Chapter 8

Discrete Time Continuous State Dynamic Models: Theory

We now turn our attention to discrete time dynamic economic models whose state spaces are closed convex sets in Euclidean space. Three classes of discrete time, continuous state dynamic economic models are examined. One class includes models of centralized decisionmaking by individuals, firms, or institutions. Examples include a central planner managing the harvest of a natural resource so as to maximize social welfare, an entrepreneur planning production and investment so as to maximize the present value of her firm, and a consumer making consumption and savings decisions so as to maximize his expected lifetime utility.

A second class of discrete time continuous state dynamic model examined includes models of strategic gaming between a small number of individuals, firms, or institutions. Dynamic game models attempt to capture the behavior of a small group of dynamically optimizing agents when the policy pursued by one agent directly affects the welfare of another. Examples include a two national grain marketing boards deciding quantities of grain to sell on world markets and two individuals deciding how much to work and invest in the presence of co-insurance.

A third class of discrete time continuous state dynamic economic model examined includes partial and general equilibrium models of collective, decentralized economic behavior. Dynamic equilibrium models characterize the behavior of a market, economic sector, or entire economy through intertemporal arbitrage conditions that are enforced by the collective action of atomistic dynamically optimizing agents. Often the behavior of agents at a given date depends on their expectations of what will happen at a future date. If it is assumed that agent's expectations are consistent with the implications of the model as a whole, then agents are said to possess rational expectations. Examples of rational expectations models include arbitrage pricing models for financial and physical assets.

Dynamic optimization and equilibrium models are closely related. The solution to a continuous state dynamic optimization may often be equivalently characterized by first-order intertemporal equilibrium conditions obtained by differentiating Bellman's equation. Conversely, many dynamic equilibrium problems can be "integrated" into equivalent optimization formulations. Whether cast in optimization or equilibrium form, most discrete time continuous state dynamic economic models pose infinite-dimensional fixed-point problems that lack closed-form solution. This chapter provides an introduction to the theory of discrete time continuous state dynamic economic models. The subsequent chapter is devoted to numerical methods that may be used to solve and analyze such models.

8.1 Continuous State Dynamic Programming

The discrete time, continuous state Markov decision model has the following structure: In every period t, an agent observes the state of an economic process s_t, takes an action x_t, and earns a reward $f(s_t, x_t)$ that depends on both the state of the process and the action taken. The state space $S \in \Re^n$, which contains all the states attainable by the process, is a closed convex nonempty set. The action space $X \in \Re^m$, which contains all actions that may be taken by the agent, may be either a finite set or a closed convex set. The state of the economic process follows a controlled Markov probability law. Specifically, the state of the economic process in period $t+1$ will depend on the state and action in period t and an exogenous random shock ϵ_{t+1} that is unknown in period t:

$$s_{t+1} = g_t(s_t, x_t, \epsilon_{t+1}).$$

The agent seeks a policy $\{x_t^*\}_{t=1}^T$ of state-contingent actions $x_t = x_t^*(s_t)$ that will maximize the present value of current and expected future rewards over time, discounted at a per-period factor δ:

$$E \sum_t \delta^t f(s_t, x_t).$$

Discrete time, continuous state Markov decision models also may have a finite or infinite horizon, and may be stochastic or deterministic. If the model is stochastic, the exogenous random shocks ϵ_t are assumed identically distributed over time, mutually independent, and independent of past states and actions. Also, δ is assumed to be less than one.

Continuous space Markov decision models may be classified according to the cardinality of their action spaces. If the action space is finite, the model is said to be a discrete choice model. If the action space is a closed convex set, the model is said to be a continuous choice model. In some instances, the set of actions available to the agent may vary with the state of the process s_t. In such cases, the restricted action space will be denoted $X(s_t)$.

Like the discrete Markov decision problem, the discrete time continuous state Markov decision problem may be analyzed using dynamic programming methods based on Bellman's Principle of Optimality. The Principle of Optimality applied to the discrete time continuous state Markov decision model yields Bellman's recursive functional equation:

$$V_t(s) = \max_{x \in X(s)} \{f(s,x) + \delta E_\epsilon V_{t+1}(g(s,x,\epsilon))\}, \qquad s \in S.$$

Here, the value function $V_t(s)$ specifies the maximum attainable sum of current and expected future rewards, given that the process is in state s in t.

For the finite horizon discrete time continuous state Markov decision model to be well posed, a post-terminal value function V_{T+1} must be specified by the analyst. The post-terminal value function is fixed by some economically relevant terminal condition. In many applications, V_{T+1} will be identically zero, indicating that no rewards are earned by the agent beyond the terminal decision period; while in other applications, V_{T+1} may specify a salvage value earned by the agent after making his final decision. Given the post-terminal value function, the finite horizon discrete time continuous state Markov decision model may be solved recursively, at least in principle, by repeated application of Bellman's equation: Having V_{T+1}, solve for $V_T(s)$ for all states s; having V_T, solve for $V_{T-1}(s)$ for all states s; having V_{T-1}, solve for $V_{T-2}(s)$ for all states s; and so on, until $V_1(s)$ is derived for all states s.

The value function of the infinite horizon discrete time continuous state Markov decision model will the same for every period and thus may be denoted simply by V. The infinite horizon value function V is characterized as the solution to the Bellman functional fixed-point equation

$$V(s) = \max_{x \in X(s)} \{f(s, x) + \delta E_\epsilon V(g(s, x, \epsilon))\}, \qquad s \in S.$$

If the discount factor δ is less than one and the reward function f is bounded, the mapping underlying Bellman's equation is a strong contraction on the space of bounded continuous functions and, thus, by The Contraction Mapping Theorem, will possess an unique solution.

In practice, however, solving the Bellman equation for either the finite or infinite horizon discrete-time continuous state Markov decision problem is not analytically straightforward. The problem lies with the state space, which contains an infinite number of points. Except in rare special cases, it is not possible to derive analytically an explicit closed form expression for the period t value function of the finite horizon model, even if the period $t+1$ value function is known and possesses a closed-form. Thus, in the typical case, solving the Bellman's equation requires explicitly solving an infinite number of optimization problems, one for each state. This is an impracticable task. As a rule, one can only solve a discrete time continuous state Bellman equation numerically, a matter that we take up the following chapter.

8.2 Euler Equilibrium Conditions

Like many optimization problems, the solution to a continuous state continuous choice Markov decision problem can often be characterized by "first-order" equilibrium conditions. Characterizing the solution to a Markov decision problem through its equilibrium conditions, widely called the Euler conditions, serves two purposes. First, the Euler conditions admit an intertemporal arbitrage interpretation that help the analyst understand and explain the essential features of the optimized dynamic economic process. Second, the Euler conditions can, in many instances, be solved more easily than Bellman's equation for the optimal solution of the Markov decision model.

The equilibrium conditions of the continuous state continuous choice Markov decision problem involve, not the value function, but its derivative

$$\lambda_t(s) = V_t'(s).$$

We call λ_t the shadow price function. It represents the value of the marginal unit of state variable to the optimizer or, equivalently, the price that the optimizer imputes to the state variable.

Assume that both the state and action spaces are closed convex nonempty sets ant that the reward functions f and the state transition functions g are continuously differentiable of all orders there. The equilibrium conditions for discrete time continuous state continuous choice Markov decision problem are derived by applying the Karush-Kuhn-Tucker and Envelope Theorems to the optimization problem embedded in Bellman's equation. Assuming actions are unconstrained, the Karush-Kuhn-Tucker conditions for the embedded unconstrained optimization problem imply that the optimal action x, given state s in period t, satisfies the equimarginality condition:

$$f_x(s,x) + \delta E_\epsilon \left[\lambda_{t+1}(g(s,x,\epsilon)) g_x(s,x,\epsilon)\right] = 0.$$

The Envelope Theorem applied to the same problem implies:

$$\lambda_t(s) = f_s(s,x) + \delta E_\epsilon \left[\lambda_{t+1}(g(s,x,\epsilon)) \cdot g_s(s,x,\epsilon)\right].$$

Here, f_x, g_x, f_s, and g_s denote partial derivatives. If the horizon is infinite, then the shadow price functions λ_t will be the same in every period and the time subscript may be deleted.

The Euler conditions take a different form when actions are subject to constraints. Suppose, for example, that feasible actions are described by the constrained action set

$$X(s) = \{x | h(s,x) \leq 0\},$$

where h is a smooth function on S and X. Then the Karush-Kuhn-Tucker and Envelope theorems imply that the optimal action x, given state s in period t, satisfies the complementarity conditions:

$$f_x(s,x) + \delta E_\epsilon \left[\lambda_{t+1}(g(s,x,\epsilon)) \cdot g_x(s,x,\epsilon)\right] - \mu h_x(s,x) = 0$$

$$h(s,x) \perp \mu \geq 0$$

$$\lambda_t(s) = f_s(s,x) + \delta E_\epsilon \left[\lambda_{t+1}(g(s,x,\epsilon)) \cdot g_s(s,x,\epsilon)\right] - \mu h_s(s,x)$$

where μ is the shadow price associated with the inequality constraints at the optimal action for the given state. Similar conditions can be obtained when equality constraints are present.

In many applications, the constraints on the actions are simple bounds of the form

$$X(s) = \{x \mid a(s) \leq x \leq b(s)\},$$

where a and b are differentiable functions of the state s. In these instances, the Euler conditions simplify to the complementarity conditions:

$$f_x(s,x) + \delta E_\epsilon \left[\lambda_{t+1}(g(s,x,\epsilon)) \cdot g_x(s,x,\epsilon)\right] \perp a(s) \leq x \leq b(s)$$

$$\lambda_t(s) = f_s(s,x) + \delta E_\epsilon \left[\lambda_{t+1}(g(s,x,\epsilon)) \cdot g_s(s,x,\epsilon)\right] + \mu^- a'(s) + \mu^+ b'(s).$$

where

$$\mu = f_x(s,x) + \delta E_\epsilon \left[\lambda_{t+1}(g(s,x,\epsilon)) \cdot g_x(s,x,\epsilon)\right]$$

is the marginal value of increasing the action level. The positive and negative parts of μ, $\mu^+ = \max(0,\mu)$ and $\mu^- = \min(0,\mu)$, are the shadow prices of the upper and lower bounds, respectively.

An analyst is often interested with the longrun tendencies of the optimized process. If the model is deterministic, it may possess a well-defined steady-state to which the process will converge over time. The steady-state is characterized by the solution to a nonlinear equation. More specifically, the steady-state of an unconstrained deterministic problem, if it exists, consists of a state s^*, an action x^*, and shadow price λ^* such that

$$f_x(s^*,x^*) + \delta\lambda^* g_x(s^*,x^*) = 0$$

$$\lambda^* = f_s(s^*,x^*) + \delta\lambda^* g_s(s^*,x^*)$$

$$s^* = g(s^*,x^*).$$

The steady-state of a constrained deterministic dynamic optimization problem can be similarly stated, except that it takes the form of a nonlinear complementarity problem, rather than a nonlinear equation.

Knowledge of the steady-state of a deterministic Markov decision problem is often very useful. For most well-posed deterministic problems, the optimized process will converge to the steady-state, regardless of initial condition. For this reason, the analyst will often be satisfied to understand the dynamics of the process around the steady-state. The steady-state conditions, moreover, are equations or complementarity conditions that can be analyzed algebraically. In particular, the derivative of the longrun value of an endogenous variable with respect to model parameters can often be derived using standard differential calculus, even if the dynamic model itself lacks a closed-form solution.

If the discrete time continuous state model is stochastic, the model will not converge to a specific state and action and the longrun behavior of the model can only be described probabilistically. In these cases, however, it is often practically useful to derive the steady-state of the deterministic "certainty-equivalent" problem obtained by fixing all exogenous random shocks at their respective means. Knowledge o the certainty-equivalent steady-state can assists the analyst by providing a reasonable initial guess for the optimal policy, value, and shadow price functions in iterative numerical solution algorithms. Also, one can often solve a hard stochastic dynamic model by first solving the certainty-equivalent model, and then solving a series of models obtained by gradually perturbing the variance of the shock from zero back to its true level, always using the solution of one model as the starting point for the algorithm used to solve the subsequent model.

8.3 Linear-Quadratic Control

Before proceeding to more complicated continuous state Markov decision models we discuss a special case: the linear-quadratic control model. The linear quadratic control problem is a Markov decision model with a quadratic reward function

$$f(s, x) = F_0 + F_s s + F_x x + 0.5 s' F_{ss} s + s' F_{sx} x + 0.5 x' F_{xx} x$$

and a linear state transition function with additive shock

$$g(s, x, \epsilon) = G_0 + G_s s + G_x x + \epsilon.$$

Here, s is an n-by-1 state vector, x is an m-by-1 action vector, F_0 is a known constant, F_s is a known n-by-1 vector, F_x is a known m-by-1 vector, F_{ss} is a known n-by-n matrix, F_{sx} is a known n-by-m matrix, F_{xx} is a known m-by-m matrix, G_0 is a known constant, G_s is a known n-by-1 vector, and G_x is a known m-by-1 vector. Without loss of generality, the shock ϵ is assumed to have a mean of zero. The linear-quadratic control problem admits no constraints on the action.

The linear-quadratic is of special importance because it is one of the few discrete time continuous state Markov decision models with known analytic solution. By a conceptually simple but algebraically burdensome proof omitted here, one can show that the solution to the infinite horizon linear quadratic control model takes a particularly simple form. Specifically, both the optimal policy and shadow price functions are linear in the state variable:

$$x(s) = X_0 + X_s s$$

$$\lambda(s) = \Lambda_0 + \Lambda_s s.$$

Here, X_0 is an m-by-1 vector, X_s is an m-by-n matrix, Λ_0 is an n-by-1 vector, and Λ_s is an n-by-n matrix.

The parameters of the shadow price function are characterized by the nonlinear vector-fixed point equations

$$\begin{aligned}
\Lambda_0 &= -[\delta G_s'\Lambda_s G_x + F_{sx}][\delta G_x'\Lambda_s G_x + F_{xx}']^{-1}[\delta G_x'[\Lambda_s G_0 + \Lambda_0] + F_x] \\
&\quad +\delta G_s'[\Lambda_s G_0 + \Lambda_0] + F_s \\
\Lambda_s &= -[\delta G_s'\Lambda_s G_x + F_{sx}][\delta G_x'\Lambda_s G_x + F_{xx}']^{-1}[\delta G_x'\Lambda_s G_s + F_{sx}'] \\
&\quad +\delta G_s'\Lambda_s G_s + F_{ss}.
\end{aligned}$$

These fixed-point equations can typically be solved in practice using a simple function iteration scheme. One first solves for Λ_s by applying function iteration to the second equation, and then solves for Λ_0 by applying function iteration to the first equation. Once the parameters of the shadow price function have been computed, one can easily compute the parameters of the optimal policy:

$$X_0 = -[\delta G_x'\Lambda_s G_x + F_{xx}']^{-1}[\delta G_x'[\Lambda_s G_0 + \Lambda_0] + F_x]$$

$$X_s = -[\delta G_x'\Lambda_s G_x + F_{xx}']^{-1}[\delta G_x'\Lambda_s G_s + F_{sx}']$$

The relative simplicity of the linear quadratic control problem derives from the fact that the optimal policy and shadow price functions are known to belong to a finite parameter family. The parameters, moreover, are characterized as the solution to a well-defined nonlinear vector fixed-point equation. Thus, the apparently infinite dimensional Euler functional fixed-point equation is converted into finite dimensional vector fixed-point equation that may be solved using standard nonlinear equation solution methods. This simplification, unfortunately, is not generally possible for other types of discrete time continuous state Markov decision models.

A second simplifying feature of the linear-quadratic control problem is that the shadow price and optimal policy functions do not depend on the distribution of the state shock. This is known as the certainty-equivalence property of the linear quadratic control problem. It asserts that the solution of the stochastic problem is the same as the solution of the deterministic problem obtained by fixing the state shock ϵ at its mean of zero. Certainty equivalence also is not a property of more general discrete time continuous state Markov decision models.

Because linear quadratic control models are easy to solve, many analysts compute approximate solutions to more general Markov decision models using the method of linear quadratic approximation. Linear quadratic approximation calls for all constraints of the general problem to be discarded, for its reward function to be replaced with its second-order quadratic approximation about the steady-state

$$\begin{aligned} f(s,x) \;\approx\; & f_0 + f_s(s-\bar{s}) + f_x(x-\bar{x}) + 0.5(x-\bar{x})' f_{xx}(x-\bar{x}) \\ & +0.5(s-\bar{s})' f_{ss}(s-\bar{s}) + (s-\bar{s})' f_{sx}(x-\bar{x}) \end{aligned}$$

and for its state transition function to be replaced with its first-order, certainty-equivalent linear approximation about the steady state

$$g(s,x,\epsilon) \approx \bar{s} + g_s(s-\bar{s}) + g_x(x-\bar{x}).$$

Here, $\bar{s}$, $\bar{x}$, and $\bar{\lambda}$ denote the steady-state state, action, and shadow price of the original model, f_0 represents the value of the of the reward at the steady state, and f_s, f_x, f_{ss}, f_{sx}, f_{xx}, g_s, and g_x, represent derivatives of the reward and state transition functions at the steady state.

Exploiting properties of the derivative of the reward and state transition functions at the steady state, the solution to the linear quadratic approximation may be simplified to

$$\lambda(s) = \bar{\lambda} + \Lambda_s(s-\bar{s})$$

$$x(s) = \bar{x} + X_s(s-\bar{s})$$

where

$$\begin{aligned}
\Lambda_s &= -[\delta g_s' \Lambda_s g_x + f_{sx}][\delta g_x' \Lambda_s g_x + f_{xx}']^{-1}[\delta g_x' \Lambda_s g_s + f_{sx}'] \\
&\quad + \delta g_s' \Lambda_s g_s + f_{ss} \\
X_s &= -[\delta g_x' \Lambda_s g_x + f_{xx}']^{-1}[\delta g_x' \Lambda_s g_s + f_{sx}']
\end{aligned}$$

These equations can be easily solved using function iteration methods.

8.4 Economic Examples

8.4.1 One Sector Optimal Growth

Consider an economy comprising a single composite good. Each year t begins with a predetermined amount of the good s_t, of which an amount x_t is invested and the remainder is consumed. The social welfare derived from consumption in year t is $u(s_t - x_t)$. The amount of good available in year $t+1$ is $s_{t+1} = \gamma x_t + \epsilon_{t+1} f(x_t)$ where γ is the capital survival rate, f is the aggregate production function, and ϵ_{t+1} is a positive production shock with mean 1. What consumption-investment policy maximizes the sum of current and expected future welfare over an infinite horizon?

This is an infinite horizon, stochastic model with time $t \in \{0, 1, 2, \ldots\}$ measured in years. The model has a single state variable

$$\begin{aligned}
s_t &= \text{stock of good at beginning of year } t \\
s_t &\in [0, \infty)
\end{aligned}$$

and a single action variable

$$x_t = \text{amount of good invested in year } t$$

subject to the constraint

$$0 \leq x_t \leq s_t.$$

The reward earned by the optimizing agent is

$$u(s_t - x_t) = \text{social utility in } t.$$

State transitions are governed by

$$s_{t+1} = \gamma x_t + \epsilon_{t+1} f(x_t)$$

where

$$\epsilon_t = \text{productivity shock in year } t.$$

The value function

$$V(s) = \text{sum of current and expected future social welfare}$$

satisfies Bellman's equation

$$V(s) = \max_{0 \leq x \leq s} \{u(s-x) + \delta EV(\gamma x + \epsilon f(x))\}, \qquad s > 0.$$

Assuming $u'(0) = -\infty$ and $f(0) = 0$, the solution to Bellman's equation will always be internal, and the shadow price function

$$\lambda(s) = \text{shadow price of stock}$$

satisfies the Euler equilibrium conditions:

$$u'(s-x) - \delta E\left[\lambda(\gamma x + \epsilon f(x)) \cdot (\gamma + \epsilon f'(x))\right] = 0$$

$$\lambda(s) = u'(s-x).$$

Thus, along the optimal path,

$$u'_t = \delta E_t\left[u'_{t+1} \cdot (\gamma + \epsilon_{t+1} f'_t)\right]$$

where u'_t is marginal utility and $\epsilon_{t+1} f'_t$ is the ex-post marginal product of capital. That is, on the margin, the utility derived from a unit of good today must equal the discounted expected utility derived from investing the good and consuming it and its product tomorrow.

The certainty-equivalent steady-state is obtained by fixing ϵ, its mean 1. The certainty-equivalent steady-state stock of good s^*, investment level x^*, and shadow price λ^* are characterized by the nonlinear equation system

$$u'(s^* - x^*) = \delta \lambda^* (\gamma + f'(x^*))$$

$$\lambda^* = u'(s^* - x^*)$$

$$s^* = \gamma x^* + f(x^*).$$

The certainty-equivalent steady-state conditions imply the golden rule: $1 - \gamma + r = f'(x^*)$. That is, in deterministic steady-state, the marginal product of capital equals the capital depreciation rate plus the interest rate. Totally differentiating the equation system above with respect to the interest rate r:

$$\frac{\partial s^*}{\partial r} = \frac{1+r}{f''} < 0$$

$$\frac{\partial x^*}{\partial r} = \frac{1}{f''} < 0$$

$$\frac{\partial \lambda^*}{\partial r} = \frac{u'' r}{f''} > 0.$$

That is, a one-time rise in the interest rate will reduce the deterministic steady-state supply and investment, and will raise the shadow price.

8.4.2 Nonrenewable Resource Problem

A social planner wishes to maximize the discounted sum of net social surplus from harvesting a nonrenewable resource over an infinite horizon. For year t, let s_t denote the resource stock at the beginning of the year, let x_t denote the amount of the resource harvested, let $c_t = c(x_t)$ denote the total cost of harvesting, and let $p_t = p(x_t)$ denote the market clearing price. What is the socially optimal harvest policy?

This is an infinite horizon, deterministic model with time $t \in \{0, 1, 2, \ldots\}$ measured in years. There is one state variable,

$$\begin{aligned} s_t &= \text{stock of resource at beginning of year } t \\ s_t &\in [0, \infty), \end{aligned}$$

and one action variable,

$$x_t = \text{amount of resource harvested in year } t,$$

subject to the constraint

$$0 \leq x_t \leq s_t.$$

The reward earned by the optimizing agent is

$$\int_0^{x_t} p(\xi)\, d\xi - c(x_t).$$

State transitions are governed by

$$s_{t+1} = s_t - x_t.$$

The value function

$$V(s) = \text{net social value of resource stock}$$

satisfies Bellman's equation

$$V(s) = \max_{0 \le x \le s} \{ \int_0^x p(\xi)\ d\xi - c(x) + \delta V(s - x) \}, \qquad s \ge 0.$$

Assuming $p(0) = \infty$ and $g(0) = 0$, the solution to the optimization problem embedded in Bellman's equation will be internal. Under these assumptions, the shadow price function

$$\lambda(s) = \text{shadow price of resource}$$

satisfies the Euler conditions, which stipulate that for every stock level $s > 0$ there is a harvest level x such that

$$p(x) = c'(x) + \delta\lambda(s - x)$$

$$\lambda(s) = \delta\lambda(s - x).$$

Thus, along the optimal path

$$p_t = c'_t + \lambda_t$$

$$\lambda_t = \delta\lambda_{t+1}$$

where p_t is the market price and c'_t is the marginal harvest cost at t. That is, the market price of the harvested resource equals the marginal value of the unharvested resource plus the marginal cost of harvesting it. Also, the price of the harvested resource grows at the rate of interest. The steady-state, which occurs when stock is $s^* = 0$, is an uninteresting case.

Figure. Optimal harvest path of a renewable resource.

8.4.3 Renewable Resource Problem

A social planner wishes to maximize the discounted sum of net social surplus from harvesting a renewable resource over an infinite horizon. For year t, let s_t denote the resource stock at the beginning of the year, let x_t denote the amount of the resource harvested, let $c_t = c(x_t)$ denote the total cost of harvesting, and let $p_t = p(x_t)$ denote the market clearing price. Growth in the stock level is given by $s_{t+1} = g(s_t - x_t)$. What is the socially optimal harvest policy?

This is an infinite horizon, deterministic model with time $t \in \{0, 1, 2, \ldots\}$ measured in years. There is one state variable,

$$\begin{aligned} s_t &= \text{stock of resource at beginning of year } t \\ s_t &\in [0, \infty), \end{aligned}$$

and one action variable,

$$x_t = \text{amount of resource harvested in year } t,$$

subject to the constraint

$$0 \leq x_t \leq s_t.$$

The reward earned by the optimizing agent is

$$\int_0^{x_t} p(\xi)\ d\xi - c(x_t).$$

State transitions are governed by

$$s_{t+1} = g(s_t - x_t).$$

The value function

$$V(s) = \text{net social value of resource stock}$$

satisfies Bellman's equation

$$V(s) = \max_{0 \leq x \leq s} \{ \int_0^x p(\xi)\ d\xi - c(x) + \delta V(g(s - x)) \}.$$

Assuming $p(0) = \infty$ and $g(0) = 0$, the solution to the optimization problem embedded in Bellman's equation will be internal. Under these assumptions the shadow price function

$$\lambda(s) = \text{shadow price of resource}$$

satisfies the Euler conditions, which stipulate that for every stock level $s > 0$ there is a harvest level x such that

$$p(x) = c'(x) + \delta\lambda(g(s-x))g'(s-x)$$

$$\lambda(s) = \delta\lambda(g(s-x))g'(s-x).$$

Thus, along the optimal path

$$p_t = c'_t + \lambda_t$$

$$\lambda_t = \delta\lambda_{t+1}g'_t$$

where p_t is the market price, c'_t is the marginal harvest cost, and g'_t is the marginal future yield of stock in t. Thus, the market price of the harvested resource must cover both the marginal value of the unharvested resource and the marginal cost of harvesting it. Moreover, the value of one unit of resource today equals the discounted value of its yield tomorrow.

The steady-state resource stock s^*, harvest x^*, and shadow price λ^* solve the equation system

$$p(x^*) = c'(x^*) + \delta\lambda^* g'(s^* - x^*)$$

$$\lambda^* = \delta\lambda^* g'(s^* - x^*)$$

$$s^* = g(s^* - x^*).$$

These conditions imply $g'(s^* - x^*) - 1 = r$. That is, in steady-state, the marginal yield equals the interest rate.

Totally differentiating the equation system above:

$$\frac{\partial s^*}{\partial r} = \frac{1+r}{g''} < 0$$

$$\frac{\partial x^*}{\partial r} = -\frac{r}{g''} < 0$$

$$\frac{\partial \lambda^*}{\partial r} = \frac{(c'' - p')r}{g''} < 0.$$

That is, as the interest rate rises, the steady-state stock, the steady-state harvest, and the steady-state shadow price all fall.

Figure. Steady-state optimal harvest of a renewable resource.

8.4.4 Feedstock Problem

An animal weighing s_1 pounds in period $t = 1$ is to be fed up to period $T+1$, at which time it will be sold at a price of p dollars per pound. The cost of increasing the animal's weight by an amount x_t during period t is given by $c(s_t, x_t)$ where s_t is the animal's weight at the beginning of t. What feeding strategy maximizes the present value of profit?

This is a finite horizon, deterministic model with time $t \in \{1, 2, \ldots, T\}$ measured in feeding periods. There is one state variable,

$$\begin{aligned} s_t &= \text{weight of animal at beginning of period } t \\ s_t &\in [0, \infty), \end{aligned}$$

and one action variable,

$$x_t = \text{weight gain during period } t,$$

subject only to a nonnegativity constraint.

The reward earned by the hog farmer in feeding periods is

$$-c(s_t, x_t).$$

State transitions are governed by

$$s_{t+1} = s_t + x_t.$$

The value function

$$V_t(s) = \text{value of animal weighing pounds } s \text{ in period } t$$

satisfies Bellman's equation

$$V_t(s) = \max_{x \geq 0} \{-c(s, x) + \delta V_{t+1}(s + x)\},$$

subject to the terminal condition

$$V_{T+1}(s) \equiv ps.$$

The shadow price function

$$\lambda_t(s) = \text{shadow price of animal mass in period } t$$

satisfies the Euler conditions, which stipulate that for each decision period t and weight level $s > 0$, the optimal weight gain x satisfies

$$\delta\lambda_{t+1}(s+x) - c_x(s,x) \perp x \geq 0$$

$$\lambda_t(s) = -c_s(s,x) + \delta\lambda_{t+1}(s+x).$$

For the post-terminal period,

$$\lambda_{T+1}(s) = p.$$

Thus, along an optimal path, assuming an internal solution, we have:

$$\delta\lambda_{t+1} = c_x(s_t, x_t)$$

$$c_s(s_t, x_t) = \lambda_t - \delta\lambda_{t+1}.$$

In other words, the marginal cost of feeding the animal this period must equal the discounted value of the additional body mass obtained the following period. Also, the marginal value of body mass declines at the same rate at which it weight gains become increasingly more costly.

Figure. Feedstock problem dynamics.

8.4.5 Optimal Growth with Debt

Reconsider the optimal growth problem when the central planner can carry an external debt load d_t whose unit cost $\eta_0 + \eta_1 q_t$ rises with the debt to asset ratio $q_t = d_t/s_t$.

This is an infinite horizon, stochastic model with time $t \in \{0, 1, 2, \ldots\}$ measured in years. There are two state variables:

$$\begin{aligned} s_t &= \text{stock of good at beginning of year } t \\ s_t &\in [0, \infty) \end{aligned}$$

and

$$\begin{aligned} d_t &= \text{debt load at beginning of year } t \\ d_t &\in (-\infty, \infty). \end{aligned}$$

Here, $d_t < 0$ implies that the economy runs a surplus. There are two action variables:

$$x_t = \text{amount of good invested in year } t$$

$$c_t = \text{amount of good consumed in year } t,$$

both subject to nonnegativity constraints.

The reward earned by the optimizing agent is

$$u(c) = \text{social utility in } t.$$

Supply state transitions are governed by

$$s_{t+1} = \gamma x_t + \epsilon_{t+1} f(x_t)$$

where

$$\epsilon_t = \text{productivity shock in year } t.$$

Debt state transitions are governed by

$$d_{t+1} = d_t + b_t,$$

where

$$b_t = c_t + x_t + (\eta_0 + \eta_1 d_t / s_t) \cdot d_t - s_t,$$

indicates net borrowing in year t.

The value function

$$V(s, d) = \text{sum of current and expected future social welfare}$$

satisfies Bellman's equation

$$V(s, d) = \max_{x \geq 0, c \geq 0} \{u(c) + \delta E V(\gamma x + \epsilon f(x), d + b)\}$$

where $b = x + c + (\eta_0 + \eta_1 d/s) \cdot d - s)$ is net borrowing.

Assuming $u'(0) = -\infty$ and $f(0) = 0$, the solution to Bellman's equation will always be internal, and the shadow price and cost functions

$$\lambda(s, d) = \frac{\partial V}{\partial s}(s, d) = \text{shadow price of stock}$$

and

$$\mu(s,d) = \frac{\partial V}{\partial d}(s,d) = \text{shadow cost of debt}$$

satisfy the Euler equilibrium conditions, which stipulate that for every stock level $s > 0$ and debt level d,

$$u'(c) + \delta E\mu(\gamma x + \epsilon f(x), d + b) = 0$$

$$\delta E\left[\lambda(\gamma x + \epsilon f(x), d + b) \cdot (\gamma + \epsilon f'(x))\right] + \delta E\mu(\gamma x + \epsilon f(x), d + b) = 0$$

$$\lambda(s,d) = -\delta E[\mu(\gamma x + \epsilon f(x), d + b) \cdot (1 + \eta_1 q^2)]$$

$$\mu(s,d) = \delta E[\mu(\gamma x + \epsilon f(x), d + b) \cdot (1 + \eta_0 + 2\eta_1 q)]$$

where $q = d/s$ is the debt to asset ratio.

The certainty-equivalent steady-state is obtained by assuming $\epsilon = 1$ with probability 1. The certainty-equivalent steady-state stock of good s^*, debt load d^*, debt-asset ratio $q^* = d^*/s^*$, investment level x^*, consumption level c^*, stock shadow price λ^*, and debt shadow cost μ^* solve the equation system

$$u'(c) + \delta\mu^* = 0$$

$$\delta\lambda^*(\gamma + f'(x^*)) + \delta\mu^* = 0$$

$$\lambda^* = -\delta\mu^*(1 + \eta_1 q^{*2})$$

$$\mu^* = \delta\mu^*(1 + \eta_0 + 2\eta_1 q^*)$$

$$s^* = \gamma x^* + f(x^*)$$

$$s^* = x^* + c^* + (\eta_0 + \eta_1 q^*)d^*$$

$$q^* = d^*/s^*.$$

These conditions imply a steady-state optimal debt load $q^* = (r - \eta_0)/(2\eta_1)$, which increases with the discount rate r but falls with the base cost of debt η_0.

8.4.6 A Production-Inventory Problem

The output price faced by a competitive firm follows a first-order autoregressive process:

$$p_{t+1} = \alpha + \gamma p_t + \epsilon_{t+1}, \qquad |\gamma| < 1, \epsilon_t \text{ i.i.d.}$$

The cost of producing q_t units in period t is $c(q_t)$. The firm may store across periods at a constant unit cost k. Assuming p_t is known at the time the period t production-inventory decision is made, what production-inventory policy maximizes the sum of current and expected future profits?

This is an infinite horizon, stochastic model with time $t \in \{0, 1, 2, \ldots\}$ measured in years. There are two state variables:

$$b_t = \text{beginning inventories,}$$

$$p_t = \text{current market price.}$$

There are two action variables:

$$q_t = \text{current production}$$

$$x_t = \text{ending inventories}$$

subject to the constraints

$$q_t \geq 0$$

$$x_t \geq 0$$

$$x_t \leq q_t + b_t;$$

that is, production, inventories, and deliveries must be nonnegative.

The reward earned by the optimizing agent is

$$p_t \cdot (q_t + b_t - x_t) - c(q_t) - kx_t.$$

State transitions are governed by

$$p_{t+1} = \alpha + \gamma p_t + \epsilon_{t+1}$$

where

$$\epsilon_t = \text{price process innovation in year } t$$

and

$$b_{t+1} = x_t.$$

The value function

$$V(b,p) = \text{value of firm given inventories } b \text{ and price } p$$

satisfies Bellman's equation

$$V(b,p) = \max_{0\leq q, 0\leq x\leq q+b} \{p(q+b-x) - c(q) - kx + \delta EV(x, \alpha + \gamma p + \epsilon)\}.$$

The shadow price function

$$\lambda(b,p) = V_b(b,p) = \text{marginal value of inventories}$$

satisfies the Euler conditions, which require that for every beginning inventory level b and price p, there is a production level q, ending inventory level x, and material balance shadow price μ such that

$$\delta E\lambda(b, \alpha + \gamma p + \epsilon) - p - k - \mu \perp x \geq 0$$

$$p - c'(q) \perp q \geq 0$$

$$q + b - x \perp \mu \geq 0$$

$$\lambda(b,p) = p - \mu$$

Along the optimal path, if deliveries and storage are positive,

$$\delta E_t p_{t+1} - p_t - k = 0$$

$$p_t = c'_t.$$

That is, marginal revenue equals the marginal production cost and the discounted expected future price equals the current output price plus the cost of storage.

The certainty-equivalent deterministic problem is obtained by assuming p is fixed at its longrun mean $\alpha/(1-\gamma)$. The certainty-equivalent steady-state inventories are 0 and production is constant and implicitly defined by the short-run profit maximization condition:

$$p = c'(q).$$

8.5 Rational Expectations Models

By definition, agents in rational expectations models take into account how their actions will affect them in the future and form expectations that coincide with those implied by the model as a whole. Most discrete time rational expectation models take the following form: At the beginning of period t, an economic system emerges in a state s_t. The agents in the economic system observe the state of the system and, by pursuing their individual objectives, formulate a collective behavioral response x_t. The economic system then evolves to a new state s_{t+1} that depends on the current state s_t and response x_t, and an exogenous random shock ϵ_{t+1} that is realized only after the agents respond at time t.

More formally, the behavioral responses of economic agents and the state transitions of the economic system are governed by a structural law of the form

$$f(s_t, x_t, E_t x_{t+1}) = 0,$$

and the dynamic law

$$s_{t+1} = g(s_t, x_t, \epsilon_{t+1}).$$

The stipulation that only the expectation of the subsequent period's behavioral response is relevant to the current response of agents is more general than first appears. By introducing new accounting variables, the current response can be made to depend on the expectation of any function of future states and responses, including states and responses more than one period into the future.

The state space $S \in \Re^n$, which contains all the states attainable by the economic system, and the response space $X \in \Re^m$, which contains all behavioral responses that may be made by the economic agents, are both assumed to be closed convex nonempty sets. In some instances, the range of admissible responses may vary with the state of the process s_t. In such cases, the restricted response space will be denoted $X(s_t)$ and will be assumed to be a closed convex nonempty set. The structure f and dynamic law g are assumed to be twice continuously differentiable on S and X and the per-period discount factor δ is assumed to be less than one. The exogenous random shocks ϵ_t are assumed identically distributed over time, mutually independent, and independent of past states and responses.

The primary task facing an economic analyst is to explain the behavioral response $x = x(s)$ of agents in each state s attainable by the process. The response function $x(\cdot)$ is characterized implicitly as the solution to a functional equation:

$$f(s, x(s), Ex(g(s, x(s), \epsilon)) = 0 \qquad \forall s \in S.$$

In many instances, this functional equation will not possess a closed-form solution and can only be solved numerically.

8.5.1 Lucas-Prescott Asset Pricing Model

The basic rational expectations asset pricing model has been studied extensively by macroeconomists. The model assumes the existence of a pure exchange economy in which a representative infinitely-lived agent allocates real wealth between immediate consumption q_t and investment in an index asset i_t. The agent's objective is to maximize expected lifetime utility subject to an intertemporal budget constraint:

$$\begin{aligned} \max \quad & E_t\{\sum_{k=0}^{\infty} \delta^k u(q_{t+k})\} \\ \text{s.t.} \quad & q_t + i_t = i_{t-1} r_t. \end{aligned} \tag{8.1}$$

Here, E_t is the conditional expectation operator given information available at time t, δ is the agent's subjective discount rate, i_t is the amount of asset held by the agent at the end of period t, and r_t is the asset's return in period t.

Under mild regularity conditions, the agent's dynamic optimization problem has an unique solution that satisfies the first-order Euler condition:

$$\delta E_t[u'(q_{t+1}) r_{t+1}] = u'(q_t).$$

The Euler condition asserts that along an optimal consumption path the marginal utility of consuming one unit of wealth today equals the marginal benefit of investing the unit of wealth and consuming it and its dividends tomorrow.

The asset pricing model may be completed by specifying the utility function, introducing a production sector, and imposing a market clearing condition. Assume that the agent's preferences exhibit constant relative risk-aversion $\gamma > 0$:

$$u(q) = \frac{q^{1-\gamma}}{1-\gamma}.$$

We assume that aggregate output y_t is exogenous and follows a stationary first-order autoregressive process whose innovation ϵ_t is normally distributed white noise with standard deviation σ_ϵ:

$$y_t = \alpha + \beta y_{t-1} + \epsilon_t.$$

And we assume that output is consumed entirely in the period that it is produced:

$$y_t = q_t.$$

A formal solution to the rational expectations asset pricing model is a rule that gives the equilibrium asset return r_t as a function of current and past realizations of the driving exogenous output process. Lucas demonstrated that when the output process is stationary and first-order Markovian, as assumed here, the rule is well-defined. In particular, the equilibrium return in period t will be a stationary deterministic function of the contemporaneous output level y_t:

$$r_t = \lambda(y_t).$$

From the dynamic equilibrium conditions, it follows that the asset return function λ is characterized by the equilibrium condition:

$$E_\epsilon \delta(\alpha + \beta y + \epsilon)^{-\gamma} \lambda(\alpha + \beta y + \epsilon) = y^{-\gamma} \qquad \forall y.$$

The Euler functional equation of the asset pricing model is nonlinear and lacks a known a closed-form solution. It can only be solved approximately using numerical functional equation methods.

8.5.2 Competitive Storage Under Uncertainty

The centerpiece of the classical theory of storage is the competitive intertemporal arbitrage equation

$$\delta E_t p_{t+1} - p_t = c(x_t).$$

The intertemporal arbitrage equation asserts that, in equilibrium, expected appreciation in the commodity price p_t must equal the unit cost of storage $c(x_t)$. Dynamic equilibrium in the commodity market is enforced by competitive expected-profit-maximizing storers. Whenever expected appreciation exceeds the storage cost, the attendant profits induce storers to increase their stockholdings until the equilibrium is restored. Conversely, whenever the storage cost exceeds expected appreciation, the attendant loses induce storers to decrease their stockholdings until the equilibrium is restored.

According to the classical theory, the unit storage cost $c(x_t)$ is a nondecreasing function of the amount stored x_t. The unit storage cost represents the marginal physical cost of storage less the marginal "convenience yield", which is the amount processors are willing to pay to have sufficient stocks available to avoid costly production adjustments. If stock levels are high, the marginal convenience yield is zero and the unit storage cost equals the physical storage cost. As stock levels approach zero, however, the marginal convenience yield rises, eventually resulting in a negative unit storage cost. The classical storage model has received strong empirical support over the years and captures the key stylized fact of markets for storable commodities: the coincidence of negative intertemporal price spreads and low, but positive, stock levels.

The modern theory of storage extends the classical model to a partial equilibrium model of price-quantity determination by appending supply, demand, and market clearing conditions to the intertemporal arbitrage equation. For the sake of discussion, let us consider a simple agricultural commodity market model with exogenous production. Denote quantity consumed by q_t, quantity harvested h_t, available supply by s_t, and the per-period discount factor by δ. Assume that the market clearing price is a decreasing function of the quantity consumed:

$$p_t = p(q_t);$$

that available supply is either consumed in the current period or stored:

$$s_t = q_t + x_t;$$

and that the supply available next period will be the sum of current carryout and next period's harvest:

$$s_{t+1} = x_t + h_{t+1}.$$

The modern storage model is closed by assuming that price expectations are consistent with the other structural assumptions of the model. The so-called rational expectations assumption endogenizes the expected future price while preserving internal consistency of the model.

The solution of the nonlinear rational expectations commodity market model is illustrated in figures (*)-(*). These figures show, respectively, equilibrium price and carryout in terms of available supply. For comparison, the first figure also shows the inverse consumption demand function $p(\cdot)$, which gives the market price that would prevail in the absence of storage. At low supply levels, there is effectively no storage and the equilibrium price coincides with the inverse consumption demand function. Over this range, acreage supply is not significantly affected by variations in available supply. At sufficiently high supply levels, incentives for speculative storage begin to appear. Over this range, the equilibrium price, which reflects both consumption and storage demand, exceeds the inverse consumption demand function.

The nonlinear rational expectations commodity market model cannot be solved using standard algebraic techniques. To see this, let $\lambda(s)$ denote the equilibrium price implied by the model for a given available supply s. Having the equilibrium price function $\lambda(\cdot)$, the rational ex-ante expected price could be computed by integrating over the harvest distribution:

$$E_t p_{t+1} = E_y \lambda(x_t + h_{t+1})$$

Appending this equation to the previous three market equations would result in a system of four nonlinear algebraic equations that in principle could be solved for all the unknowns.

Unfortunately, the equilibrium price function $\lambda(\cdot)$ is not known a priori and deriving it, the key to solving the commodity market model, is a nontrivial functional equation problem. Combining all the behavioral relations, we see that $\lambda(\cdot)$ must simultaneously satisfy an infinite number of conditions. Specifically, for every realizable supply s,

$$\lambda(s) = p(s - x)$$

where stock x solves

$$\delta E_y \lambda(x + h) - p(s - x) = c(x)$$

In the general framework developed for rational expectations models above, available supply is only state variable, price and carryout are the response variables, and harvest is the random driving shock. Only the relationship between price and supply needs to be derived, since only future price expectations affect behavior, and once the price and supply are known, the carryout may be computed from the inverse demand function.

An alternative way to pose the rational expectations commodity storage model is to integrate it into an equivalent optimization problem. Consider the problem of maximizing the discounted expected sum of consumer surplus less storage costs. The resulting dynamic optimization problem, with state variable s and action variable x, yields the following Bellman equation:

$$V(s) = \max_{0 \leq x \leq s} \{ \int_0^{s-x} p(\xi)\ d\xi - \int_0^x c(\xi)\ d\xi + \delta EV(x+h) \}, \qquad s \geq 0.$$

One may verify that the Euler equilibrium conditions for this dynamic optimization problem are precisely the equilibrium conditions of the original rational expectations model, provided that the shadow price of the optimization problem is identified with the rational expectations equilibrium market price.

Finally, one might compute certainty-equivalent steady-state supply s^*, storage x^*, and price p^* by solving the equation system

$$\delta p^* = f(s^* - x^*) + c(x^*)$$

$$p^* = f(s^* - x^*)$$

$$s^* = x^* + h^*$$

where h^* is the expected harvest.

8.5.3 Spatial-Temporal Equilibrium

8.6 Dynamic Games

Dynamic game models attempt to capture strategic interactions among a small number of dynamically optimizing agents when the actions of one agent affects the welfare of another. For the sake of brevity, we consider only two agent games. The theory and methods developed below, however, can be generalized to accommodate an arbitrary number of agents.

Denote by s_i the state of the process controlled by agent i, and denote by x_i the action taken by agent i. In a dynamic game setting, agent i receives a reward that depends not only on the state of his own process and the action he takes, but also the state s_j of the other agent's process and the action x_j that he takes. Specifically, the reward earned by agent i at any point in time is $f_i(s_i, s_j, x_i, x_j)$.

As with a static game, the equilibrium solution to a dynamic game depends on the information available to the agents and the class of strategies they are allowed to pursue. For simplicity, we consider only the most common game structure. Specifically, we will seek a noncooperative Nash game equilibrium under the assumption that each agent knows the other agent's state at any point in time, and that each agent also knows the policy followed by the other agent. A dynamic Nash game equilibrium exists when each agent's policy maximizes his own stream of current and expected future rewards given that the other agent follows his policy.

The dynamic Nash game equilibrium may be formally expressed by a pair of Bellman equations, one for each agent. The Bellman equation for agent i takes the form

$$V_i(s_i, s_j) = \max_{x \in X(s_i,s_j)} \{f_i(s_i, s_j, x_i, x_j) + \delta E_\epsilon V_i(g_1(s_1, x_1, \epsilon_1), g_2(s_2, x_2, \epsilon_2))\},$$

for $s_i, s_j \in S$. Here, $V_i(s_i, s_j)$ denotes the maximum current and expected future rewards that can be earned by agent i, given that agent j remains committed to his policy. Solving for the Nash equilibrium involves finding policies x_i and x_j for every state that solve the Bellman equations of both agents simultaneously.

Let λ_{ii} denote the partial derivative of agent i's value function with respect to the state controlled by him:

$$\lambda_{ii}(s_1, s_2) = \frac{\partial V_i}{\partial s_i}(s_1, s_2) \qquad \forall s_1, s_2.$$

Also, let λ_{ij} denote the partial derivative of agent i's value function with respect to the state controlled by agent j:

$$\lambda_{ij}(s_1, s_2) = \frac{\partial V_i}{\partial s_j}(s_1, s_2) \qquad \forall s_1, s_2.$$

The shadow price function λ_{ii} represents agent i's valuation of a marginal unit of the state controlled by him; the shadow price function λ_{ij} represents agent i's valuation of a marginal unit of the state controlled by his rival.

The first-order equilibrium conditions for the Nash dynamic game are derived by applying the Karush-Kuhn-Tucker and Envelope Theorems to the optimization problems embedded in the two Bellman equations. Assuming actions are unconstrained, the Karush-Kuhn-Tucker conditions for the embedded unconstrained optimization problems imply that the optimal action x_i for agent i, given state s_i, s_j, must satisfy the equimarginality condition:

$$\frac{\partial f_i}{\partial x_i}(s_1, s_2, x_1, x_2) + \delta E_\epsilon \left[\lambda_{ii}(s_1', s_2') \frac{\partial g_i}{\partial x_i}(s_i, x_i, \epsilon_i) \right] = 0$$

where $s_i' = g_i(s_i, x_i, \epsilon_i)$. The Envelope Theorem applied to the same problem implies:

$$\lambda_{ii}(s_1, s_2) = \frac{\partial f_i}{\partial s_i}(s_1, s_2, x_1, x_2) + \frac{\partial f_i}{\partial x_j}(s_1, s_2, x_1, x_2) \frac{\partial x_j}{\partial s_i}(s_1, s_2) +$$

$$\delta E_\epsilon \left[\lambda_{ii}(s_1', s_2') \frac{\partial g_i}{\partial s_i}(s_i, x_i, \epsilon_i) + \lambda_{ij}(s_1', s_2') \frac{\partial g_j}{\partial x_j}(s_j, x_j, \epsilon_j) \frac{\partial x_j}{\partial s_i}(s_1, s_2) \right]$$

$$\lambda_{ij}(s_1, s_2) = \frac{\partial f_i}{\partial s_j}(s_1, s_2, x_1, x_2) + \frac{\partial f_i}{\partial x_j}(s_1, s_2, x_1, x_2) \frac{\partial x_j}{\partial s_j}(s_1, s_2) +$$

$$\delta E_\epsilon \left[\lambda_{ij}(s_1', s_2') \frac{\partial g_j}{\partial s_j}(s_j, x_j, \epsilon_j) + \lambda_{ij}(s_1', s_2') \frac{\partial g_j}{\partial x_j}(s_j, x_j, \epsilon_j) \frac{\partial x_j}{\partial s_j}(s_1, s_2) \right].$$

The Euler conditions for a two agent dynamic game thus comprise six functional equations in six unknown functions: the two own-shadow price functions, the two cross-shadow price functions, and the two optimal policy functions.

8.6.1 Redistribution Game

Consider an economy comprising two agents and a single composite good. Each year t begins with predetermined amounts of the good s_{1t} and s_{2t} held by the two agents, respectively. Given the amounts on hand, each agent selects an amount x_{it} to be invested, and consumes the rest. The utility derived from consumption in year t by agent i is $u_i(s_{it} - x_{it})$. Given each agent's investment decision, the amount of good available in year $t+1$ to agent i will be $s_{i,t+1} = g_i(x_{it}, \epsilon_{i,t+1}) = \gamma x_{it} + \epsilon_{i,t+1} f_i(x_{it})$ where γ is the capital survival rate, f_i is agent i's production function, and $\epsilon_{i,t+1}$ is a positive production shock with mean 1 that is specific to agent i.

Suppose now that the two agents agree to insure against a string of production disasters by entering into a contract to share collective wealth in perpetuity. Specifically, the agents agree that, in any given period t, the wealthier of the two agents will transfer a certain proportion σ of the wealth differential to the poorer agent. Under this scheme, if agent i begins period t with wealth s_{it}, his post-transfer wealth will be $\hat{s}_{it} = s_{it} - \sigma(s_{it} - s_{jt})$. If the wealth transfer is enforceable, but agents are free to pursue consumption and investments freely, moral hazard will arise. In particular, both agents will have incentives to change their consumption and investment policies upon introduction of insurance. How will insurance affect the agents' investment behavior, and for what initial wealth states s_{1t} and s_{2t} and share parameter σ will both agents be willing to enter into the insurance contract?

The essence of the dynamic Nash game equilibrium for the redistribution game is captured by a pair of Bellman equations, one for each agent. The Bellman equation for agent i takes the form

$$V_i(s_i, s_j) = \max_{0 \leq x_i \leq \hat{s}_i} \{u_i(\hat{s}_i - x_i) + \delta E_\epsilon V_i(g_i(x_i, \epsilon_i), g_j(x_j, \epsilon_j))\},$$

where $\hat{s}_i = s_i - \sigma(s_i + s_j)$, for $s_i, s_j \in S$. Here, $V_i(s_i, s_j)$ denotes the maximum current and expected future rewards that can be earned by agent i, given that agent j remains committed to his policy.

The first-order equilibrium conditions for the Nash dynamic game are derived by applying the Karush-Kuhn-Tucker and Envelope Theorems to the optimization problems embedded in the two Bellman equations. Assuming an internal solution to each agent's investment problem, the Karush-Kuhn-Tucker conditions imply that the optimal investment x_i for agent i, given wealths s_i, s_j, must satisfy the equimarginality condition:

$$-u_i'(\hat{s}_i - x_i) + \delta E_\epsilon \left[\lambda_{ii}(s_1', s_2') \frac{\partial g_i}{\partial x_i}(x_i, \epsilon_i) \right] = 0$$

where $s_i' = g_i(x_i, \epsilon_i)$. The Envelope Theorem applied to the same problem implies:

$$\lambda_{ii}(s_1, s_2) = (1 - \sigma) u_i'(\hat{s}_i - x_i) + \delta E_\epsilon \left[\lambda_{ij}(s_1', s_2') \frac{\partial g_j}{\partial x_j}(x_j, \epsilon_j) \frac{\partial x_j}{\partial s_i}(s_1, s_2) \right]$$

$$\lambda_{ij}(s_1, s_2) = \sigma u_i'(\hat{s}_i - x_i) + \delta E_\epsilon \left[\lambda_{ij}(s_1', s_2') \frac{\partial g_j}{\partial x_j}(x_j, \epsilon_j) \frac{\partial x_j}{\partial s_j}(s_1, s_2) \right],$$

The Euler conditions for a two agent dynamic game thus comprise six functional equations in six unknown functions: the two own-shadow price functions, the two cross-shadow price functions, and the two optimal policy functions.

8.6.2 Marketing Board Game

Assume that there are only two countries that can supply a given commodity on the world market. In each country, a government marketing board has the exclusive power to sell the commodity on the world market. The marketing boards compete with each other, using storage as a strategy variable to maximize the present value of current and expected future income from commodity sales.

For each exporting country $i = 1, 2$ and period t, let s_{it} denote the supply available at the beginning of period, let q_{it} denote the quantity exported, let x_{it} denote the stocks held at the end of the period, let y_{it} denote new production, let p_t denote the world price, let c_{it} denote total storage costs, and let δ denote the discount factor.

Formally, each marketing board $i = 1, 2$, solves

$$\max E \sum_{t=0}^{\infty} \delta [p_t q_{it} - c_{it}]$$

subject to the following conditions: Available supply is the sum of beginning stocks and new production:

$$s_{it} = x_{it-1} + y_{it}.$$

Available supply is either exported or stored:

$$s_{it} = q_{it} + x_{it}.$$

The world market clearing price p_t is a decreasing function $\pi(\cdot)$ of the total amount exported:

$$p_t = \pi(q_{1t} + q_{2t}).$$

The cost of storage is an increasing function $c_i(\cdot)$ of the quantity stored:

$$c_{it} = c_i(x_{it}).$$

And production y_{it} is exogenous, stochastic, independently distributed across countries, and independently and identically distributed across time.

Each marketing board faces a dynamic optimization problem subject to the constraints. The price, and thus the payoff, for each country at time t is simultaneously determined by the quantities marketed by both boards.

In making its storage decision, each board must anticipate the storage decision of its rival. The two optimization problems must therefore be solved simultaneously to determine the equilibrium levels of stocks, exports, and price.

The noncooperative Nash equilibrium is characterized by a pair of Bellman equations, which for country i takes the form

$$V_i(s_1, s_2) = \max_{x_i}[pq_i - c_i + \delta E_y V_i(x_1 + y_1, x_2 + y_2)] \qquad \forall s_1, s_2$$

where $q_i = s_i - x_i$, $p = \pi(q_1 + q_2)$, and $c_i = c_i(x_i)$.

For each combination of $i = 1, 2$ and $j = 1, 2$, let λ_{ij}^* denote the partial derivative of country i's value function with respect to the supply in country j:

$$\lambda_{ij}(s_1, s_2) = \frac{\partial V_i}{\partial s_j}(x_1 + y_1, x_2 + y_2) \qquad \forall s_1, s_2.$$

The shadow price function λ_{ij} represents country i's valuation of a marginal unit of stock in country j.

Applying the to Envelope Theorem to Bellman equation, the own shadow price function must satisfy

$$\lambda_{ii}(s_1, s_2) = p + p' q_i[1 - \frac{\partial x_j}{\partial s_i}] + \delta E_y \lambda_{ij}(x_1 + y_1, x_2 + y_2)\frac{\partial x_j}{\partial s_i} \qquad \forall s_1, s_2$$

and the cross shadow price function must satisfy

$$\lambda_{ij}(s_1, s_2) = p' q_i[1 - \frac{\partial x_j}{\partial s_j}] + \delta E_y \lambda_{ij}(x1 + y_1, x2 + y_2)\frac{\partial x_j}{\partial s_j} \qquad \forall s_1, s_2$$

where $p' = \pi'(q_1 + q_2)$.

Another necessary condition for the dynamic feedback Nash equilibrium can be obtained by deriving the first-order condition for the optimization problem embedded in Bellman's equation:

$$p + p' q_i = \delta E_y \lambda_{ii}(x_1 + y_1, x_2 + y_2) - c_i' \qquad \forall s_1, s_2$$

where $c_i' = c_i'(x_i)$. This condition asserts that along an equilibrium path, the marginal payoff from selling this period $p + p' q_i$ must equal the expected marginal payoff from storing and selling next period $\delta E_y \lambda_{ii} - c_i'$.

The noncooperative feedback Nash equilibrium for the game between the two marketing boards is characterized by six functional equations in six unknowns: the equilibrium feedback strategies x_1 and x_2 and the equilibrium shadow price functions λ_{11}, λ_{12}, λ_{21}, and λ_{22}.