) and (16.18) (for $k = 2, \dots, K$) together represent the Euler equations. We can also combine them as follows. Multiply (16.18) by $\omega_{k,t}$ and sum over $k = 1, \dots, K$ yielding

$$\mathbb{E} \left[\left(\omega_t \tilde{R}_t \right)^{\frac{\alpha}{\rho}} \left(\frac{\tilde{x}_{t+1}}{x_t} \right)^{\frac{\alpha}{\rho}(\rho-1)} | I_t \right] - \mathbb{E} \left[\left(\omega_t \tilde{R}_t \right)^{\frac{\alpha}{\rho}-1} \left(\frac{\tilde{x}_{t+1}}{x_t} \right)^{\frac{\alpha}{\rho}(\rho-1)} \tilde{R}_{1,t} | I_t \right] = 0.$$

By equation (16.17) it follows that

$$\mathbb{E} \left[\left(\omega_t \tilde{R}_t \right)^{\frac{\alpha}{\rho}-1} \left(\frac{\tilde{x}_{t+1}}{x_t} \right)^{\frac{\alpha}{\rho}(\rho-1)} \tilde{R}_{1,t} | I_t \right] = \beta \frac{\rho}{\alpha}.$$

Therefore we can write the Euler equations as

$$\mathbb{E} \left[\beta \frac{\alpha}{\rho} \left(\omega_t \tilde{R}_t \right)^{\frac{\alpha}{\rho}-1} \left(\frac{\tilde{x}_{t+1}}{x_t} \right)^{\frac{\alpha}{\rho}(\rho-1)} \tilde{R}_{k,t} | I_t \right] = 1 \quad (16.19)$$

for $k = 1, \dots, K$. In either form these Euler equations have been used as point of departure for estimating ρ and α by comparing interest on individual assets with interest gained on the market portfolio (see in particular ?, ?, Vissing-Jørgensen & Attanasio 2003, Basal & Yaron 2004).

We first discuss the special case obtained in the intertemporally additive expected utility standard model where $\alpha = \rho$. Then, equation (16.19) becomes

$$\mathbb{E} \left[\beta \left(\frac{\tilde{x}_{t+1}}{x_t} \right)^{\rho-1} \tilde{R}_{k,t} | I_t \right] = 1. \quad (16.20)$$

Often, the return of an asset k is its future price $P_{k,t+1}$ plus its payoff in dividends $D_{k,t+1}$, relative to its present price P_t . Then $R_{k,t} = \frac{P_{k,t+1} + D_{k,t+1}}{P_{k,t}}$ and

$$P_{k,t} = \mathbb{E} \left[\beta \left(\frac{\tilde{x}_{t+1}}{x_t} \right)^{\rho-1} (P_{k,t+1} + D_{k,t+1}) | I_t \right] = 1.$$

The equations state that an asset's return is determined by its covariance with consumption growth, or, more precisely, by its covariance with marginal utility¹¹⁵ $\frac{u'(\tilde{x}_{t+1})}{u'(x_t)} = \left(\frac{\tilde{x}_{t+1}}{x_t}\right)^{\rho-1}$. Here, the expression $m_{t,t+1} \equiv \beta \frac{u'(\tilde{x}_{t+1})}{u'(x_t)}$ is also referred to as the stochastic discount factor. It discounts future asset payoffs depending on the state of the world. We observe from equation (16.20) by analyzing the payoff of a risk free asset called $R_{f,t}$ that

$$\mathbb{E} \left[\beta \left(\frac{\tilde{x}_{t+1}}{x_t} \right)^{\rho-1} \mid I_t \right] = \frac{1}{R_{f,t}}$$

and therefore the expected value of the stochastic discount factor is

$$\mathbb{E}_t m_{t,t+1} = \mathbb{E}_t \beta \frac{u'(x_{t+1})}{u'(x_t)} = \frac{1}{R_{f,t}},$$

i.e. the inverse of the return (factor) on a certain investment.

Equation (16.20) and the idea that an asset is to be priced based on its covariance with consumption is a well known result of consumption based capital asset pricing models. However, empirical evidence has not been very supportive of the fact that such a behavior is actually observed on asset markets. Hansen and Jagannathan 1991 (?) have shown how the Euler equations (16.20) provide an upper bound for the Sharpe ratio, which is a measure for the excess return on the market portfolio per unit of market risk. For lognormally distributed consumption and power utility (as above) this Sharpe ratio would be bound by the product of the standard deviation of consumption growth and $\eta = 1 - \rho$ (the consumption elasticity of marginal utility or aversion to intertemporal substitution). As data of US stocks over the past century reveal, consumption volatility has been too low to explain high excess return of the market portfolio over the certain interest rate, at least with reasonable values for the consumption elasticity of marginal utility.¹¹⁶ This finding is referred to as the *equity premium puzzle*. If we were to accept an unreasonably high consumption elasticity of marginal utility, then we run straight into another asset pricing puzzle. Recall the stochastic

¹¹⁵Here utility characterizes utility in the sense of intertemporal aggregation, or, the utility function in the model where we write aggregation in terms of utility u and intertemporal risk aversion f .

¹¹⁶Which would have to be somewhere in the range from 10 to 50 in order to expand the bounds on the Sharpe ratio sufficiently to describe the observed returns.

Ramsey equation of section 14.5, which we will generalize further below. It tells us that the certainty equivalent interest rate is $r_f = \rho + \eta\mu_x - \eta^2\frac{\sigma_x^2}{2}$. If we plug in a lower bound for $\rho = 0$ and derive from the same data $\mu_x = 0.018$ and $\sigma_x = 0.036$ we find $r_f \approx 0\% + \eta 1.8\% - \eta^2 0.06\%$. For a value of $\eta = 10$ which is a lower bound to the consumption elasticity of substitution needed to explain the equity premium we find that the certain interest rate must already be $r_f \approx 12\%$, a value that is significantly too high. This finding is referred to as the *risk-free rate puzzle*.¹¹⁷

In contrast to the consumption based asset pricing model resulting in equation (16.20), the simple static capital asset pricing model (CAPM) based on mean variance analysis tells us that the return of an asset is determined by its covariance with the market portfolio.¹¹⁸ Equation (16.19) shows that in the general case with disentangled preferences the asset's return is determined by the covariance with both of these factors. More precisely, what counts for the asset's return is its covariance with a geometric mean (with weighting factor $\frac{\alpha}{\rho}$) of marginal utilities of consumption and the return of the market portfolio.

Finally, we return to the stochastic Ramsey equation determining the risk free rate of interest in an uncertain world (see section 14.5). For this purpose, let the first asset be a bond with a fixed interest rate r^* . Moreover, denote the pure rate of time preference by $\delta = -\ln \beta$, the consumption growth rate by $\tilde{y}_t = \ln \frac{\tilde{x}_{t+1}}{x_t}$, and the interest gained on the market portfolio ωR by $\tilde{r}_t = \ln(\omega_t R_t)$. The inverse of the intertemporal elasticity of substitution generally employed to parameterize CIES functions is $\eta = 1 - \rho$. Then

¹¹⁷For more recent data the puzzle seems to be slightly less strong than indicated by the cited historic US data. See Pennachi 2008 p.90 ? for the values employed here and further sources.

¹¹⁸Note that the standard CAPM and the consumption based pricing model coincide in a static setting with quadratic utility or normal asset returns. If utility is quadratic, marginal utility is linear in consumption and (end of period) consumption is linear in the return of the market portfolio. Then marginal utility is perfectly negatively correlated with the return on the market portfolio and consumption based asset pricing coincides with the CAPM approach. For general utility, normal asset returns, and a static model, end of period consumption is perfectly correlated with the return of the market portfolio, which, therefore, is perfectly negatively correlated with marginal utility. Once more the asset pricing prediction coincides with the beta formula of the standard CAPM.

equation (16.19) for $k = 1$ can be rewritten as

$$\mathbb{E} \left[e^{-\frac{\alpha}{\rho} \delta} e^{(\frac{\alpha}{\rho}-1)\tilde{r}_t} e^{-\eta \frac{\alpha}{\rho} \tilde{y}_t} | I_t \right] = e^{-r^*} \quad (16.21)$$

$$\Leftrightarrow \mathbb{E} \left[e^{-[\frac{\alpha}{\rho} \delta + \frac{\alpha}{\rho} \eta \tilde{y}_t + (1-\frac{\alpha}{\rho})\tilde{r}_t]} | I_t \right] = e^{-r^*}$$

$$\Leftrightarrow \mathbb{E} \left[e^{-[(\delta + \eta \tilde{y}_t) - (1-\frac{\alpha}{\rho})(\delta + \eta \tilde{y}_t - \tilde{r}_t)]} | I_t \right] = e^{-r^*} \quad (16.22)$$

We can spell the equation out further if we assume that the market portfolio return and consumption growth in $t + 1$ are jointly lognormally distributed, conditional on information at time t . Then

$$\begin{aligned} r^* &= \frac{\alpha}{\rho} \delta + \frac{\alpha}{\rho} \eta \mu_{y_t} + \left(1 - \frac{\alpha}{\rho}\right) \mu_{r_t} - \left(\frac{\alpha}{\rho}\right)^2 \eta^2 \frac{\sigma_{y_t}}{2} - \left(1 - \frac{\alpha}{\rho}\right)^2 \frac{\sigma_{r_t}}{2} \\ &\quad + \eta^2 \frac{\alpha}{\rho} \left(1 - \frac{\alpha}{\rho}\right) \text{cov}_{y_t, r_t} \\ &= \frac{\alpha}{\rho} \left[\delta + \eta \mu_{y_t} - \frac{\alpha}{\rho} \eta^2 \frac{\sigma_{y_t}}{2} \right] \\ &\quad + \left(1 - \frac{\alpha}{\rho}\right) \left[\mu_{r_t} - \left(1 - \frac{\alpha}{\rho}\right) \frac{\sigma_{r_t}}{2} + \eta^2 \frac{\alpha}{\rho} \text{cov}_{y_t, r_t} \right]. \end{aligned} \quad (16.23)$$

For the standard model where risk aversion and aversion to intertemporal substitution coincide we have $\alpha = \rho$ and equation (16.23) collapses to

$$r^* = \delta + \eta \mu_{y_t} - \eta^2 \frac{\sigma_{y_t}^2}{2},$$

which is the stochastic Ramsey equation that we already encountered in section 14.5. In general, however, the certainty equivalent discount rate is determined not only by time preference and consumption growth, but as equations (16.22) and (16.23) show the risk free rate depends on the joint distribution of consumption growth and return of the market portfolio. It thereby is a sort of an arithmetic mean between the two contribution with weight $1 - \frac{\alpha}{\rho}$. The latter characterizes the difference between Arrow Pratt risk aversion and intertemporal substitutability. The next section derives an interpretation of the term $1 - \frac{\alpha}{\rho}$ as a measure of relative intertemporal risk aversion.

16.3 Intertemporal Risk Aversion

This section derives a measure for the difference between Arrow Pratt risk aversion and intertemporal substitutability. This measure will have an interesting interpretation of risk aversion itself. Let us recall the general stationary representation in equation (16.6)

$$\mathbf{u}(x_{t-1}, p_t) = (1 - \beta)u(x_{t-1}) + \beta f^{-1} [E_{p_t} f \circ \mathbf{u}_t(x_t, p_{t+1})] .$$

In section 16.1 we elaborated the interpretation of u as characterizing the attitude toward intertemporal consumption fluctuations and the interpretation of $f \circ u \equiv h$ as characterizing Arrow Pratt risk aversion. Thus, the concavity of $f = h \circ u^{-1}$ is a measure for the difference between aversion to intertemporal fluctuations and aversion to risk in the Arrow Pratt sense. For example, f concave is equivalent to h being more concave than u . This relation corresponds to a possible definition of ‘more concave than’ and is based on the fact that if f is concave then $h = f \circ u$ is a concave transformation of u and thus ‘more concave’.¹¹⁹

An interesting alternative characterization of f is obtained by the following axiom. Let $\mathbf{X} = X^T$ be the space of consumption paths where T might be infinite. Let a preferences \succeq be represented by equation (16.6) (over certain paths as well as probability trees of the according length). For two given consumption paths \mathbf{x} and \mathbf{x}' define the ‘best off combination’ path $x^{\text{high}}(\mathbf{x}, \mathbf{x}')$ by $(x^{\text{high}}(\mathbf{x}, \mathbf{x}'))_t = \operatorname{argmax}_{x \in \{\mathbf{x}_t, \mathbf{x}'_t\}} u(x)$ for all t and the ‘worst off combination’ path $x^{\text{low}}(\mathbf{x}, \mathbf{x}')$ by $(x^{\text{low}}(\mathbf{x}, \mathbf{x}'))_t = \operatorname{argmin}_{x \in \{\mathbf{x}_t, \mathbf{x}'_t\}} u(x)$ for all t . In every period the consumption path $x^{\text{high}}(\mathbf{x}, \mathbf{x}')$ picks out the better outcome of \mathbf{x} and \mathbf{x}' , while $x^{\text{low}}(\mathbf{x}, \mathbf{x}')$ collects the inferior outcomes. We call a decision maker (weakly)¹²⁰ *intertemporal risk averse* in period t if and only if for all $\mathbf{x}, \mathbf{x}' \in \mathbf{X}$ holds

$$\mathbf{x} \sim \mathbf{x}' \quad \Rightarrow \quad \mathbf{x} \succeq_t \frac{1}{2} x^{\text{high}}(\mathbf{x}, \mathbf{x}') + \frac{1}{2} x^{\text{low}}(\mathbf{x}, \mathbf{x}'). \quad (16.24)$$

The premise states that a decision maker is indifferent between the certain consumption paths \mathbf{x} and \mathbf{x}' . Then, an intertemporal risk averse decision

¹¹⁹For such a definition of more concave see for example ?.

¹²⁰We can define strict intertemporal risk aversion by assuming in addition that there exists some period t^* such that $u(\mathbf{x}_{t^*}) \neq u(\mathbf{x}'_{t^*})$ and requiring a strict preference \succ rather than the weak preference \succeq in equation (16.24).

maker prefers the consumption path x (or equivalently x') with certainty over a lottery that yields with equal probability either a path combining all the best outcomes, or a path combination of all the worst outcomes. Traeger (2010a) and Traeger (2007) spell out the above axiom as well as an alternative formulation in terms of preferences.¹²¹ The same papers show (for the non-stationary and the stationary case respectively) that equation (16.24) holds if and only if the function f in the representation (16.6) is concave¹²².

Another useful interpretation of intertemporal risk aversion is simply as risk aversion with respect to utility gains and losses. This interpretation is true if preferences are represented in the form where the aggregation over time in every recursion is additive as in the form chosen here.¹²³ Then, the utility as expressed by u and by v characterizes how much the decision maker likes a particular outcome x or a particular (degenerate) situation in the future. If the decision maker is intertemporally risk averse, he dislikes taking risk with respect to gains and losses of such utility. Note that in difference to the Arrow-Pratt measure of risk aversion, the function f is always one dimensional and its curvature is well defined. Thus we can use it as a commodity independent risk measure also in a multi-commodity setting. In particular, we can define a measure of relative intertemporal risk aversion

$$\text{RIRA}_t(z) = -\frac{f_t''(z)}{f_t'(z)} z.$$

However, the measure $\text{RIRA}_t(z)$ depends on the choice of zero in the definition of the utility function u . That is equivalent to the fact that the measure of relative Arrow Pratt risk aversion depends on the choice of the zero consumption or wealth level (i.e. when defining a von Neumann-Morgenstern utility index over monetary lotteries the relative risk aversion measure depends on what is considered to be the zero payoff). For the generalized isoelastic model that we discussed in the previous section where $f(z) = z^{\frac{\alpha}{\rho}}$

¹²¹That is, the representation in equation (16.6) is not assumed but axiomatically derived and x^{high} and x^{low} are defined purely in terms of preferences without the use of u . While Traeger (2010a) deals with the general non-stationary setting where the functions f_t and u_t can be time dependent, Traeger (2007) analyzes simplifications that arise for stationary preferences.

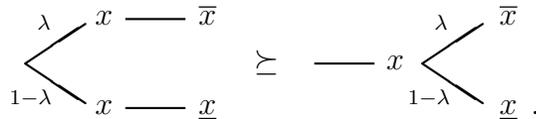
¹²²Respectively strictly concave for strict intertemporal risk aversion as defined in footnote 120.

¹²³See Traeger (2010a) how the same preferences can be represented by making uncertainty aggregation linear at the cost of incorporating a nonlinear aggregation over time.

we find under the assumption $\rho > 0$ ¹²⁴ that $RIRA_t(z) = 1 - \frac{\alpha}{\rho}$ which coincides with the proportionality factor of the new term in the discount rate in equation (16.21).

16.4 Preference for the Timing of Risk Resolution

Let us briefly mention another characteristic of choice that is closely related to intertemporal risk aversion, respectively the disentanglement of Arrow Pratt risk aversion from aversion to intertemporal fluctuation. Let $\lambda(x_t, p_{t+1}) + (1 - \lambda)(x_t, p'_{t+1})$ denote a lottery in period t that delivers x_t in period t and a future described by (the probability tree) p_{t+1} with probability λ and that delivers the outcome x_t in period t and a future described by p'_{t+1} with probability $(1 - \lambda)$. Note that, however the lottery turns out, the decision maker consumes x_t in period t . We would like to compare this first lottery to a second lottery that writes as $(x_t, \lambda p_{t+1} + (1 - \lambda)p'_{t+1})$. In this second lottery the uncertainty over the future does not resolve within period t , but only in period $t + 1$. Otherwise both lotteries coincide. Figure ?? depicts the comparison between two such lotteries for the two period setting. A



decision maker with intertemporally additive expected utility preferences will always be indifferent between the two depicted lotteries. However, in general a decision maker might for example prefer the lottery with the earlier resolution of uncertainty depicted on the left. Note that we are characterizing an intrinsic preference for the timing of uncertainty resolution. There is no choice that the decision maker can take after the resolution of uncertainty. If adapting the optimal choice to resolving information leads to a better expected outcomes in the future causing a preference for an early resolution of uncertainty we say that such preference for early resolution is instrumental (as it is in the end simply a preference for better outcomes). An intrumental

¹²⁴For $\rho < 0$ the function $u(x) = x^\rho$ decreases and f aggregates disutility. In this case we either have to pick a representational equivalent utility function that is strictly increasing or switch around some of the concavity conditions on f . The above discussion assumed that utility is strictly increasing in consumption.

preference for early resolution of uncertainty also prevails in the standard model. The analysis of an intrinsic preference for the timing of uncertainty resolution goes back to ?. New is an interpretation of this finding in terms of intertemporal risk aversion (?). For this purpose we take the special case of the representation (16.5) where u_t is stationary but f_t is not.¹²⁵ Then, it can be shown that the preference depicted in figure ?? or, in general, the preference for early resolution of uncertainty for consumption level x_t as defined by

$$\lambda(x_t, p_{t+1}) + (1 - \lambda)(x_t, p'_{t+1}) \succeq_t (x_t, \lambda p_{t+1} + (1 - \lambda)p'_{t+1}) \quad (16.25)$$

for all p_{t+1}, p'_{t+1} and $\lambda \in [0, 1]$ holds if and only if the expression

$$f_t [(1 - \beta)u(x_t) + \beta f_{t+1}^{-1}(z)] \quad (16.26)$$

is convex in z .¹²⁶ There are two distinct effects driving a preference for an early resolution of uncertainty. For the first one, assume that $u(x_t) = 0$ and $\beta = 1$. Then the condition states that $f_t [f_{t+1}^{-1}(z)] \equiv h$ convex which is equivalent to $f_t = h \circ f_{t+1}$ or f_{t+1} more concave (less convex) than f_t . The intuitive interpretation is that a decision maker prefers an early resolution of uncertainty in the (intrinsic) sense of equation (16.25) if intertemporal risk aversion is increasing over time. The second effect that can drive a preference for an early resolution of uncertainty is observed by setting $f_t = f_{t+1}$. Then, $u(x_t)$ increases the welfare level at which intertemporal risk aversion in period t is evaluated. In addition, β can either increase or decrease the welfare level at which intertemporal risk aversion is evaluated in period t , depending on the value of $u(x_t)$. Assume that the decision maker's preferences exhibit decreasing absolute intertemporal risk aversion. Then, if the effective welfare level at which he evaluates the lottery in period t (incorporating $u(x_t)$ and discounting) is large enough, he will be less risk averse evaluating the lottery in period t than evaluating it in period $t + 1$ with the same function

¹²⁵Axiomatically this representation is obtained by only requiring a stationarity evaluation of certain consumption paths (Traeger 2007).

¹²⁶In general, equation (16.26) is the condition for the infinite time horizon setting corresponding to equation (16.5). But it also holds for the two period setting depicted in figure ?. In the two period setting representation, however, the constant β is not the pure time preference discount factor but the transformation $\beta = \frac{\beta^*}{1 + \beta^*}$ of the pure time preference discount factor β^* . Thus, setting $\beta = 1$ as done below only has an interpretation as a limiting case in the infinite time horizon setting.

f but at a lower effective welfare level. Also in this case the decision maker would prefer an early resolution of uncertainty. In general, any combination of intertemporal risk aversion and timing preference is possible. However, in the case of the generalized isoelastic preferences used by ? an early resolution of uncertainty is preferred if and only if $\alpha > \rho$,¹²⁷ i.e. if the decision maker is intertemporal risk averse respectively more Arrow Pratt risk averse than averse to intertemporal consumption fluctuations. Therefore, it is a widespread believe that a disentanglement of Arrow Pratt risk aversion and attitude with respect to intertemporal substitution is only possible in combination with a non-trivial timing preference. However, ? points out that this is not true in general.

There is an interesting relation between intertemporal risk aversion, the assumptions of indifference to the timing of uncertainty resolution, and discounting. In particular, in an open loop evaluation of the future, intertemporal risk aversion can devalue future utility in a similar way as pure time preference. Moreover, under a stationary evaluation of risk standard assumption lead to the result that a decision maker can only either have a positive rate of pure time preference or be timing indifferent. From a normative perspective the result can be used to derive a zero rate of pure time preference simply from consistency assumptions and widespread axioms on decision making under uncertainty. The underlying interpretation of the result is that intertemporal risk aversion has to be constant in present value terms to satisfy indifference to the resolution of uncertainty while it has to be constant over time in current value terms to satisfy risk stationarity.¹²⁸ However, intertemporal risk aversion can only be constant in both, present value and current value terms, if the pure rate of time preference is zero.

¹²⁷With the exception of a rather special case pointed out in ?.

¹²⁸Note that in the standard model risk stationarity does not add any restriction in addition to certainty stationarity. Only with a nontrivial attitude to intertemporal risk aversion the axiom gets its own bite.

17 Ambiguity

This section discusses decision in the absence of (unique) probabilistic beliefs over the future. We focus on the smooth ambiguity aversion model that is combining aspects of the model discussed in the preceding section with a Bayesian model of decision making. As a simple application we discuss how the setting changes (once more) the social discount rate. We briefly discuss an extension of the model and a sketch of another climate change related application.

17.1 Introduction to ambiguity

Most of the literature on ambiguity is founded on the so called Ellsberg paradox. (?) proposed a thought experiment featuring bets on drawing colored balls from different urns that challenge the standard expected utility model.¹²⁹ A simplified version of this experiment that gets at the intuition¹³⁰ is as follows. You face two different urns. You know that the first urn contains 50 blue and 50 red balls. You know that the second urn only contains blue and red balls, but you do not know their number. You can bet on drawing a blue ball from one of the urns. If the ball is indeed blue you receive \$100, otherwise nothing. Ellsberg's intuition suggested that people would prefer to bet on drawing a blue ball from the first urn. Here, they know the distribution of balls and that the probability of drawing a blue ball equals $\frac{1}{2}$. We don't know anything about the second urn but that it contains blue and red balls. Thus, we can also arrive at a probability of $\frac{1}{2}$ by the principle of insufficient reason, i.e. because we don't have any good reason to assume that blue is more or less probable than red. However, the idea is that people tend to feel different about the two urns and about a situation where the probabilities are known and where they are simply assumed for lack of knowledge. In the expected utility model probabilities are probabilities, so the different

¹²⁹The experiment is conducted to yield a contradiction to Savage's subjective derivation of expected utility theory. While Ellsberg himself didn't actually carry out the experiment, many have done so (in various variants) later on and generally find that some significant fraction of the participants violate the expected utility hypothesis. More than that, some well known economists were unwilling to change their choices also after they learned that it would violate expected utility.

¹³⁰However, this simplified experiment as is would not suffice to prove the contradiction to Savage's axiomatization of expected utility.

urns cannot be distinguished and neither can the two bets be distinguished. In consequence, a strict preferences for betting on one of the urns cannot be represented.¹³¹ We feel that a similar, possibly even more convincing, situation of ambiguity arises in the environmental economic context. For example, specifying a (unique) probability for the climate sensitivity, or for a possible collapse of the gulf stream seems to much harder than in the case of a coin toss. Probably few people would contradict that there are qualitatively very different types of probabilities. For a reason the literature started to distinguish a long time ago objective probabilities from subjective probabilities, where the first refer to probabilities that are derived mostly from symmetry reasonings of sufficiently good frequency data. Objective probabilities are often also considered to be known probabilities, even though that is a somewhat intricate statement.¹³² However, the crucial distinction in economic models of ambiguity is that individuals also evaluate expected outcomes differently when they are attached to these different types of probabilities. One way to approach this setting is by distinguishing objective and subjective lotteries. Essentially, that is what Klibanoff et al. (2005) did. At the end of the section we also discuss the idea that there might be more than just two different classes of probabilities (objective and subjective). Before we introduce the corresponding smooth ambiguity model, we give a brief mentioning of some of the other more widespread approaches to model ambiguity.

The decision-theoretic literature has developed different concepts to deal with situations where uncertainty is not sufficiently captured by unique probability measures. Apart from tagging these situations with the word ambiguity, they are also referred to as situations of Knightian uncertainty, hard or deep uncertainty and are often contrasted with situations of risk (used to denote the standard setting). One way to characterize non-risk uncertainty is by extending the concept of probabilities to a form of more general set function called capacities. These set functions weigh possible events but are not necessarily additive (in the union of disjoint events). In analogy to expected utility theory that aggregates utility over states respectively events we can define a way to aggregate utility over capacities. The according procedure

¹³¹The real Ellsberg paradox setting asks for slightly more sophisticated bets so that we can as well rule out that the individuals place differing probabilities on the two urns, at least one of which would not be 50 : 50.

¹³²A better characterization might be intersubjectively agreed probabilities. On an individual level, a person might be completely convinced that a particular probability is known to him and still be off. Probably most wouldn't classify such a probability as objective.

is known as Choquet integration. A second way is to define an evaluation functional that expresses beliefs in form of sets of probability distributions rather than unique probability distributions. The first and simplest such representation goes back to Gilboa and Schmeidler (1989). Here a decision maker evaluates a scenario by taking expected values with respect to every probability distribution deemed possible and then identifies the scenario with the minimal expected value in this set.¹³³ The most general representation of this type is given by Gilboa and Schmeidler (1989) and, in an intertemporal framework by Gilboa and Schmeidler (1991). There are several equivalence results between the Choquet approach and that of multiple priors as well as rank dependent utility theory where a decision maker uses distorted probabilities in an expected utility approach increasing the weights given to small probability events. Axiomatically, all of these models relax the independence axiom in one way or the other.

17.2 The smooth ambiguity model

We will focus our discussion on a recent representation result by Klibanoff et al. (2005), and, in an intertemporal setting Klibanoff et al. (2009). The authors model ambiguity as second order probability distributions that is, probabilities over probabilities. The model almost resembles a standard Bayesian decision model. The likelihood function, which delivers a probability distribution given some (unknown) parameter, is identified with objective or first order risk. The Bayesian prior, which gives the distribution over the unknown parameter of the likelihood function, is referred to as subjective probability or second order probability. In difference to the standard Bayesian model, Klibanoff et al. (2005) introduce a separate characterization of ambiguity attitude that comes into play when aggregating over second order lotteries. Translated into notation similar to our recursive objective function in the previous section we can define the welfare function as

$$V(x_t, I_t) = u(x_t) + \beta \Phi^{-1} \left\{ \int_{\Theta} \Phi [E_{\Pi_{\theta}(x_{t+1}|x_t, I_t)} V(x_{t+1}, I_{t+1})] d\mu(\theta|x_t, I_t) \right\}.$$

I_t characterizes again the information that determines the (first and second order) probabilities over the future. Here Π denotes first order or ‘objective’ probabilities. However, these are not known uniquely and depend on

¹³³ Gilboa and Schmeidler (1989) give conditions under which this approach is equivalent to what is known as robust control or model uncertainty.

a parameter θ that is unknown and subjective. The probability measure μ denotes the prior over the parameter $\theta \in \Theta$.¹³⁴ The utility function u corresponds to the utility function of the standard model. It jointly captures aversion to intertemporal substitutability and ‘objective’ or first order risk. The function Φ captures additional aversion with respect to second order uncertainty which is called *ambiguity aversion*. Note that for Φ linear the model collapses to the standard Bayesian model. If the objective uncertainty measure Π is degenerate, there is a close formal similarity to the model of intertemporal risk aversion. The welfare function ?? can be applied in the infinite time horizon setting employing dynamic programming. The application is similar to the model discussed in section 16.2, only that in general we want to update the Bayesian prior as we move into the future.

We will limit our attention to a simple ‘certain \times uncertain’ two period application of the model that brings out the characteristics of the utility aggregation without complicating the setting by introducing learning. Moreover, we will return to a one commodity setting and assume once more an isoelastic utility specification $u(x) = x^\rho$. Moreover we assume constant relative ambiguity aversion corresponding to $\Phi(z) = (\rho z)^\varphi$. Then our evaluation functional in the first period can be written as

$$V(x_1, I_1) = \frac{x_1^\rho}{\rho} + \beta \frac{1}{\rho} \left\{ \int_{\Theta} [\mathbb{E}_{\Pi_\theta(x_2|x_1, I_1)} x_2^\rho]^\varphi d\mu(\theta|x_1, I_1) \right\}^{\frac{1}{\varphi}}. \quad (17.1)$$

We can define a coefficient of relative ambiguity aversion similarly to the standard coefficient of Arrow Pratt risk aversion and the coefficient of relative intertemporal risk aversion. Here (as in the case of relative intertemporal risk aversion), we have to be careful that the argument of our risk aggregation can potentially be negative. In particular, for $\rho < 0$ the chosen isoelastic utility function measures welfare in negative units. That was the reason why we have included the parameter ρ in the definition of $\Phi(z) = (\rho z)^\varphi$. That way the argument of the isoelastic aggregator is positive and the aggregation is well defined. However, the generalized mean now aggregates over the negative of what really makes up welfare. Thus, if we have a parameterization $\varphi < 1$ we now generate a lower than expected outcome as the generalized mean in positively measured units, but that turns into a lower than expected welfare when translated back into welfare units (as done by $\frac{1}{\rho}$ in front of the integral

¹³⁴In (Klibanoff et al. 2009) axiomatization of the model the parameter space Θ is finite.

that stems from Φ^{-1}). Thus, in general we have to define the coefficient of relative ambiguity aversion as

$$\text{RAA} = \frac{\Phi''(z)}{\Phi'(z)} |z| \begin{cases} 1 - \varphi & \text{if } \rho > 0 \\ \varphi - 1 & \text{if } \rho < 0. \end{cases}$$

17.3 The social discount rate under ambiguity

We employ the extended welfare framework represented in equation (17.1) to analyze how the social discount rate changes in a setting with ambiguity. Our setting will be similar to the one employed to derive the stochastic Ramsey equation in section 14.5. Consumption growth of the economy is stochastic and governed by the growth rate $g = \ln \frac{x_2}{x_1}$ that is assumed to be normally distributed with $g \sim N(\theta, \sigma^2)$. Introducing second order uncertainty we assume that σ^2 is known while $\theta \sim N(\mu, \tau^2)$ is itself random and normally distributed.

We are interested in the rate at which consumption is discounted over the future. We obtain this rate by analyzing the ratio of a small certain consumption gain dx_2 in the second period over a small consumption loss dx_1 in the first period that leaves the decision maker indifferent:

$$dV(x_1, I_1) = x_1^{\rho-1} dx_1 + \beta \frac{1}{\rho} \frac{1}{\varphi} \left\{ \int_{\Theta} [\mathbb{E}_{\Pi_{\theta}(y|x_1, I_1)} x_2^{\rho}]^{\varphi} d\mu(\theta|x_1, I_1) \right\}^{\frac{1}{\varphi}-1} \\ \int_{\Theta} \varphi [\mathbb{E}_{\Pi_{\theta}(y|x_1, I_1)} x_2^{\rho}]^{\varphi-1} \mathbb{E}_{\Pi_{\theta}(y|x_1, I_1)} \rho x_2^{\rho-1} dx_2 d\mu(\theta|x_1, I_1) \stackrel{!}{=} 0$$

$$\begin{aligned}
\Rightarrow x_1^{\rho-1} dx_1 &= -\beta \left\{ \int_{\Theta} [\mathbb{E}_{\Pi_{\theta}(y|x_1, I_1)} x_2^{\rho} dx_2]^{\varphi} d\mu(\theta|x_1, I_1) \right\}^{\frac{1}{\varphi}-1} \\
&\quad \int_{\Theta} [\mathbb{E}_{\Pi_{\theta}(y|x_1, I_1)} x_2^{\rho}]^{\varphi-1} \mathbb{E}_{\Pi_{\theta}(y|x_1, I_1)} x_2^{\rho-1} dx_2 d\mu(\theta|x_1, I_1) \\
\Rightarrow \frac{dx_1}{dx_2} &= -\beta \left\{ \int_{\Theta} \left[\mathbb{E}_{\Pi_{\theta}(y|x_1, I_1)} \left(\frac{x_2}{x_1} \right)^{\rho} \right]^{\varphi} d\mu(\theta|x_1, I_1) \right\}^{\frac{1}{\varphi}-1} \\
&\quad \int_{\Theta} \left[\mathbb{E}_{\Pi_{\theta}(y|x_1, I_1)} \left(\frac{x_2}{x_1} \right)^{\rho} \right]^{\varphi-1} \mathbb{E}_{\Pi_{\theta}(y|x_1, I_1)} \left(\frac{x_2}{x_1} \right)^{\rho-1} d\mu(\theta|x_1, I_1) \\
\Rightarrow \frac{dx_1}{dx_2} &= -\beta \left\{ \int_{\Theta} \left[\mathbb{E}_{\Pi_{\theta}(y|x_1, I_1)} e^{\rho \ln \frac{x_2}{x_1}} \right]^{\varphi} d\mu(\theta|x_1, I_1) \right\}^{\frac{1}{\varphi}-1} \\
&\quad \int_{\Theta} \left[\mathbb{E}_{\Pi_{\theta}(y|x_1, I_1)} e^{\rho \ln \frac{x_2}{x_1}} \right]^{\varphi-1} \mathbb{E}_{\Pi_{\theta}(y|x_1, I_1)} e^{(\rho-1) \ln \frac{x_2}{x_1}} d\mu(\theta|x_1, I_1) \\
\Rightarrow \frac{dx_1}{dx_2} &= -\beta \left\{ \int_{\Theta} \left[e^{\rho\theta + \rho^2 \frac{\sigma^2}{2}} \right]^{\varphi} d\mu(\theta|x_1, I_1) \right\}^{\frac{1}{\varphi}-1} \\
&\quad \int_{\Theta} \left[e^{\rho\theta + \rho^2 \frac{\sigma^2}{2}} \right]^{\varphi-1} e^{(\rho-1)\theta + (\rho-1)^2 \frac{\sigma^2}{2}} d\mu(\theta|x_1, I_1) \\
\Rightarrow \frac{dx_1}{dx_2} &= -\beta e^{[(1-\varphi)\rho^2 + (\varphi-1)\rho^2 + (\rho-1)^2] \frac{\sigma^2}{2}} \left\{ \int_{\Theta} [e^{\rho\theta\varphi}] d\mu(\theta|x_1, I_1) \right\}^{\frac{1}{\varphi}-1} \\
&\quad \int_{\Theta} [e^{\rho\theta(\varphi-1)}] e^{(\rho-1)\theta} d\mu(\theta|x_1, I_1) \\
\Rightarrow \frac{dx_1}{dx_2} &= -\beta e^{(\rho-1)^2 \frac{\sigma^2}{2}} \left\{ \int_{\Theta} [e^{\rho\theta\varphi}] d\mu(\theta|x_1, I_1) \right\}^{\frac{1}{\varphi}-1} \\
&\quad \int_{\Theta} e^{\theta[\rho\varphi - \rho + \rho - 1]} d\mu(\theta|x_1, I_1)
\end{aligned}$$

$$\begin{aligned}
\Rightarrow \frac{dx_1}{dx_2} &= -\beta e^{(\rho-1)^2 \frac{\sigma^2}{2}} \left\{ e^{\rho\varphi\mu + \rho^2\varphi^2 \frac{\tau^2}{2}} \right\}^{\frac{1}{\varphi}-1} \\
&\quad \left[e^{(\rho\varphi-1)\mu + (\rho\varphi-1)^2 \frac{\tau^2}{2}} \right] \\
\Rightarrow \frac{dx_1}{dx_2} &= -\beta e^{(\rho-1)^2 \frac{\sigma^2}{2}} e^{[\rho(1-\varphi) + \rho\varphi-1]\mu} \\
&\quad e^{[\rho^2\varphi(1-\varphi) + \rho^2\varphi^2 - 2\rho\varphi + 1] \frac{\tau^2}{2}} \\
\Rightarrow \frac{dx_1}{dx_2} &= -\beta e^{(\rho-1)^2 \frac{\sigma^2}{2}} e^{(\rho-1)\mu} e^{[\rho^2\varphi - 2\rho\varphi + \varphi - \varphi + 1] \frac{\tau^2}{2}} \\
\Rightarrow \frac{dx_1}{dx_2} &= -\beta e^{(\rho-1)^2 \frac{\sigma^2}{2}} e^{(\rho-1)\mu} e^{[\varphi(\rho-1)^2 + (1-\varphi)] \frac{\tau^2}{2}} \\
\Rightarrow \frac{dx_1}{dx_2} &= -\beta e^{-(1-\rho)\mu} e^{(1-\rho)^2 \frac{\sigma^2 + \tau^2}{2}} e^{[-(1-\varphi)(1-\rho)^2 + (1-\varphi)] \frac{\tau^2}{2}}.
\end{aligned}$$

We want to express this welfare neutral trade off between a small positive amount dx_2 and a small negative amount dx_1 in terms of the corresponding (risk free) consumption discount rate $r = \ln \frac{dx_2}{-dx_1} \Big|_{\bar{V}}$. Moreover we define the pure rate of time preference $\delta = -\ln \beta$ and the consumption elasticity of marginal utility characterizing aversion to intertemporal substitution as $\eta = 1 - \rho$. Then we obtain

$$r = \delta + \eta\mu - \eta^2 \frac{\sigma^2 + \tau^2}{2} - (1 - \varphi)(1 - \eta^2) \frac{\tau^2}{2}.$$

Finally, note that $1 - \eta^2 > 0 \Leftrightarrow \rho > 0$ which implies $1 - \varphi = \text{RAA}$ and similarly $1 - \eta^2 < 0 \Leftrightarrow \rho < 0$ implies $1 - \varphi = -\text{RAA}$. Thus $(1 - \varphi)(1 - \eta^2) = \text{RAA} |1 - \eta^2|$. In consequence the certainty equivalent social discount rate in the above 2 period setting with constant relative ambiguity aversion and isoelastic utility is

$$r = \delta + \eta\mu - \eta^2 \frac{\sigma^2 + \tau^2}{2} - \text{RAA} |1 - \eta^2| \frac{\tau^2}{2}. \quad (17.2)$$

The first two terms reflect the discount rate in the standard Ramsey equation under certainty. The third is term $-\eta^2 \frac{\sigma^2 + \tau^2}{2}$ is only a slight modification of the well known extension to risk. On top of the immediate variance σ^2 of the growth rate itself we have in addition the variance τ^2 . It is a straightforward

consequence of making the growth process more uncertain by introducing a prior (second order uncertainty) over some parameter of the growth process. In the case of the normal distributions and second order uncertainty over expected growth adopted here, the variance simply adds up. The fourth term, finally, is the interesting new contribution. It reduces the social discount rate proportionally to the coefficient of relative ambiguity aversion and the variance of the subjective prior. In consequence, a decision maker in a situation where he faces ambiguity with respect to baseline growth is willing to invest into a certain projects with relatively lower productivity than is a decision maker who just faces (first order) risk or is ambiguity neutral.

(Weitzman 2009) famously used a Bayesian decision framework to discuss the consequences of structural uncertainty, fat tails, and catastrophic climate events. It is insightful to discuss his findings in the above framework. Weitzman (2009) does not employ a decision theoretic framework of ambiguity. So in discussing the social discount rate and the willingness to pay for the future Weitzman relies on the first three terms in equation (17.2). The only difference between these remaining terms and the standard stochastic Ramsey equation is the additional variance τ^2 in the third term on the right hand side (standard risk term). Note more precisely Weitzman (2009) assumes that the variance of the first order distribution rather than its expected value is unknown, an uncertainty he loosely relates to climate sensitivity. From our discussion in section 14.5 we know that the contribution of the stochastic term $\eta^2 \frac{\sigma^2}{2}$ is generally negligible. Instead of a doubling, a factor of 10 – 100 is needed. Thus, in order to make uncertainty play a serious role in the standard model Weitzman has to increase the variance of the prior significantly. Effectively, this is what he does in deriving what Weitzman calls a dismal theorem. He introduces a fat tailed (improper) prior whose moments do not exist. Consequently, the risk-free social discount rate in equation (14.5) goes to minus infinity implying an infinite willingness to transfer (certain) consumption into the future. Weitzman limits this willingness by the value of a (or society's) statistical life.¹³⁵ Instead of tickling infinity our derivation here which is based on Traeger (2008) introduces ambiguity aversion, i.e. the term $\text{RAA} |1 - \eta^2| \frac{\tau^2}{2}$, into social discounting, reflecting experimental ev-

¹³⁵Note that Weitzman (2009) puts the prior on the variance σ rather than on the expected value of growth. He loosely relates the uncertainty to climate sensitivity. The above is a significantly simplified, but insightful, perspective on Weitzman's approach – abstracting from learning.

idence that economic agents tend to be more afraid of unknown probabilities than they are of known probabilities.

17.4 Epilogue

In the introduction we motivated ambiguity by pointing to the fact that there are different types of probabilities. (Klibanoff et al. 2009) setting can be interpreted as world with two different types of probabilities either objective, or subjective. They use the Bayesian framework and identify the objective probabilities with the likelihood function and subjective probabilities with the prior. Now if we start to distinguish the quality of or confidence into our priors we probably find a much wider range of classifications than two. Moreover, it is not a priori clear that, within a period, we always face more subjective lotteries (prior) over less subjective lotteries (likelihood function). For example, we might have a pretty good guess of the probability distribution of atmospheric global mean temperatures over the next decades, but we only have a very crude guess of the probability that a particular temperature level causes us to cross a tipping point in the climate system, i.e. a point where we change the climate system significantly and irreversibly, like with stopping the gulf stream, or disintegrating major ice sheets. If we model such an ambiguous regime shift, the ambiguous probability of the regime shift is actually a function of the better known temperature distribution. It is more convenient to turn the order of subjective and objective probabilities around in this situation.¹³⁶ Finally, we have motivated ambiguity by means of the Ellsberg paradox and the fact that people behave in a way that cannot be captured within the standard model. However, is it legitimate to employ such a behaviorally founded model to address a question like the social discount rate and a cost benefit analysis of climate change?

Traeger 2010 Traeger (2010*b*) addresses these questions. First, the paper extends the idea of ambiguity aversion to a concept of aversion to the subjectivity of belief with an arbitrary number of subjectivity classes of probabilities. Second, the paper detaches the degree of subjectivity of a lottery from there order (first or second order etc. lottery). Third, the paper's setting takes place in only a minimalistic deviation from the standard von Neumann-Morgenstern axioms. The basic idea is that probabilities (or lot-

¹³⁶See Lemoine & Traeger 2009 ?.

teries) obtain a subjectivity (or confidence) label. Then we adapt the von Neumann-Morgenstern axioms so that they only mix lotteries of the same degree of subjectivity. The result is that our measure of risk aversion becomes a function of the degree of subjectivity (or confidence) of the lottery. While there isn't an immediate normative reason that the degrees of risk aversion should depend on the degree of confidence into the lottery, the framework also shows that the von Neumann-Morgenstern axioms don't really tell us that we should have the same degree of risk aversion with respect to different lotteries either, as soon as we acknowledge that there are different types of probabilities out there. Finally, we pointed out that the smooth ambiguity framework by Klibanoff et al. (2009) implicitly assumes that Arrow-Pratt risk aversion (with respect to objective lotteries) is equivalent to the (inverse of the) intertemporal elasticity of substitution. The cited paper also relaxes this assumption.

References

- Andreoni, J. (1989), 'Giving with impure altruism: applications to charity and ricardian equivalence', *Journal of Political Economy* **97**, 1447–1458.
- Basal, R., Kiku, D. & Yaron, A. (2010), '*', *American Economic Review: Papers & Proceedings* **100**, 542–546.
- Basal, R. & Yaron, A. (2004), 'Risks for the long run: A potential resolution of asset pricing puzzles', *The Journal of Finance* **59**(4), 1481–509.
- Blanchard, O. J. (1985), 'Debts, deficits and finite horizons', *Journal of Political Economy* **93**, 223–247.
- Brock, W. & Starrett, D. (2003), 'Managing systems with non-convex positive feedback', *Environmental and Resource Economics* **26**(4), 575–602.
- Broome, J. (1992), *Counting the Cost of Global Warming*, White Horse Press, Cambridge.
- Calvo, G. & Obstfeld, M. (1988), 'Optimal time-consistent fiscal policy with finite lifetimes', *Econometrica* **56**, 411–32.

- Cropper, M. L., Ayded, S. K. & Portney, P. R. (1994), 'Preferences for life-saving programs: How the public discounts time and age', *Journal of Risk and Uncertainty* **8**, 243–266.
- Dasgupta, P. & Mäler, K.-G. (2003), 'The economics of non-convex ecosystems: Introduction', *Environmental and Resource Economics* **26**(4), 499–525.
- Ekeland, I. & Lazrak, A. (2010), 'The golden rule when preferences are time inconsistent', *Mathematical and Financial Economics* **4**(1).
- Freeman, M. C. (2010), 'Yes, we should discount the far-distant future at its lowest possible rate: A resolution of the weitzman-gollier puzzle', *The Open-Access, Open-Assessment E-Journal* **4**.
- Gierlinger, J. & Gollier, C. (2008), 'Socially efficient discounting under ambiguity aversion', *Working Paper* .
- Gollier, C. (2002), 'Discounting an uncertain future', *Journal of Public Economics* **85**, 149–166.
- Gollier, C. (2004), 'Maximizing the expected net future value as an alternative strategy to gamma discounting', *Finance Research Letters* **1**(2), 85–89.
- Gollier, C. (2009), 'Should we discount the far-distant future at its lowest possible rate?', *The Open-Access, Open-Assessment E-Journal* **3**.
- Gollier, C. & Weitzman, M. L. (2010), 'How should the distant future be discounted when discount rates are uncertain?', *Economic Letters* **107**, 350–353.
- Harris, C. & Laibson, D. (2001), 'Dynamic choices of hyperbolic consumers', *Econometrica* **69**(5), 935–957.
- Harrod, S. R. F. (1948), *Towards a dynamic economics*, Macmillan, London.
- Hoel, M. & Sterner, T. (2007), 'Discounting and relative prices', *Climatic Change* **84**, 265–280.
- Karp, L. (2007), 'Non-constant discounting in continuous time', *Journal of Economic Theory* **132**, 557 – 568.

- Karp, L. (2009), 'Sacrifice, discounting and climate policy: five questions', Giannini Foundation working paper.
- Klibanoff, P., Marinacci, M. & Mukerji, S. (2005), 'A smooth model of decision making under ambiguity', *Econometrica* **73**(6), 1849–1892.
- Klibanoff, P., Marinacci, M. & Mukerji, S. (2009), 'Recursive smooth ambiguity preferences', *Journal of Economic Theory* **144**, 930–976.
- Kocherlakota, N. R. (1996), 'The equity premium: It's still a puzzle', *Journal of Economic Literature* **34**, 42–71.
- Koopmans, T. C. (1963), On the concept of optimal economic growth, Cowles Foundation Discussion Papers 163, Cowles Foundation, Yale University.
- Kossioris, G., Plexousakis, M., Xepapadeas, T. & de Zeeuw K.-G. Mäler, A. (2008), 'Feedback nash equilibria for non-linear differential games in pollution control', *Journal of Economic Dynamics and Control* **32**, 1312–1331.
- Laibson, D. (1997), 'Golden eggs and hyperbolic discounting', *The Quarterly Journal of Economics* **112**(2), 443–477.
- Millner, A. (2011), 'On welfare frameworks and catastrophic climate risks', Working Paper.
- Mäler, K.-G., Xepapadeas, A. & de Zeeuw, A. (2003), 'The economics of shallow lakes', *Environmental and Resource Economics* **26**(4), 603–624.
- Nordhaus, W. (2008), *A Question of Balance: Economic Modeling of Global Warming*, Yale University Press, New Haven. Online preprint: A Question of Balance: Weighing the Options on Global Warming Policies.
- Nordhaus, W. D. (2007), 'A review of the Stern review on the economics of climate change', *Journal of Economic Literature* **45**(3), 686–702.
- Pigou, A. C. (1932), *The economics of welfare*, 4th edn, Macmillan, London.
- Pindyck, R. S. (1980), 'Uncertainty in exhaustible resource markets', *The Journal of Political Economy* **88**(6), 1203–1225.

- Ramsey, F. P. (1928), 'A mathematical theory of saving', *The Economic Journal* **38**(152), 543–559.
- Ray, D. (1987), 'Nonpaternalistic intergenerational altruism', *Journal of Economic Theory* (41), 112–132.
- Schneider, M., Traeger, C. P. & Winkler, R. (2012), 'Trading off generations: Infinitely lived agent versus olig', *CESifo Working Paper No 3743*.
- Skiba, A. K. (1978), 'Optimal growth with a convex-concave production function', *Econometrica* **46**(3), 527–539*.
- Solow, R. M. (1974), 'The economics of resources of the resources of economics', *American Economic Review* **64**(2), 1–14.
- Stern, N., ed. (2007), *The Economics of Climate Change: The Stern Review*, Cambridge University Press, Cambridge.
- Traeger, C. (2007), Intertemporal risk aversion, stationarity and the rate of discount. Working Paper.
- Traeger, C. (2010a), 'Wouldn't it be nice to know whether Robinson is risk averse?'. CUDARE Working Paper No. 1102.
- Traeger, C. P. (2008), 'Why uncertainty matters - discounting under intertemporal risk aversion and ambiguity', *CUDARE Working Paper No 1092*.
- Traeger, C. P. (2010b), 'Subjective risk, confidence, and ambiguity', *CUDARE Working Paper No 1103*.
- Traeger, C. P. (2011a), 'Discounting and confidence', *CUDARE Working Paper No 1117*.
- Traeger, C. P. (2011b), 'Sustainability, limited substitutability and non-constant social discount rates', *Journal of Environmental Economics and Management* **62**(2), 215–228.
- Traeger, C. P. (2012), 'What's the rate? Disentangling the Weitzman and the Gollier effect', *CUDARE Working Paper No 1121*.

- Vissing-Jørgensen, A. & Attanasio, O. P. (2003), 'Stock-market participation, intertemporal substitution, and risk-aversion', *The American Economic Review* **93**(2), 383–391.
- von Neumann, J. & Morgenstern, O. (1944), *Theory of Games and Economic Behaviour*, Princeton University Press, Princeton.
- Wagener, F. (2003), 'Skiba points and heteroclinic bifurcations, with applications to the shallow lake system', *Journal of Economic Dynamics and Control* **27**, 1533–1561.
- Weitzman, M. (2009), 'On modeling and interpreting the economics of catastrophic climate change', *The Review of Economics and Statistics* **91**(1), 1–19. 06.
- Weitzman, M. L. (1998), 'Why the far-distant future should be discounted at its lowest possible rate', *Journal of Environmental Economics and Management* **36**, 201–208.
- Yaari, M. E. (1965), 'Uncertain lifetime, life insurance and the Theory of the Consumer', *Review of Economic Studies* **32**, 137–150.

Appendix

A Formal Background on Dynamic Programming in Discrete Time

This chapter provides a formal background for discrete time dynamic programming. We start by explaining the equivalence between the dynamic programming equation and an agent's utility maximization over an infinite consumption stream. We then proceed to discuss existence and characterize the solutions to the dynamic programming problem. The chapter draws heavily on chapters 3 (Mathematical Preliminaries) and 4 (Dynamic Programming under Certainty) of (?) *Recursive Methods in Economic Dynamics*.

A.1 Two Statements of the Dynamic Optimization Problem

Consider an agent who maximizes utility over an infinite consumption stream. In every period his choice is to either consume or to invest. If he invests k_t units of the generic consumption good in period t he receives $k_{t+1} = f(k_t)$ units in the next period. The problem's state variable is k_t and consumption c_t is the control variable. Our first characterization of the dynamic optimization problem is as a maximization of the infinite sum of utility payoffs subject to the equation of motion for the stock k_t :

$$\begin{aligned} \max_{\{(c_t, k_t)\}_{t=0}^{\infty}} & \sum_{t=0}^{\infty} \beta^t U(c_t) \\ \text{s.t.} & \quad c_t + k_{t+1} \leq f(k_t), \\ & \quad c_t, k_t \geq 0 \quad \forall t \in \mathbb{N}, \\ & \quad k_0 \text{ given.} \end{aligned}$$

Eliminating the control the problem can be rewritten in terms of the state variables as

$$\begin{aligned} \max_{\{(k_t)\}_{t=0}^{\infty}} \quad & \sum_{t=0}^{\infty} \beta^t U(f(k_t) - k_{t+1}) \\ \text{s.t.} \quad & f(k_t) \geq k_{t+1} \geq 0 \quad t \in \mathbb{N}, \\ & k_0 \text{ given,} \end{aligned}$$

or, more generally, as (SP)

$$\begin{aligned} \sup_{\{(k_t)\}_{t=0}^{\infty}} \quad & \sum_{t=0}^{\infty} \beta^t F(k_t, k_{t+1}) \\ \text{s.t.} \quad & k_{t+1} \in \Gamma(k_t) \quad t \in \mathbb{N} \\ & k_0 \in K \text{ given,} \end{aligned} \tag{A.1}$$

where $\Gamma(k_t)$ denotes the feasibility of the next period state given k_t .

The dynamic programming approach restates the optimization problem in an alternative formulation. It builds on a principle formulated by Richard Bellman who recognized that an optimal sequence of choices satisfies: whenever the initial choice based on the initial state is optimal the remaining decisions must constitute an optimal sequence of choices with regard to the state variables that result from the first decision. This assertion translates formally into the dynamic programming or Bellman equation

$$\begin{aligned} v(k) &= \max_{c,y} [U(c) + \beta v(y)] \\ \text{s.t.} \quad & c + y \leq f(k), \\ & c, y \geq 0. \end{aligned} \tag{A.2}$$

For the general case we write the dynamic programming equation as (FE)

$$v(k) = \sup_{y \in \Gamma(k)} [F(k, y) + \beta v(y)] \quad \forall k \in K. \tag{A.3}$$

Facing a stationary utility function and an infinite time horizon the trade-off between consuming and augmenting the state variable for the continuation

path is identical in every period so that we dropped the time index. Here k denotes the current stock and y denotes the stock resulting for the next period. Equations (A.2) and (A.3) are functional equations. We will establish that the solution maps a given state k into the value of the best achievable infinite consumption stream. The subsequent section establishes equivalence between the two optimization approaches (SP) and (FE).

A.2 Establishing Equivalence

We now state the precise assumptions that result in the equivalence of the optimization problems (SP) formulated in equation (A.1) and (FE) formulated in equation (A.3).¹³⁷ The first assumption requires merely that the set of feasible controls is non-empty.

Assumption 3: $\Gamma(k)$ is nonempty for all $k \in K$.

The second assumption requires that the discounted sum over the instantaneous payoff $F(k_t, k_{t+1})$ converges for all feasible policy paths.

Assumption 4: For all $k_0 \in K$ and sequences $\{k_t\}_{t=0}^\infty$ satisfying $k_{t+1} \in \Gamma(k_t) \forall t \in \mathbb{N}$ the limit $\lim_{n \rightarrow \infty} \sum_{t=0}^n \beta^t F(k_t, k_{t+1})$ exists (it may be plus or minus infinity).

We refer to ? for different ways to ensure that assumption 4 holds. The assumptions guarantee that the solution to the maximization problem (SP) are also solutions to the dynamic programming problem (FE).

Theorem 1 (SP \rightarrow FE, value function): Let K, Γ, F, β satisfy assumption 3 and 4. Then the function $v^*(k_0)$ solving problem (SP) in equation (A.1) satisfies:

1. If $|v^*(k_0)| < \infty$ then

$$v^*(k_0) \geq F(k_0, y) + \beta v^*(y) \quad \forall y \in \Gamma(k_0)$$

and for any $\epsilon > 0$

$$v^*(k_0) \leq F(k_0, y) + \beta v^*(y) + \epsilon \quad \text{for some } y \in \Gamma(k_0).$$

¹³⁷See chapter 4.1 of ? for proofs.

2. If $v^*(k_0) = \infty$ then there exists a sequence y^n in $\Gamma(k_0)$ such that

$$\lim_{n \rightarrow \infty} F(k_0, y^n) + \beta v^*(y^n) = \infty.$$

3. If $v^*(k_0) = -\infty$ then

$$F(k_0, y) + \beta v^*(y) = -\infty \quad \forall y \in \Gamma(k_0).$$

We are most interested in the first case where the maximized welfare from future consumption is finite. In this case, the theorem establishes that the maximized sum of welfare from the infinite consumption stream in problem (SP) is arbitrarily close to satisfying the functional equation (FE). Note that we have not established, yet, that the problem (FE) has a solution and when this solution would be unique. If the problem function equation (FE) has a solution, then the subsequent theorem states that it characterizes the optimal welfare from problem (SP) under one additional condition.

Theorem 2 (FE \rightarrow SP, value function): Let K, Γ, F, β satisfy assumptions 3-4. If the function v is a solution to problem (FE) in (A.3) and satisfies

$$\lim_{t \rightarrow \infty} \beta^t v(k_t) = 0 \text{ for all feasible sequences } \{k_t\}_0^\infty, \quad (\text{A.4})$$

then v also solves problem (SP) formulated in equation (A.1).

We note that condition (A.4) has to hold for all feasible sequences. It plays a similar role as the transversality condition in Pontryagin's maximum principle (chapter 7) and, e.g., rules out solutions to (FE) that go along with an ever increasing borrowing against the future.¹³⁸

Finally we are interested in the equivalence of the optimal policy plans $\{k^*\}_0^\infty$ solving the two different approaches.

^{138?} give an example of such a solution that satisfies the functional equation (A.3) but violates the boundary (or boundedness) condition (A.4) on page 74 et seq. Exercise 4.3 on page 75 gives an alternative theorem establishing equivalence if equation (A.4) is not satisfied for all feasible paths.

Theorem 3 (SP \rightarrow FE, policy sequence): Let K, Γ, F, β satisfy assumptions 3 and 4. Let $\{k_t^*\}_0^\infty$ be a feasible plan that attains the supremum in (A.3) for initial state k_0 . Then

$$v^*(k_t^*) = F(k_t^*, k_{t+1}^*) + \beta v^*(k_{t+1}^*), \quad t \in \mathbb{N}.$$

The result states that every optimal plan $\{k_t^*\}_0^\infty$ is generated¹³⁹ by the the optimal policy correspondence $G^*(k)$ defined as

$$G^*(k) = \{y \in \Gamma(k) : v^*(k) = F(k, y) + \beta v^*(y)\}.$$

The “inverse statement” is that any plan $\{k_t^*\}_0^\infty$ generated by the the optimal policy correspondence G^* is an optimal plan (solving SP). The next theorem establishes that this statement holds under one additional condition (equation A.5).

Theorem 4 (FE \rightarrow SP, policy sequence): Let K, Γ, F, β satisfy assumptions 3-4. Let $\{k_t^*\}_0^\infty$ be a feasible plan starting with initial state k_0 satisfying

$$v^*(k_t^*) = F(k_t^*, k_{t+1}^*) + \beta v^*(k_{t+1}^*), \quad t \in \mathbb{N}$$

and

$$\limsup_{t \rightarrow \infty} \beta^t v^*(k_t^*) \leq 0. \quad (\text{A.5})$$

Then $\{k_t^*\}_0^\infty$ attains the supremum in problem (A.1) for initial state k_0 .

A.3 Banach's Contraction Mapping

This section introduces Banach's fixed point theorem and closely related mathematical results that underlie the proofs establishing existence and uniqueness of the value function v solving problem (FE) formulated in equation

¹³⁹The sequence (k_0, k_1, \dots) is said to be generated from k_0 by a policy correspondence G if it satisfies $k_{t+1} \in G(k_t) \forall t \in \mathbb{N}$.

(A.3). The contraction mapping approach not only helps to prove characteristics of the solution to the dynamic programming problem (section A.5), but it also underlies one of the widely used numeric solution methods introduced in chapter 6.

Recall that a map takes elements from one space into another space. A contraction map takes elements from the space into itself and at the same time ‘shrinks’ the distance between any two elements when mapping them to their counterparts. In order to define such a concept of “shrinking distance” we must be able to measure a distance between any two elements in a space. That is possible in a metric and, in particular, in a normed space.

We will use the contraction mapping to find the solution to a function equation. Thus, in our case the elements of the space will be functions and the contraction map will map functions into functions. In our main application, the elements will be bounded functions on $X \subseteq \mathbb{R}^l$. The space X will correspond to the state space, and the functions on X will be candidates for the value function. To apply the contraction mapping theorem, we need to introduce a distance measure, i.e., a metric on our set of bounded functions $B(X)$. We will define our metric ρ using the supremum norm on $B(X)$ ¹⁴⁰

$$\rho(f, g) = \|f - g\|_\infty = \sup_x |f(x) - g(x)|,$$

which measures the distance of any two functions by the supremum of the absolute of the pointwise differences.

We can show that the set of bounded functions on $X \subseteq \mathbb{R}^l$ equipped with the supremum norm is a Banach space, i.e. a complete normed vector space.¹⁴¹ Similarly, the subspace of all bounded and continuous functions equipped with the supremum norm is a Banach space. Completeness is of great interest to us because we will prove the existence of a solution to our dynamic programming problem as the limit of the repeated application of a contracted

¹⁴⁰A metric ρ satisfies the following three properties: (i) $\rho(f, g) = 0 \Leftrightarrow f = g$ (two different elements cannot have a zero distance), (ii) $\rho(f, g) = \rho(g, f)$ (symmetry), and (iii) $\rho(f, g) = \rho(g, f)$ (triangle inequality). A norm induces a metric by defining the metric as the norm of the difference of two elements.

¹⁴¹A metric space is complete if every Cauchy sequence converges to an element in the space. If X is compact, boundedness is trivial because of continuity. For details see theorem 3.1 of ?, though Stokey and Lucas look at the subspace of bounded and continuous functions.

map. Only if the space is complete, we know that the limit of this sequence actually exists (and is in the same space sharing the same properties as the other functions).

Once we established that our space is complete metric we can forget about the precise structure for the moment. Let (S, ρ) be an arbitrary metric space, i.e. a space S equipped with a metric ρ . A map $T : S \rightarrow S$ on a metric space (S, ρ) is a contraction mapping (with modulus β) if for some $\beta \in (0, 1) : \rho(Tv, Tv') \leq \beta\rho(v, v') \forall v, v' \in S$.¹⁴² A fixed point $v^* \in S$ of a map $T : S \rightarrow S$ is a point that satisfies $Tv^* = v^*$.

Theorem 5 (Banach's Contraction Mapping of Fixed Point Theorem):

If (S, ρ) is a complete metric space and $T : S \rightarrow S$ is a contraction mapping with modulus β , then

- i) T has exactly one fixed point v^* in S , and
- ii) for any $v_0 \in S$: $\rho(T^n v_0, v^*) \leq \beta^n \rho(v_0, v^*)$, $n \in \mathbb{N}$.

Part ii) of the theorem tells us that iterative application of a contraction map will bring us arbitrarily close to its fixed point (from any starting point).

We return to our case where the elements of the metric space are bounded functions. The following theorem helps us to verify that a map is a contraction. Maps that take functions to functions are frequently referred to as operators.

Theorem 6 (Blackwell's sufficient conditions for a contraction):

Let $X \subseteq \mathbb{R}^I$ and let $B(X)$ be the space of all bounded functions $f : X \rightarrow \mathbb{R}$ with the supremum norm. Let $T : B(X) \rightarrow B(X)$ be an operator satisfying

- a. (monotonicity) $f, g \in B(X)$ and $f(x) \leq g(x) \forall x \in X$ imply

$$Tf(x) \leq Tg(x) \quad \forall x \in X$$

- b. (discounting) $\exists \beta \in (0, 1)$ such that for all $f \in B(X)$:

$$[T(f + a)](x) \leq Tf(x) + \beta a \quad \forall a \geq 0, x \in X.$$

Then T is a contraction with modulus β .

¹⁴²We follow (?) definition of the contraction mapping (see p. 50 for the theorem below). Sometimes, a contraction mapping is defined weaker and the existence of a (uniform) β characterizing the contraction as above is referred to as a uniform contraction.

A.4 Applicability of the Contraction Mapping

This section discusses how we employ the contraction mapping theorem to prove the existence of a solution to the dynamic programming problem (FE) formulated in equation (A.3). For this purpose we define the following map T taking a function v on X into the following function

$$Tv(k) = \max_{y \in \Gamma(k)} [F(k, y) + \beta v(y)] \text{ for all feasible } k.$$

This map is widely referred to as the Bellman operator. Before we even check whether T is a contraction mapping we have to verify that Tv is again an element of our original space. For the case of bounded functions the case is straight forward, we simply have to bound F or, equivalently, require continuity and assume compactness of Γ . If we are seeking a continuous value function, we have to make sure that Tv is continuous if v is continuous.

Let $C(X)$ denote the set of all continuous functions on X . Before applying the fixed point theorem to T we have to verify that if $v \in C(X)$ then $Tv \in C(X)$. The affirmative answer will depend on our assumptions with respect to the correspondence $\Gamma : X \rightarrow X$, describing feasibility constraints, and the return function F . Clearly, if F is continuous the function $F(k, y) + \beta v(y)$ is continuous. The question is whether the maximized function stays continuous. The answer is stated below in the Theorem of the Maximum. We state it for the general form that contains our particular problem in the case where $f(k, y) = F(k, y) + \beta v(y)$.

Let the function h be defined by the maximization

$$h(x) = \max_{y \in \Gamma(x)} f(x, y). \quad (\text{A.6})$$

If the maximum is attained, the set of values y attaining the maximum

$$G(x) = \{y \in \Gamma(x) : f(x, y) = h(x)\} \quad (\text{A.7})$$

is nonempty. The following theorem establishes restrictions on f and Γ such that the function h and the correspondence G vary continuously in x .

Theorem 7 (Theorem of the Maximum): Let $X \subseteq \mathbb{R}^l$ and $Y \subseteq \mathbb{R}^m$, let $f : X \times Y \rightarrow \mathbb{R}$ be a continuous function, and let $\Gamma : X \rightarrow Y$ be

a compact valued and continuous¹⁴³ correspondence. Then the function $h : X \rightarrow \mathbb{R}$ defined in equation (A.6) is continuous, and the correspondence $G : X \rightarrow Y$ defined in equation (A.7) is nonempty, compact-valued, and upper hemi-continuous.

If in addition the correspondence Γ is convex-valued and the function f is strictly concave in y , then G is single-valued. This fact will translate into an according theorem to establish uniqueness of the policy function in the dynamic programming context. In particular, the Theorem of the Maximum establishes that $Tv \in C(X)$ if $v \in C(X)$.

We proceed to show that Blackwell's conditions (Theorem 6) for a contraction mapping apply to the Bellman operator T . We start by showing monotonicity. For all $f, g \in B(X)$ with $f(k) \leq g(k) \forall k \in X$ we have

$$\begin{aligned} [F(k, y) + \beta f(y)] &\leq [F(k, y) + \beta g(y)] \quad \forall k, y \text{ feasible} \\ \Leftrightarrow \max_{y \in \Gamma(k)} [F(k, y) + \beta f(y)] &\leq \max_{y \in \Gamma(k)} [F(k, y) + \beta g(y)] \quad \forall k \text{ feasible} \\ \Leftrightarrow Tf(k) &\leq Tg(k) \quad \forall x \in X. \end{aligned}$$

We show that discounting is satisfied as by Assumption 6 we have $\beta \in (0, 1)$ and for all $f \in B(X)$ and for all feasible k

$$\begin{aligned} [T(f + a)](k) &= \max_{y \in \Gamma(k)} [F(k, y) + \beta(f(y) + a)] \\ &= \max_{y \in \Gamma(k)} [F(k, y) + \beta f(y)] + \beta a \\ &= Tf(k) + \beta a. \end{aligned}$$

Hence, the Bellman operator defines a contraction mapping on the Banach (complete and metric vector) space of all bounded functions.

The, property ii) of the contraction mapping theorem states that repeated application of the contraction mapping on any initial guess of the function brings us closer and closer to the fixed point, i.e., the value function solving (FE). We close the present section with a theorem ensuring not only that the value function converges from its initial guess to the fixed point, but that also the optimal control converges.

¹⁴³A correspondence is continuous if it is upper and lower hemi-continuous, see ?.

Theorem : Let $X \subseteq \mathbb{R}^l$ and $Y \subseteq \mathbb{R}^m$, let $f : X \times Y \rightarrow \mathbb{R}$ be a continuous function, and let $\Gamma : X \rightarrow Y$ be a nonempty, compact- and convex valued, and continuous correspondence with graph A .¹⁴⁴ Let $\{f_n\}$ be a sequence of continuous (real valued) functions on A . Assume that for each n and $x \in X$ the function $f_n(x, \cdot)$ is strictly concave in its second argument. Assume that f has the same properties and that $f_n \rightarrow f$ uniformly (in the sup norm). Define the functions

$$g_n(x) = \operatorname{argmax}_{y \in \Gamma(x)} f_n(x, y), \quad n \in \mathbb{N} \text{ and}$$

$$g(x) = \operatorname{argmax}_{y \in \Gamma(x)} f(x, y).$$

Then $g_n \rightarrow g$ pointwise. If X is compact, then $g_n \rightarrow g$ uniformly.

A.5 Existence of the Dynamic Programming Solution

This section proves the existence and uniqueness of a solution to the dynamic programming equation (FE) under the assumption that F is continuous and bounded and future value is discounted. We restate the corresponding function equation

$$v(k) = \max_{y \in \Gamma(k)} [F(k, y) + \beta v(y)] \quad \text{with } 0 \leq \beta < 1. \quad (\text{A.8})$$

We denote the correspondence mapping present states into feasible future states by $\Gamma : X \rightarrow X$ with graph A . The instantaneous payoff function is $F : A \rightarrow \mathbb{R}$. Existence and uniqueness of a solution require the following sufficient conditions on Γ and F .

Assumption 5: $X \subset \mathbb{R}^l$ is convex and $\Gamma : X \rightarrow X$ is nonempty, compact-valued and continuous.

Assumption 6: The function $F : A \rightarrow \mathbb{R}$ is bounded and continuous and $0 \leq \beta < 1$.

¹⁴⁴The graph A of a correspondence $\Gamma : X \rightarrow Y$ is defined as $A = \{(x, y) \in X \times Y : y \in \Gamma(x)\}$. In equation (A.6) f is a real valued function on the graph of Γ .

Note that assumptions 5-6 imply assumptions 3-4. Moreover, by Theorems 1-4 the solution of the functional equation also solves the optimization.¹⁴⁵ We define the Bellman operator T on $C(X)$ by

$$T v(k) = \max_{y \in \Gamma(k)} [F(k, y) + \beta v(y)] . \quad (\text{A.9})$$

The previous section established that, under Assumptions 5 and 6, the Bellman operator is a contraction mapping. Therefore, Banach's Fixed Point Theorem (Theorem 5) and the Theorem of the Maximum (Theorem 7) deliver the following result.

Theorem 8 (Unique Solution of FE) : Let K, Γ, F, β satisfy Assumptions 5-6. Let $C(K)$ be the space of bounded and continuous real valued functions over $K \subset \mathbb{R}^l$ equipped with the supremum norm. Then

- i) The operator T defined in equation (A.9) maps $C(K)$ into itself.
- ii) The operator T defined in equation (A.9) has a unique fixed point $v \in C(K)$.
- iii) For all $v_0 \in C(K)$ it is $\|T^n v_0 - v\| \leq \beta^n \|v_0 - v\|$, $n \in \mathbb{N}$
- iv) Given v , the policy correspondence $G : X \rightarrow X$ defined by

$$G(k) = \{y \in \Gamma(k) = F(k, y) + \beta v(y)\} \quad (\text{A.10})$$

is compact-valued and u.h.c.

The theorem ensures a solution to our dynamic programming problem. Moreover, Theorem 2 shows that our solution to the dynamic programming problem (FE) is also a solution to the maximization problem (SP) of the infinite sum of future payoffs. Thus, Theorems 2 and 8 establish that under assumptions 5-6 the supremum function of problem (SP) in equation (A.1) is bounded and continuous. Adding Theorem 4, we have established that there exists at least one optimal plan and that any plan generated by the (nonempty) correspondence G is optimal.

¹⁴⁵The transversality condition is not needed because $F \leq B$ is bounded so that v is bounded by a converging geometric sum to $v \leq \frac{B}{1-\beta}$ implying $\lim_{t \rightarrow \infty} \beta^t v(k_t) = 0$ so that the transversality condition is satisfied.

A.6 Further Properties of the Solution

This section analyzes additional assumptions on the primitives that ensure additional properties of the value function v and the policy correspondence G . We start by characterizing an increasing value function.

Assumption 7: For each y the function $F(\cdot, y)$ is strictly increasing in each of its first l arguments.

Assumption 8: Γ is monotone: $k < k'$ implies $\Gamma(k) \subset \Gamma(k')$.

Adding these assumption, we find that the value function is strictly increasing in the state.

Theorem 9 (strictly increasing v): Let K, Γ, F, β satisfy assumptions 5-8 and let v be the unique solution to equation (A.8). Then, v is strictly increasing.

Next, we seek conditions under which the value function is strictly concave and the policy correspondence is single valued, i.e., a function.

Assumption 9: F is strictly concave: For any $0 < \theta < 1$ and $(k, y), (k', y') \in A$:

$$F[\theta(k, y) + (1 - \theta)(k', y')] \geq \theta F(k, y) + (1 - \theta)F(k', y')$$

and the inequality is strict if $k \neq k'$.

Assumption 10: Γ is convex: For any $0 \leq \theta \leq 1$ and $k, k' \in K$:

$$y \in \Gamma(k) \text{ and } y' \in \Gamma(k') \Rightarrow \theta y + (1 - \theta)y' \in \Gamma[\theta k + (1 - \theta)k'] .$$

Assumption 10 implies that for each $k \in K$ the set $\Gamma(k)$ is convex and that there are no ‘increasing returns’. Since K is convex by Assumption assumption is equivalent to assuming that the graph A is convex.

Theorem 10: Let K, Γ, F, β satisfy assumptions 5-6 and 9-10 and let v satisfy equation (A.8) and G satisfy equation (A.10). Then, v is strictly concave and G is a continuous and single-valued function.

We note that the proof of these theorems employs the fact that the operator T preserves certain properties.

Finally, it would be convenient to establish differentiability properties of the value function. For example, in the one-sector growth model from section [A.1](#)

$$v(k) = \max_{0 \leq y \leq f(k)} U[f(k) - y] + \beta v[y]$$

differentiability of v allows us to write the maximization condition as

$$U'[f(k) - g(k)] = \beta v'[g(k)] .$$

Note that the one sector growth problem falls into the category of bounded problems analyzed here if we add the assumptions that $f(0) = 0$ and the existence of $\bar{k} > 0$ such that $k \leq f(k) \leq \bar{k} \forall 0 \leq k \leq \bar{k}$ and $f(k) < k \forall k > \bar{k}$.

For continuous differentiability of the value function we add an assumption on the continuous differentiability onto our objective function F .

Assumption 11: F is continuously differentiable on the interior of A .

Theorem 11 (Differentiability of v): Let K, Γ, F, β satisfy assumptions [5-6](#) and [9-11](#) and let v satisfy equation [\(A.8\)](#) and g satisfy equation [\(A.10\)](#). If $k_0 \in \text{int } K$ and $g(k_0) \in \text{int } \Gamma(k_0)$, then v is continuously differentiable at k_0 . The derivatives are given by

$$v_i(k_0) = F_i[k_0, g(k_0)], \quad i = 1, \dots, l.$$

Whereas it is straight forward to establish differentiability of the value function v , conditions ensuring twice differentiability and, thus, differentiability of g are very strong, see [?](#).

B The Basics of Ordinary Differential Equations

The objective of this chapter is to provide sufficient background in differential equations to enable readers unfamiliar with this topic to understand and use

the subsequent chapters on optimal control. We begin by introducing several basic terms and results in the field of ordinary differential equations (ODEs), and then discuss phase portrait analysis using examples. We then discuss the use of linear approximations to non-linear differential equations, emphasizing the role of these approximations in determining the stability of a steady state.

Continuous time optimal control models use differential equations to describe the change in the state variable, and an integral to describe the objective. The analysis of these models requires a basic understanding of ODEs. Many optimal control models, particularly those that are designed for qualitative analysis, involve a single state variable. The necessary conditions for optimality in these scalar models can be written as a pair of ordinary differential equations, together with boundary conditions. Two-dimensional systems of differential equations can be analyzed qualitatively, using phase portrait analysis.

Some interesting dynamic models do not involve optimal control. Those problems can be studied using the methods described here. We provide two examples of such problems

B.1 Elements of ODEs

An ordinary differential equation is an equation that involves an ordinary (as distinct from a partial) derivative. For example,

$$g\left(y, x, \frac{dx}{dy}\right) = 0$$

is an ODE. In many cases of interest, we can write the derivative as a function of x, y , in which case we have an explicit ODE. Often, in models involving optimal control, the dependent variable, here y , is time, denoted t . In these cases, the ODE takes the form

$$\dot{x} \equiv \frac{dx}{dt} = f(x, t). \quad (\text{B.1})$$

A dot over a variable indicates the total derivative of that variable with respect to time. The solution to an ODE is a family of curves, $x_t = \phi(t)$, whose derivative satisfies the original equation, i.e.

$$\frac{d\phi}{dt} = f(\phi(t), t).$$

In many cases it is not possible to find the explicit form of the function ϕ . However, in certain special cases, a known solution exists. There are procedures for finding solutions in special types of problems. The simplest circumstance is where the function f is multiplicatively separable, i.e. $f(x, t) = b(x)c(t)$, so that we can write the differential equation as

$$\frac{dx}{b(x)} = c(t)dt.$$

By integrating both sides of this identity, we obtain a solution of x in terms of an integral. This integral typically involves a constant of integration. For example, if $b(x) = x$ and $c(t) = c$, the differential equation is linear with constant coefficient and we can write it as

$$\frac{dx}{x} = cdt. \tag{B.2}$$

Integrating this equation gives

$$\ln x = ct + A \implies x = ae^{ct},$$

where A is the unknown constant of integration and $a = \exp(A)$. The solution is a *family* of curves, because it depends on the parameter a , which can take any value.

We check the solution by differentiating it and confirming that the result satisfies the original differential equation:

$$\frac{dx}{dt} = \frac{d(ae^{ct})}{dt} = cae^{ct} = cx,$$

as required. A common approach to solving differential equations is to “guess” the form of the solution, substitute this guess into the differential equation, and then observe what must be true in order for the guess to be correct. For example, if we guess that the solution to equation B.2 is of the form $x = ae^{a_1t}$, then upon substitution of this guess into the original ODE we find that

$$\frac{dx}{dt} = \frac{d(ae^{a_1t})}{dt} = a_1ae^{a_1t} = a_1x.$$

This procedure shows that it must be the case that $a_1 = c$. As this example suggests, the guess is not random. We use this method when we know, or

at least have a good idea of the form of the solution, and need only to learn something specific about this solution; here, the specific element we want to learn is the value of a_1 . We use this method of (informed) guessing several times in subsequent chapters.

A given value of x at a particular time, t , is known as a boundary value. When $t = 0$ it is usually referred to as an initial condition. In the example above, if we know that $x(7) = 4$, we can substitute this information into the solution to obtain $4 = ae^{7c}$ to obtain $a = \frac{4}{e^{7c}}$, and then write the solution as $x(t) = 4e^{(t-7)c}$. This example shows how a boundary condition determines the value of the constant of integration. The solution to an ODE with a boundary condition is a function, rather than a family of functions. This solution is often called a path; the graph of the solution, as a function of time, is a curve in the (t, x) plane. When we want to indicate a path that goes through the point (t_0, x_0) we denote the solution as $x(t) = \phi(t; t_0, x_0)$.

We now return to the general ODE, equation B.1. This equation is *autonomous* if and only if $f(x)$ is independent of t , i.e. when $\dot{x} = f(x)$. In most cases where we care about a steady state, we are dealing with autonomous equations. A steady state, denoted x_∞ , is a value of x at which $\dot{x} = 0$, i.e. it is a root of the algebraic (not differential) equation $f(x) = 0$. In the case where f is linear in x there is a single steady state, but for non-linear f there can be multiple steady states.

There are three types of stability in this setting. A steady state x_∞ is said to be *stable* if trajectories that begin close to the steady state remain close. Formally, for all $\varepsilon > 0$ and $t_0 \geq 0$, there exists $\delta(\varepsilon, t_0)$ such that

$$|x_0 - x_\infty| < \delta \implies |\phi(t; t_0, x_0) - x_\infty| < \varepsilon.$$

A steady state is *asymptotically stable* if all paths that begin close to the steady state eventually become and remain arbitrarily close to the steady state. Formally, there exists δ such that

$$|x_0 - x_\infty| < \delta \implies \lim_{t \rightarrow \infty} \phi(t; t_0, x_0) = x_\infty.$$

Asymptotic stability is stronger than stability, because the former implies that the path approaches the steady state, whereas the latter means only that the path remains close to the steady state. For example, a path that remains a constant distance from its steady state without approaching the steady state is stable but not asymptotically stable. A steady state is *globally*

asymptotically stable if the path approaches the steady state regardless of the initial condition. Global asymptotic stability implies that there is a unique steady state.

We have defined the three types of stability in the context of a scalar problem, but the same definitions apply when considering a system of ODEs, with $n > 1$ variables. There, however, we need to recognize that the neighborhood of a steady state is an n -dimensional set rather than an interval on the real line. When we apply the test for stability we need to consider all initial conditions in the neighborhood of a steady state. For example, let \mathbf{x} be an n -dimensional vector of state variables, \mathbf{f} a vector of functions, with $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$, and let \mathbf{x}_∞ be a steady state. In order for \mathbf{x}_∞ to be asymptotically stable (for example), it must be the case that for all $t_0 \geq 0$, there exists δ such that

$$\sqrt{\sum_{i=1}^n (x_{i0} - x_{i\infty})^2} < \delta \implies \lim_{t \rightarrow \infty} \phi_i(t; t_0, \mathbf{x}_0) = x_{i\infty} \text{ for all } i.$$

The term to the left of the inequality is the Euclidean distance between the initial condition \mathbf{x}_0 and the steady state \mathbf{x}_∞ . The inequality in the hypothesis states that the Euclidean distance between the initial condition and the steady state is small. The implication of the hypothesis is that in the limit each element of the vector converges to its corresponding steady state, which of course is equivalent to stating that the Euclidean distance between the state and the steady state approaches 0.

In most cases we will be interested in two-dimensional systems of autonomous ODEs, i.e. those of the form

$$\dot{x} = F(x, y) \text{ and } \dot{y} = G(x, y).$$

The variables x and y might be the biomass of two stocks of fish that interact, e.g. as predator and prey. In optimal control settings, one variable is typically the state variable, such as the stock of fish, and the other variable might be the shadow value of the stock, or the optimal harvest. We have the following existence theorem for the *initial value problem*, in which the boundary conditions are specified at the same point time, which without loss of generality can be taken as the initial time. Later we will be concerned with other types of boundary value problems, such as those arising in an optimal control context.

Theorem 12: If the functions F and G are continuous, bounded, and with continuous first derivatives, then for every initial condition $x(t_0) = x_0$ and $y(t_0) = y_0$ there exists a unique solution to the system that satisfies this initial condition. The solution might be bounded, or unbounded. In the latter case, the Euclidean norm $\sqrt{(x_t^2 + y_t^2)} \rightarrow \infty$ as $t \rightarrow \infty$.

B.2 Phase portrait analysis

We will repeatedly use phase portrait analysis to study the qualitative features of a pair of differential equations. The idea is as follows. Given an initial value, the solution to the pair of differential equations is a pair of functions $x_t = \phi(t; t_0, x_0, y_0)$ and $y_t = \psi(t; t_0, x_0, y_0)$. Note that the boundary condition specifies the value of x and y at a time t_0 . We can graph these paths in (x, y, t) space. The triple $(\phi(t; t_0, x_0, y_0), \psi(t; t_0, x_0, y_0), t)$ describes a curve in three dimensional space, intersecting the point (x_0, y_0, t_0) . Without loss of generality we can set $t_0 = 0$, so that the curve starts on the (x, y) plane, where $t = 0$. We can project this curve on to the (x, y) plane, obtaining a curve in two dimensional space. This curve is referred to as a trajectory; it is the projection of the path in three dimensions on to the two-dimensional (x, y) plane. The family of all such trajectories, i.e. for all initial conditions, is called the phase portrait. The remarkable fact is that we can obtain the phase portrait using information about the primitive functions, F and G , without having to solve the differential equations. In view of the importance of this technique, we illustrate it using two examples, of increasing complexity.

An *isocline* is a set of points, typically a curve or a set of curves, at which \dot{x} or \dot{y} is constant. We are particularly interested in two isoclines, one where $\dot{x} = 0$ and the other where $\dot{y} = 0$. Along such an isocline, x or y is constant, because its time derivative is zero. We refer to these as the 0-isoclines, or when the meaning is clear, simply as the isoclines.

For the first example (taken from Clark (cite)) let

$$\dot{x} = y^2 \text{ and } \dot{y} = x^2.$$

We use this example to illustrate the recipe for drawing the phase portrait. The simplicity of the example makes it easy to follow the steps of the recipe.

The *first step* is to find the 0-isoclines. For this example, the $\dot{x} = 0$ isocline is simply the x axis (where $y = 0$) and the $\dot{y} = 0$ isocline is the y axis (where $x = 0$). These two isoclines intersect once and therefore divide the (x, y) plane into four *isosectors*, which in this example happen to be the orthants of the plane. More generally, the isoclines might intersect more than once, and therefore might divide the plane into several regions. Also, more generally a particular isocline might consist of more than a single curve, in which case the division of the plane is more complex. The intersection of the 0-isoclines determines the steady states. In this example, there is a single steady state, $(0, 0)$. At any other point, the trajectory changes over time.

The *second step* is to include directional arrows in the figure. Recall that these trajectories are the projections of a path in three dimensional space, on to the (x, y) plane. The convention is that an arrow pointing East indicates that x is increasing over time, whereas an arrow pointing West indicates that x is decreasing over time. Similarly, an arrow pointing North indicates that y is increasing over time, and an arrow pointing South indicates that y is decreasing over time. Figure 1 shows the isoclines and the directional arrows for our first example.

We can determine by inspection the direction of the arrows, but for the purpose of more difficult problems it is important to understand the procedure. At any point on the $\dot{x} = 0$ isocline x is not changing over time, but for points on either side of the isocline $\dot{x} > 0$. Therefore, at all points above and below the isocline, the East-West arrow points East, indicating that x is increasing over time. The idea is that we examine the value of the function $F(x, y)$ in the neighborhood of the isocline in order to determine the change in x on either side of the isocline. For this example, a North-South arrow through any point off the x axis (where $\dot{y} = 0$) points North, regardless of the isosector; for any point off the y axis (where $\dot{x} = 0$) an East-West arrow points East, regardless of the isosector. Therefore, every trajectory follows a North-East path. A useful convention is to connect the base of the East-West and the North-South arrow in each isosector.

The slope of a trajectory through any point is given by

$$\frac{dy}{dx} = \frac{G(x, y)}{F(x, y)} = \frac{x^2}{y^2} \quad (\text{B.3})$$

Figure 1 shows three trajectories beginning in the third quadrant. The middle trajectory is a line through the origin with slope 1. Trajectories

through points on this line below the origin converge to $(0, 0)$. Trajectories through points on this line above the origin diverge to (∞, ∞) , as does any trajectory beginning off this line. As a trajectory crosses the $\dot{x} = 0$ isocline (here the x axis), its slope is infinite, and as a trajectory crosses the $\dot{y} = 0$ isocline (here the y axis) its slope is 0. The two outer trajectories illustrate these features.

For this example, we can obtain an explicit solution to a trajectory, by solving the ODE equation B.3. Equation B.3 implies

$$y^2 dy = x^2 dx \implies \int y^2 dy = \int x^2 dx \implies \frac{y^3}{3} = \frac{x^3}{3} + \tilde{c} \implies y = (x^3 + c)^{\frac{1}{3}},$$

where c is the constant of integration multiplied by 3. To find the value of y for a given value of x , of a trajectory that begins at (x_0, y_0) we solve

$$y_0 = (x_0^3 + c)^{\frac{1}{3}} \implies c = y_0^3 - x_0^3 \implies y = (x^3 + (y_0^3 - x_0^3))^{\frac{1}{3}}$$

For slightly more complicated systems we cannot obtain the solution in closed form, but we have no need to do so if our goal is merely to determine qualitatively how the trajectory changes over time.

The discussion above leads to an important point: *trajectories do not cross*. To confirm that trajectories cannot cross outside the steady state (defined as a root of $F(x, y) = G(x, y) = 0$) we show in Figure 2 a circumstance where they do cross, and explain why that circumstance is impossible. The figure shows the isoclines (the solid curves) for an unspecified system, for which the unique steady state is at (x_∞, y_∞) . Here there are four isosectors. For this example, trajectories above the $\dot{x} = 0$ isocline are increasing, and those below the isocline are decreasing, as the arrows indicate. The North-South arrows show the change in y in the different isosectors. The figure also shows an impossibility, a case where two trajectories, the dashed curves labelled (1) and (2), cross at point a , which is not a steady state. The slope of a trajectory through point a has the slope $\frac{G(x,y)}{F(x,y)}$. Any trajectory through a has this slope, a single number, or infinity if $F(x, y) = 0$. The figure however shows that there are two trajectories, with different slopes, as must

be the case if the trajectories cross. Therefore, the situation shown in the figure, where trajectories cross outside of a steady state, cannot occur.

Trajectories that approach or emanate from a state do not cross either, but they might appear to do so. The trajectories labelled (iii) and (iv) approach the steady state as $t \rightarrow \infty$, but they do not cross. For other types of dynamic systems there are unaccountably many trajectories that approach or diverge from the steady state, but these do not cross either.

We now consider the second example of a phase portrait, a predator-prey model. The dynamic system is

$$\begin{aligned} \text{prey: } \dot{x} &= (A - By - \omega x)x \\ \text{predator: } \dot{y} &= (Cx - D - \mu y)y \end{aligned}$$

where x is the biomass of prey and y is the biomass of predators. All parameters are positive. For example, $B > 0$ means that an increase in the biomass of predators decreases the growth rate of the prey; $C > 0$ means that an increase in the biomass of prey increases the growth of predators; ω is a measure of *congestion*: as the biomass of prey increases, members of the population face increased competition for food and other resources, so the growth rate of the population falls; $\mu > 0$ means that there is competition amongst the predators for prey, a kind of congestion. A special case of this model, where $\omega = \mu = 0$, is known as the Volterra-Lotka system.

Figure 3 shows the phase portrait for the case where $\frac{D}{C} > \frac{A}{\omega}$. We now describe its construction. Recall that the first step of the exercise is to find and graph the 0-isoclines. We have

$$\begin{aligned} \dot{x} = 0 &\implies x = 0 \text{ or } A - By - \omega x = 0 \\ \dot{y} = 0 &\implies y = 0 \text{ or } Cx - D - \mu y = 0. \end{aligned}$$

Each equation has two 0-isoclines. Figure 3 shows the graph of $A - By - \omega x = 0$, denoted as the line L , and the graph of $Cx - D - \mu y = 0$, denoted as the line M . The slopes of these lines follows from the assumption that the parameters are positive, and the relative positions of the x intercepts follows from the assumption that $\frac{D}{C} > \frac{A}{\omega}$. Figure 4 shows the phase portrait when that inequality is reversed.

The second step is to find the directional arrows. Again, due to the simplicity of the example, this is easy to do by inspection. However, we proceed a bit

more systematically in order to demonstrate a method that is useful for more complicated models. In the interior of the positive orthant (i.e. not including the axes) we know that $\dot{x} = 0$ only on the line L . We want to know how the magnitude of \dot{x} changes as we move from a point slightly to the left to a point slightly to the right of this line. We therefore examine the derivative

$$\frac{\partial \dot{x}}{\partial x|_L} = (A - By - \omega x) - \omega x = -\omega x < 0,$$

where the second equality follows from the fact that we are evaluating the derivative on L , where $A - By - \omega x = 0$. Therefore, we know that \dot{x} is decreasing in x in the neighborhood of L . Because $\dot{x} = 0$ on L , it must be the case that $\dot{x} > 0$ slightly to the left of this line, and $\dot{x} < 0$ slightly to the right of this line. We also know that \dot{x} never changes sign in the positive orthant, off the line L . We therefore conclude that $\dot{x} > 0$ everywhere to the left of this line, and $\dot{x} < 0$ everywhere to the right of this line. The arrows pointing East below L and pointing West above L reflect this information. We use exactly the same procedure to determine the East-West directional arrows in Figure 4; the only difference is that there are four rather than three isosectors there. It is important to show directional arrows in each isosector.

We could also have determined the directional arrows by investigating the change in \dot{x} as we increase y (rather than x) in the neighborhood of L . That is, we could have evaluated

$$\frac{\partial \dot{x}}{\partial y|_L} = -Bx < 0.$$

Using the same reasoning as above, we would then conclude that every point in the positive orthant above L , $\dot{x} < 0$ and at every point below L , $\dot{x} > 0$. In other words, we obtain the same information by examining the partial derivative of \dot{x} with respect to either x or y . In some cases one derivative is easier to evaluate than the other, but in this case, both derivatives are simple.

We follow the same procedure to determine that $\dot{y} < 0$ in the positive orthant above the isocline labelled M , and $\dot{y} > 0$ below this isocline. We draw in the North-South directional arrows in each isosector in Figures 3 and 4. Again, we connect the directional arrows at their base.

We also include directional arrows for trajectories crossing the isoclines. Recall that the slope of a trajectory is

$$\frac{dy}{dx} = \frac{\dot{y}}{\dot{x}} = \frac{(Cx - D - \mu y)y}{(A - By - \omega x)x}.$$

A trajectory that crosses the line L has an infinite slope because the denominator of the derivative is 0. In Figure 3, every trajectory that crosses the line L lies above the line M , where $\dot{y} = 0$. Therefore, such a trajectory is vertical, pointing South.

In both cases, $\frac{D}{C} > \frac{A}{\omega}$ and $\frac{D}{C} < \frac{A}{\omega}$, one steady state is $x = 0 = y$ and a second steady state is $x = \frac{A}{\omega}$, $y = 0$. If a population begins with $x = 0 < y$, it converges to the steady state $x = 0 = y$; in the absence of prey, the predators die out. The directional arrow on the y axis contains this information. If a population begins with $x > 0 = y$, the population converges to $\frac{A}{\omega}$, the steady state of the prey in the absence of predators, also known as the *carrying capacity* of the stock. The directional arrow on the x axis contains this information. Both of these steady states are unstable and therefore also not asymptotically stable.

Recall that in a multidimensional context, stability and asymptotic stability requires that paths from all initial conditions in the neighborhood of the steady state remain close to (under stability) or converge to (under asymptotic stability) the steady state. In the case of the $(0, 0)$ steady state, there are *some* initial conditions in the neighborhood – i.e. those on the y axis – from which trajectories converge to the steady state. But there are other initial conditions, – i.e. all those off the y axis – from which trajectories do not converge to this steady state. Therefore $(0, 0)$ is an unstable steady state, as is $(\frac{A}{\omega}, 0)$. We study the stability of the interior steady state in Figure 4 using material in the next section.

B.3 Linear approximations

By examining the linear approximation to a system of non-linear ODEs in the neighborhood of a steady state, we obtain information about the stability of that steady state under the original nonlinear system. This procedure uses the Theorem of the First Approximation, which gives conditions under which a linear approximation provides a reliable guide to the behavior of a non-linear system in the neighborhood of a steady state.

A function is *analytic* if it is sufficiently smooth that its Taylor approximation converges to the true value of the function.¹⁴⁶ If a function is analytic in a set that excludes a particular point, then that point is said to be a singularity, or a singular point. For example, the function $\frac{1}{x}$ is analytic at all points other than $x = 0$, a singularity.

We return to the general two dimensional system above:

$$\dot{x} = F(x, y) \text{ and } \dot{y} = G(x, y), \quad (\text{B.4})$$

and we assume that both functions are analytic. This assumption means that we can take a first order Taylor expansion in the neighborhood of the steady state. The first order expansion of F is

$$\begin{aligned} \dot{x} = & F(x_\infty, y_\infty) + F_x(x_\infty, y_\infty)(x - x_\infty) + F_y(x_\infty, y_\infty)(y - y_\infty) \\ & + o(x - x_\infty, y - y_\infty). \end{aligned}$$

The term $o(x - x_\infty, y - y_\infty)$ (“little oh”) contains the higher order terms of the expansion, and satisfies

$$\lim_{\Delta \rightarrow 0} \frac{o(x - x_\infty, y - y_\infty)}{\Delta} = 0,$$

where $\Delta = \sqrt{(x - x_\infty)^2 + (y - y_\infty)^2}$ is the Euclidean norm of the vector $(x - x_\infty, y - y_\infty)$. The equality states that the contribution of all higher order terms gets small faster than does Δ , as Δ approaches 0.

We simplify the expansion using $F(x_\infty, y_\infty) = 0$ and defining $X = (x - x_\infty)$ and $Y = (y - y_\infty)$, the difference between a state variable and its steady state. Also using $\dot{X} = \dot{x}$ (because x_∞ is a constant), we write the first order approximation as

$$\dot{X} = aX + bY$$

with $a = F_x(x_\infty, y_\infty)$ and $b = F_y(x_\infty, y_\infty)$. Using the same procedure and analogous definitions, we write the expansion of the \dot{y} equation as

$$\dot{Y} = cX + dY$$

with $c = G_x(x_\infty, y_\infty)$ and $d = G_y(x_\infty, y_\infty)$. Notice that by replacing the variables x, y with X, Y , their deviations from the steady state, we obtain a

¹⁴⁶See Judd page 196 for a rigorous definition, in terms of a polynomial in the complex plane.

system of linear ODEs whose steady state is $(0, 0)$. This fact simplifies the notation because it means that we do not have to carry around a symbol for the steady state.

If we had an n -dimensional system of non-linear ODEs we would use the same procedure to obtain a first order approximation, a system of linear ODEs

$$\dot{z} = Az. \quad (\text{B.5})$$

For our two-dimensional system

$$z = \begin{pmatrix} X \\ Y \end{pmatrix} \text{ and } A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

The Theorem of the First approximation, which applies for a general n -dimensional system, states

Theorem 13: If the ODEs are analytic in the neighborhood of a steady state, then: (i) If the steady state of the linear approximation, equation B.5, is asymptotically stable, then the steady state of the original non-linear system is asymptotically stable; and (ii) If the steady state of the linear approximation is unstable, then the steady state of the non-linear system is unstable.

If the steady state of the linear system is either asymptotically stable or unstable, then the steady state of the original system has the same property. The only time the approximation is uninformative is when its steady state is stable but not asymptotically stable. In that case, we can make no conclusions about the stability of the original system. This ambiguity is quite intuitive. Recall that a steady state is stable but not asymptotically stable if all paths emanating from initial conditions near the steady state remain near, but do not converge to the steady state. This behavior is a knife-edge, in the following sense. If the steady state was slightly more attractive, paths beginning near it would converge to it, and if it was slightly less attractive, paths beginning near it would diverge from it. When the approximation exhibits this kind of knife-edge behavior, we cannot tell whether the system being approximated has the same knife-edge behavior, or falls to one side (is asymptotically stable or unstable).

The next task is to determine the stability of the linear system. Here we use the n -dimensional system, equation B.5. If this system consisted of a single equation, i.e. if z were a scalar with $\dot{z} = az$, the solution would be trivial. We considered this example in Section B.1 where we saw that the solution is $z(t) = z_0 e^{at}$, where z_0 is the initial condition. Thus, the linear equation is asymptotically stable if $a < 0$, it is stable but not asymptotically stable if $a = 0$, and it is unstable if $a > 0$. We use a multi-dimensional analog, based on eigenvalues, to assess the stability of the n -dimensional system.

An eigenvalue of the matrix A , λ_i and its corresponding eigenvector, p_i , are solutions to

$$Ap_i = \lambda_i p_i.$$

Using the n by n identity matrix I we can rewrite this equation as

$$(A - I\lambda) p_i = 0,$$

which has a non-trivial solution if and only if

$$|A - I\lambda| = 0, \tag{B.6}$$

where $|A - I\lambda|$ is the determinant of the matrix $A - I\lambda$. Equation B.6 is the *characteristic equation* for the matrix A . It is an n degree polynomial in λ , and thus has n roots – not necessarily real and not necessarily unique. We construct the diagonal matrix of eigenvalues

$$\Lambda = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \dots & \\ & & & \lambda_n \end{bmatrix}$$

and the matrix whose i 'th column is the i 'th eigenvector

$$P = [p_1 \quad p_2 \quad \dots \quad p_n].$$

With this notation we write the system

$$AP = P\Lambda.$$

There are several possible cases, depending on whether the eigenvalues are real or complex, and on whether the matrix P of eigenvectors is non-singular.

The treatment here is not exhaustive; Boyce and DiPrima (??) and other ODE textbooks provide detailed treatments.

If P is non-singular (its inverse exists) we can premultiply the previous equation by P^{-1} to obtain

$$P^{-1}AP = \Lambda.$$

In this case, define $w = P^{-1}z$, or $Pw = z$, so that

$$\dot{w} = P^{-1}\dot{z} = P^{-1}Az = P^{-1}APw = \Lambda w.$$

Thus, when P is non-singular we can replace the n -dimensional system of linear ODEs with n independent linear ODEs.

If, in addition to the non-singularity of P , the roots λ_i are all real, then the solution to each of these equations is $w_i(t) = w_i(0)e^{\lambda_i t}$, where $w_i(0)$ is the initial condition for the i 'th element. Because Λ is a diagonal matrix, we can write¹⁴⁷

$$e^{\Lambda t} = \begin{bmatrix} e^{\lambda_1 t} & & & \\ & e^{\lambda_2 t} & & \\ & & \dots & \\ & & & e^{\lambda_n t} \end{bmatrix}.$$

With this notation we can stack up the solutions to the individual ODEs to write the system in matrix notation, as

$$w(t) = e^{\Lambda t}w(0).$$

Using the fact that $w = P^{-1}z$ we can rewrite this system in terms of the original variables, z :

$$\begin{aligned} P^{-1}z(t) &= e^{\Lambda t}P^{-1}z(0) \implies \\ z(t) &= Pe^{\Lambda t}P^{-1}z(0) = Pe^{\Lambda t}k \end{aligned}$$

with $k = P^{-1}z(0)$, a linear combination of the initial conditions. It is instructive to write this system out as

$$z(t) = k_1 p_1 e^{\lambda_1 t} + k_2 p_2 e^{\lambda_2 t} + k_3 p_3 e^{\lambda_3 t} + \dots + k_n p_n e^{\lambda_n t}. \quad (\text{B.7})$$

In words, the vector $z(t)$ is a linear combination of n vectors p_i each multiplied by $e^{\lambda_i t}$.

¹⁴⁷Caution: for a non-diagonal matrix A , e^A does *not* equal the matrix whose (i, j) element is $\exp(a_{ij})$.

Thus, when P is non-singular and the eigenvalues are real, we conclude that the steady state of the linear system is asymptotically stable if and only if all eigenvalues are negative. The steady state is stable but not asymptotically stable if some eigenvalues are 0 and the rest are negative. A steady state is unstable if any of the eigenvalues are positive.

A *saddle point* is a particular kind of unstable steady state, in which there are some positive and some negative eigenvalues. For example, suppose that $\lambda_i < 0$ for $i = 1, 2, \dots, s$ and the remaining eigenvalues are positive or zero. If it happens to be the case that $k_i = 0$ for $i \geq s + 1$, then using equation B.7 we have

$$z(t) = k_1 p_1 e^{\lambda_1 t} + k_2 p_2 e^{\lambda_2 t} + k_3 p_3 e^{\lambda_3 t} + \dots + k_s p_s e^{\lambda_s t}.$$

Because all of the exponents in this expression are negative, the sum converges to 0 as $t \rightarrow \infty$. Setting $t = 0$ in the equation above we have

$$z(0) = k_1 p_1 + k_2 p_2 + k_3 p_3 + \dots + k_s p_s.$$

This equation states that the initial value of z , $z(0)$, is a linear combination of the eigenvectors associated with the stable (i.e. negative) eigenvalues.

Corresponding to a saddle point, there is a *stable manifold*, defined as the set of initial conditions from which a path converges to the steady state. Any path emanating from an initial condition not in the stable manifold, diverges, i.e. it eventually moves away from the steady state. In the linear model, the stable manifold is a hyperplane. For example, if there is a single negative eigenvalue, the stable manifold has dimension 1, i.e. it is a line; if there are two negative eigenvalues, the stable manifold is a plane. Systems of non-linear ODEs can also have saddle points and corresponding stable manifolds. These tend to be curved. For example, in the case of a single negative eigenvalue, the stable manifold is a curve rather than a line, and in the case of two negative eigenvalues the stable manifold is a warped two-dimensional surface rather than a plane.

Thus far we have considered only the case where the eigenvalues are real and the matrix of eigenvectors is non-singular. If P is non-singular but some of the roots are complex, the solution becomes more complicated. For illustration, suppose that $\lambda_1 = \alpha + \beta i$, $\lambda_2 = \alpha - \beta i$, where α β are real numbers and i is the imaginary number, $\sqrt{-1}$, and all other eigenvalues are real. The first two eigenvectors are also complex, i.e. they take the form

$\eta \pm \nu i$, where η and ν are vectors. In this case, the solution to the linear system has the form

$$z(t) = k_1 \theta(t) + k_2 \rho(t) + k_3 p_3 e^{\lambda_3 t} + \dots + k_n p_n e^{\lambda_n t}$$

with

$$\theta(t) = e^{\alpha t} (\eta \cos \beta t - \nu \sin \beta t) \quad \text{and} \quad \rho(t) = e^{\alpha t} (\eta \cos \beta t + \nu \sin \beta t). \quad (\text{B.8})$$

In this case, stability depends on the real part of the complex eigenvalue, α . For example, if all of the real eigenvalues are negative, then the steady state is asymptotically stable if $\alpha < 0$, it is stable but not asymptotically stable if $\alpha = 0$, and it is unstable if $\alpha > 0$.

Finding the eigenvectors requires solving a polynomial of degree n , which may be a daunting task. Two facts about square matrices helps in some cases:

$$\text{trace}(A) = \sum_{i=1}^n \lambda_i \quad \text{and} \quad |A| = \prod_{i=1}^n \lambda_i. \quad (\text{B.9})$$

For our two-dimensional system

$$\begin{pmatrix} \dot{X} \\ \dot{Y} \end{pmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{pmatrix} X \\ Y \end{pmatrix}, \quad (\text{B.10})$$

we have $\text{trace}(A) = a + b \equiv p$ and $|A| = ad - bc \equiv q$. Using the characteristic equation [B.6](#) we have

$$\begin{aligned} \begin{vmatrix} a - \lambda & b \\ c & d - \lambda \end{vmatrix} &= \lambda^2 - \mathbb{P}\lambda + q = 0 \\ \Rightarrow \lambda &= \frac{p \pm \sqrt{\Gamma}}{2}, \end{aligned}$$

with $\Gamma = p^2 - 4q$, the discriminant of the quadratic equation.

Figure 5 (from Boyce and DiPrima) summarizes a great deal of information about the two-dimensional system, in terms of the trace, determinant and discriminant (p , q , and Γ) of the matrix A :

- If $q < 0$ then both roots are real; one root is positive and the other negative, so the steady state is a saddle point.

- If $q > 0 > p$ (the second orthant of the (p, q) plane) then both roots must be of the same sign (because $q = \lambda_1 \lambda_2 > 0$) and at least one root must be negative, or have negative real part (because $\lambda_1 + \lambda_2 = p < 0$). Therefore, both roots must be negative, or have negative real part, so the steady state is asymptotically stable.

Case 1 If $\Gamma < 0$, i.e. at a point above the curve not on the q axis, then the roots must be complex, so the solution depends on the sine and cosine of t , as in equation B.8. In this case, the trajectory is a stable spiral.

Case 2 If $\Gamma > 0$, i.e. at a point in the second quadrant below the curve, then both roots are real and negative. Here the solution is referred to as a stable improper node.

- If $q > 0$ and $p > 0$ (the first orthant of the plane), then both roots must be of the same sign, and at least one must be positive, or have positive real part. Therefore, both roots must be positive or have positive real part, so the steady state is unstable.

Case 3 If $\Gamma < 0$, i.e. at a point above the curve not on the q axis, then the roots are complex, so the solution depends on the sine and cosine of t , as in equation B.8. In this case, the trajectory is an unstable spiral.

Case 4 If $\Gamma > 0$, i.e. at a point in the first quadrant below the curve, then both roots are real and positive. Here the solution is an unstable improper node.

- If $q > 0 = p$ then the roots are imaginary (i.e. the real part is 0). In this case, the trajectory is an undamped orbit, and the steady state is known as a stable center. This steady state is stable but not asymptotically stable.
- If $q > 0 = \Gamma$ then the two roots are equal. When both are negative (orthant 2) the steady state is known as a stable node; when both are positive (orthant 1) the steady state is an unstable node.

The above list of conditions for the various types of steady states are necessary and sufficient. For example, if $q < 0$ we know that the discriminant is positive, so we know without further calculation that the roots are real.

B.4 Back to the predator-prey model

We now use the results of Section B.3 to study the steady states of the predator-prey model from Section B.2.

	\bar{z}	\hat{z}
$a = F_x$	$-\omega\bar{x}$	$A - 2\omega\hat{x}$
$b = F_y$	$-B\bar{x}$	$-B\hat{x}$
$c = G_x$	$C\bar{y}$	0
$d = G_y$	$-\mu\bar{y}$	$C\hat{x} - D$

Table 1: partial derivatives of predator-prey model

Table 1 collects the partial derivatives of the predator-prey model, evaluated at the interior steady state \bar{z} and the steady state on the x axis, \hat{z} . The variables $\bar{x}, \bar{y}, \hat{x}$ and \hat{y} are the coordinates of the two points.

Evaluated at \bar{z} , the trace is $p = -\omega\bar{x} - \mu\bar{y} < 0$ and the determinant is $q = (\omega\mu + BC)\bar{x}\bar{y}$. Because the trace is negative, at least one root must be negative, and because the determinant is positive, the roots have the same sign. Therefore, both roots are negative. Moreover, the discriminant is positive, so both roots are real. Thus, the interior equilibrium \bar{z} is a stable improper node, i.e. it is asymptotically stable. The Theorem of the First Approximation implies that the interior equilibrium of the original nonlinear system is also asymptotically stable. A problem set asks the student to determine the stability of the the other steady state, \bar{z} .

In the Volterra-Lotka model, where $\omega = \mu = 0$, there is no congestion (as defined above). In the absence of predators, the prey grows without bound. Linearization in the neighborhood of the interior equilibrium shows that there exist imaginary eigenvalues, i.e. ones for which the real part is 0. Based on figure 5, we know that in this case the linear system is stable, but not asymptotically stable, so the Theorem of the First Approximation provides no information about the stability in the original nonlinear system. Clark (pp ??) shows that this system can be solved in closed form.

B.5 More on saddle points

Autonomous optimal control models with a single state variable can often be studied using a two-dimensional phase portrait. A steady state in these

models is always (barring pathologies) a saddle point, so there is a one-dimensional stable manifold – a line in the linear approximation and a curve in the original non-linear system. Trajectories beginning on the stable manifold converge to the steady state, and all others diverge. This point will be important in the context of optimal control, in Chapter xx. Here we provide intuition for why saddle points arise in optimal control settings, and we present some additional concepts.

Consider the optimal control problem where the state variable is the biomass of fish and the control variable is harvest. Part of the data of the problem is an equation of motion for the stock of the fish, a differential equation for the stock, as a function of the stock and the harvest. Using techniques described in Chapter xx below, we can use the necessary condition for optimality to obtain a differential equation that the optimal harvest level must satisfy. This differential equation describes the change in harvest as a function of the current harvest and the current stock. This differential equation is endogenous, in the sense that it emerges from the optimality conditions; in contrast, the differential equation for the stock of fish is exogenous, because it is part of the statement of the optimization problem.

The differential equations for the stock and for the harvest comprise a pair of differential equation, much as in the predator-prey model discussed above. There is an important difference, however. In the predator-prey model both variables are stock variables, and it is natural to regard the values of both variables as predetermined at a point in time. In the optimal control context, the biomass of fish is a state variable, and is predetermined at a point in time; its initial condition is given. In contrast, the harvest level at the initial time is not given; the purpose of the analysis is to determine its value.

In the interest of simplicity, we assume here that there is a unique interior steady state to the pair of differential equations, one describing the change in the stock and the other describing the change in the harvest; moreover, we assume (as is the case in most problems) that it is optimal to drive the biomass of fish and the harvest to that steady state. This information, together with the differential equations, is enough in principal to obtain the optimal solution. Here we provide intuition for the fact that the steady state must be a saddle point.

Consider the alternatives. If all trajectories diverge from the steady state, then there is obviously no trajectory leading to the steady state. That

conclusion is inconsistent with our statement that it is optimal to drive the system to the steady state. If all trajectories converge to the steady state, then the optimality conditions, transformed into a differential equation for harvest, are useless in identifying the optimal trajectory – simply because all trajectories converge. The remaining possibility is that one trajectory converges and all others diverge. But this is precisely what it means for the steady state to be a saddle point.

Our discussion of saddle points in Section B.3 noted that the dimension of the stable manifold equals the number of stable eigenvalues. In the two-dimensional system arising from the scalar optimal control problem, the statement that the steady state is a saddle point means that there is one stable eigenvalue, i.e. the dimension of the stable manifold is 1: a curve in biomass-harvest space. This stable manifold has another name: it is the graph of the control rule, the function that gives the optimal harvest as a function of the biomass of fish.

We do not mean to give the reader the impression that all interesting economic problems, or even all two-dimensional problems, involve saddle points. For example, there is a large literature in macroeconomics on the indeterminacy of competitive equilibria. In a class of these models, there is a unique steady state, but that steady state is a stable node, meaning that there are infinitely many paths that converge to it. In this case there are unaccountably many rational expectations competitive equilibria. This situation can arise in a variety of circumstances, and is often associated with models containing strong non-convexities.

Two-dimensional models with a saddle point contain a pair of *separatrices*. One *separatrix* is the trajectory that converges to the steady state; this is the graph of the optimal control rule. The other separatrix is a trajectory that emanates from the saddle point. These two separatrices divide the phase plane into four regions. Because trajectories do not cross, and because each separatrix is a trajectory, no trajectory crosses a separatrix. Therefore, a trajectory emanating from a point inside one of the four regions, i.e. a point not on a separatrix, remains in that region. A point on the separatrix remains on that separatrix, either moving toward the saddle point or moving away from it. The regions created by the separatrices are not the same as the isosectors, created by the 0-isoclines. Trajectories do cross isoclines.

In the two-dimensional linear model, the separatrices are straight lines through

the saddle point. We already noted that the stable manifold – one of the two separatrices – is a line, and the same is true of the other separatrix. In a higher dimensional linear model with a saddle point, the convergent separatrix is the hyperplane spanned by the eigenvectors associated with the stable eigenvalues.

The fact that the separatrices are straight lines gives an alternative way of computing them, that bypasses the need to find the eigenvalues and eigenvectors. Consider the linear system in equation B.10. Each separatrix is a straight line, so it can be written as $Y = sX$ for some constant s , so $\frac{dY}{dX} = s$ on a separatrix. On any trajectory we also have

$$\frac{\dot{Y}}{\dot{X}} = \frac{dY}{dX}.$$

On a separatrix, we can rewrite the original system as

$$\dot{X} = aX + bsX \quad \text{and} \quad \dot{y} = cX + dsX.$$

Using these equations we have

$$\begin{aligned} \frac{dY}{dX} = s &= \frac{(c + ds)X}{(a + bs)X} = \frac{(c + ds)}{(a + bs)} \implies \\ s(a + bs) &= c + ds \implies \\ bs^2 + (a - d)s - c &= 0. \end{aligned} \tag{B.11}$$

We can solve this quadratic to get the two values of s , the slopes of the two separatrices. We can then identify which is the stable and which is the unstable separatrix using the fact that a trajectory on the stable path converges to the steady state, and a trajectory on the unstable path diverges. On both paths, the variable X obeys $\dot{X} = (a + bs)X$. One root of equation B.11 satisfies $(a + bs) < 0$. Denote this root as s^* , the root associated with the stable saddle path. The other root satisfies $(a + bs) > 0$; this root is the slope of the unstable separatrix.

Of course, if this linear system is the result of taking an approximation of a non-linear system, then the straight-line separatrices of the former are merely an approximation of the curved separatrices of the latter. The slope of the approximation equals the tangent to the separatrices of the original system, evaluated at the steady state where the approximation is performed.

Although the information above may appear abstract, it has practical value. If the two dimensional non-linear system is derived from a one-dimensional optimal control problem, we have described a simple means of obtaining an approximation of the optimal control rule. Recall that we defined X and Y as deviations from a steady state value. If x is the biomass of fish, and y is the harvest, then X and Y are the deviations of the stock and harvest from their steady state values. Setting the two original, nonlinear, differential equations equal to 0 gives two algebraic equations in X and Y . Solving these equations gives the (or a) steady state. The linear approximation of this system at the (or a) steady state gives the parameters of the linear system, a, b, c and d . Using these values we obtain the stable root, s^* . Our linear approximation of the optimal control rule is $y = s^*X$.

The reader should notice that the algorithm sketched here for approximating an optimal control rule parallels the algorithm that we illustrated, for a problem in discrete time, in Section ... (the last section in Chapter 1). See the problem set at the end of this chapter.

B.6 Examples of dynamic systems in economics

Here we use two examples that illustrate the power of the methods described in this chapter. The first example, due to Brander and Taylor (1998) uses a model similar to the predator-prey model to study population dynamics in a resource-constrained economy. The second example, due to Krugman (1991) uses a linear model to illustrate the possibility of multiple rational expectations equilibria; for a set of initial conditions there are different rational expectations competitive that take the economy to different steady states. (Benabou and Fukau (1993) correct a technical mistake in Krugman's analysis.) We then consider a third example (Karp and Paul 20xx), related to the first two, which indicates the kinds of complexities that arise in higher dimensional systems, and shows how the intuition obtained from lower dimensional systems may not be robust.

B.6.1 Population dynamics on Easter Island

A simple general equilibrium model that includes a renewable resource describes the rise and fall of an isolated civilization, such as the one on Easter

Island.¹⁴⁸ The economy produces two goods and has Ricardian technology. The environment, a stock variable, affects the marginal productivity of labor in one sector. The other state variable is the stock of human population. An increase in the environment encourages growth of human population because it becomes easier to get food. An increase in human population degrades the environment, since extraction increases.

We can re-interpret the predator-prey model described in Section B.2, treating the environment as the prey and the human population as the predator. The analysis proceeds by linearizing dynamics around the interior steady state and checking whether the roots are real or imaginary. The authors show that a stable interior equilibrium can be either an improper node (real roots) with monotonic adjustment or a spiral node (imaginary roots), with trajectories spiralling into the steady state. Spirals occur if the intrinsic growth rate of the environment is sufficiently small. Using plausible parameter values, they show that it is possible to get a trajectory in which human population increases for a time, then crashes, eventually reaching a low level. [Details to be added later.]

B.6.2 Multiple rational expectations equilibria

The economy consists of two sectors, Agriculture and Manufacturing. There is fixed stock of a single input, labor, normalized to 1. There is full employment, so if L units of labor are in the Manufacturing sector, then $1 - L$ are in Agriculture. Output prices are fixed, so this model represents a small open economy. There are constant returns to scale in Agriculture, with the value of marginal productivity equal to c . There are increasing returns to labor in Manufacturing, where the value of marginal product equals $a + bL$. In both sectors the wage equals the value of marginal product. Costly adjustment between sectors makes it possible, outside of a steady state, for there to be workers in both sectors even though the wage is different in the two sectors. In a steady state either one sector must close down or the wages in the two sectors must be the same.

The flow of migrants into Manufacturing is \dot{L} , which is negative if workers are leaving the sector and positive if workers are entering the sector. The

¹⁴⁸Later add a note discussing the controversy about Easter Island, see Jared Diamond and critics of his interpretation.

total cost of migration is $\frac{\gamma}{2} (\dot{L})^2$. Here we have a model of convex adjustment costs, similar to that considered in Chapter 1. Each migrant pays the marginal cost, which equals $\gamma \left| \dot{L} \right|$. The reason for the absolute value sign will be apparent shortly. All parameters are positive. The wage differential at a point in time is $m(L) = a - c + bL$. We assume that

$$c > a \text{ and } a + b > c. \quad (\text{B.12})$$

The first inequality means that if there are no workers in Manufacturing, the wage in Agriculture is greater than the wage obtained by a single worker who moves to Manufacturing. The second equality means that if all workers are in Manufacturing, the wage in that sector is greater than the wage obtained by a single worker who moves to Agriculture. These parameter assumptions imply that there are two stable steady states: $L = 0$ and $L = 1$. If all workers are in a particular sector, no worker has an incentive to leave that sector.

We now consider the representative worker's migration decision. Suppose that all migration stops at time T ; this variable could be infinite, but in this model turns out to be finite. At or after time T migration for a single worker is costless, because the price of migration is $\left| \dot{L} \right| = 0$. By moving from Agriculture to Manufacturing at time t rather than at time T a worker obtains the present value of the wage differential

$$\chi_t = \int_t^T e^{-r(\tau-t)} m(\tau) d\tau, \quad (\text{B.13})$$

where r is the worker's discount rate. This quantity could be either negative or positive. Differentiating, we have

$$\dot{\chi} = r\chi - m. \quad (\text{B.14})$$

Equation B.14 is an example of a no-arbitrage relation that appears throughout economics. We can think of χ_t as the value or price of the "asset": being in the Manufacturing sector at time t . Therefore, $\dot{\chi}$ is the capital gain of this asset. The wage differential m is analogous to a dividend: it equals the payment that the asset holder obtains at time t . If the price of the asset is χ , and the opportunity cost of income per unit of time is r , then the opportunity cost of holding the asset is $r\chi$. The equation states that the capital

gains plus dividend, $\dot{\chi} + m$, equals the opportunity cost of holding the asset $r\chi$.

The variable χ is sometimes referred to as a *jump variable*. Unlike physical stocks, whose values at a point in time are predetermined, the initial condition of a jump variable is endogenous; it depends on what happens in the future. Notice from equation B.13 that the value of χ_t depends on future values of m .

An individual takes prices and decisions of all agents, both current and future, as given. In an equilibrium with non-zero migration, each individual who migrates must be indifferent between migration and remaining in the same sector. This indifference requires (when $\dot{L} \neq 0$)

$$\gamma\dot{L} = \chi, \text{ or } \dot{L} = \frac{\chi}{\gamma},$$

i.e. a migrant pays exactly what it is worth to migrate. For example, if $\dot{L} > 0$ then workers are moving into the Manufacturing sector, and χ is positive. If workers are moving into Agriculture ($\dot{L} < 0$) migrants pay $-\gamma\dot{L}$ and obtain the asset “not being in Manufacturing”, whose value is $-\chi$.

In matrix form, the equilibrium conditions for this model are

$$\begin{pmatrix} \dot{\chi} \\ \dot{L} \end{pmatrix} = \begin{bmatrix} r & -b \\ \frac{1}{\gamma} & 0 \end{bmatrix} \begin{pmatrix} \chi \\ L \end{pmatrix}. \quad (\text{B.15})$$

The problem set asks the reader to confirm several features of this model, which we summarize here. There is an interior equilibrium ($0 < L_* < 1$) which is always unstable, and two stable boundary equilibria, $L = 0$ and $L = 1$. (We already explained the reason for the boundary equilibria.) The interior equilibrium is an unstable spiral if

$$r^2\gamma - 4b < 0. \quad (\text{B.16})$$

Figure 6 shows the phase portrait in this case. One equilibrium trajectory spirals counter-clockwise away from the unstable steady state, eventually hitting the stable steady state $L = 0$. The figure shows the largest value of L , denoted L_2 , on this spiral. A second equilibrium trajectory spirals clockwise away from the unstable steady state, eventually reaching the stable steady state $L = 1$. The smallest value of L on this spiral is shown as L_1 . There are uncountably many non-equilibrium trajectories, not shown, that satisfy

equation B.15. The problem set asks the reader to discuss the equilibrium condition that eliminates all of the spirals not shown. For initial conditions $L_0 \in (L_1, L_2)$ there are many rational expectations equilibria. One family of equilibria takes the economy to $L = 1$ and the other family takes the economy to $L = 0$. For each family, there is one monotonic trajectory, the one that starts on the outer-most spiral, and many trajectories that involve non-monotonic adjustment.

If inequality B.16 is reversed, the unstable steady state is an unstable node; see Figure 7. In that case, all trajectories are monotonic and there is a unique equilibrium from all initial conditions except for the knife-edge case where $L_0 = L_*$. If the initial condition is $L_0 < L_*$ the economy eventually specializes in Agriculture, and if the inequality is reversed, the economy eventually specializes in Manufacturing.

Inequality B.16 has an economic interpretation. The inequality holds if the parameter that determines the extent of non-convexity, b , is large relative to the discount rate, r , and the cost-of-adjustment parameter, γ . When the inequality holds, “expectations matter”, in the sense that the equilibrium that results depends on what agents believe will happen in the future. When agents are patient (r is small) or when it is inexpensive to migrate (γ is small), or when the degree of increasing returns is high (b is large) the optimal decision of an agent today is sensitive to what she thinks other agents will do in the future, because those future actions determine the value of being in a particular sector. In these circumstances, the optimal decision today is sensitive to the agent’s beliefs about aggregate decisions in the future. In contrast, if the agent is impatient, or adjustment costs are large, or the degree of increasing returns is modest, current decisions are rather insensitive to beliefs about the future. In this case, the rational expectations equilibrium is unique and depends only on “history”, which determines the current level of L .

A final point is worth emphasizing. In all cases in this model, there are two steady states, with all labor employed either in Agriculture or in Manufacturing. However, when inequality B.16 is satisfied there are multiple equilibrium trajectories given initial conditions $L_0 \in (L_1, L_2)$, some of these equilibrium trajectories approach one steady state, and other trajectories approach the other steady state. In contrast, when inequality B.16 is not satisfied there is generically a unique equilibrium trajectory from any initial condition. The fact that there are multiple steady states does not mean

that there are multiple equilibria. Multiple equilibria exist if and only if there is more than one equilibrium trajectory emanating from a given initial condition.

B.6.3 History and expectations with two stock variables

Here we consider a model with two stock variables. In one interpretation of this model, there are constant returns to scale in both Agriculture and Manufacturing. However, Manufacturing creates pollution, a stock variable, which decreases productivity in the agricultural sector, and there are convex migration costs (as in Section B.6.2). In this model, the manufacturing wage is constant, but the agricultural wage depends on the stock of pollution. The wage differential depends not on the current allocation of labor, but on the entire history of the past allocation: more labor in Manufacturing in the past means higher past pollution flows, which create a higher current stock of pollution and lower current agricultural productivity and wage.

In the second interpretation of this model, there is no pollution stock, but a larger Manufacturing work force creates a higher flow of knowledge in the Manufacturing sector, leading to a higher stock of knowledge in the future, and thus higher future Manufacturing wage. Here, there are dynamic increasing returns to scale in Manufacturing, unlike the static increasing returns in Section B.6.2 where the manufacturing wage depends only on current, not on past allocations of labor.

Despite their apparent differences, these two models are formally equivalent. In both cases increased labor in Manufacturing today contributes to a stock variable (pollution, which is bad of Agriculture, or knowledge, which is good for Manufacturing) that increases the Manufacturing wage differential. In the interest of simplicity, we consider here only the environmental interpretation of the model.

The model contains three variables, the two stock variables and a forward-looking or jump variable (analogous to χ in Section B.6.2). Graphical methods are of little use in this context. A linear version of the model is a straightforward extension of the model in Section B.6.2, obtained by including the additional stock variable. As in the simpler one-state model, in the two-state model there is an interior unstable steady state and all steady

states are on the boundary, i.e. all labor ends up in either Manufacturing or Agriculture.

As with the one-state model, there can be a unique equilibrium. Just as in the one-state model, a lower cost of migration (γ in the one-state model) increases the measure of the set of other parameter values for which there are multiple equilibria. Moreover, in the one-state model, a lower cost of migration increases the measure of the set of initial conditions ($L_2 - L_1$) for which there are multiple equilibria. In this sense, multiplicity is “more likely” for lower costs of adjustment. In contrast, in the two-state model, a decrease in migration costs can decrease the measure of the set of initial conditions (a two-dimensional set) for which there are multiple equilibria. In this sense, low costs of adjustment make multiplicity less likely.

To understand this reversal, consider what happens if we hold all parameters constant except that we reduce the parameter that determines migration costs (γ , above). This reduction makes possible large changes in labor allocation over a short period of time. However, this parameter change does not alter the function determining the dynamics of the environment (although of course the actual trajectory does change). Even though labor adjusts quickly, the environment remains sluggish. In this circumstance the future wage differential depends – at least for an “appreciable period of time” – primarily on the predetermined environmental variable, because in the future labor will adjust quickly in response to wage differentials. Thus, it is more likely (with the lower adjustment costs) that the equilibrium depends on history, via the predetermined stock variables, than on beliefs about the future. That is, lower adjustment costs make multiplicity of equilibria “less likely” in the two-state model, whereas the opposite occurs in the one-state model.

B.7 Problems

1. (Section B.4) Confirm that the steady state \hat{z} is a saddle point, and identify the separatrix.
2. (Section B.5) Elaborate on the parallels in the method of approximating the control rule in the discrete time problem discussed at the end of Chapter 1, and the method of approximating the control rule of the continuous time optimal control rule in Section B.5.
3. Verify the description of the equilibrium to the rational expectations model in Section B.6.2. (a) Confirm the directional arrows in Figures 6 and 7. (b) Show that inequality B.16 determines whether the unstable equilibrium is a spiral or a node. (c) Notice that all the equilibria in either figure intersects the boundary ($L = 0$ or $L = 1$) at $\chi = 0$. Explain why a competitive equilibrium must have this boundary condition. (d) If you were given numerical values for the parameters, how would you numerically determine the values of L_1 and L_2 in the case where inequality B.16 holds. (Hint: think “backwards”.)