stochastically, and (54) is a natural representation.[13] Although $x(t)$ will fluctuate stochastically, it is of interest to study the probability density for x in steady-state equilibrium.

In steady-state equilibrium, ϕ does not depend on x_0, t_0, or t in the forward equation (46). Write the stationary density function as $\phi_\infty(x)$. Then the equation becomes (after integrating once):

$$\tfrac{1}{2}\frac{d}{dx}[\sigma^2(x)\,\phi_\infty(x)] = [f(x) - q(x)]\,\phi_\infty(x).$$

This can be rewritten as

$$\frac{d[\sigma^2(x)\,\phi_\infty(x)]}{\sigma^2(x)\,\phi_\infty(x)} = \frac{2}{\sigma^2(x)}\,[f(x) - q(x)]\,dx. \tag{55}$$

This can be integrated to give the following equation for the steady-state density:[14]

$$\phi_\infty(x) = \frac{m}{\sigma^2(x)}\,\exp\left[2\int^x \frac{f(v) - q(v)}{\sigma^2(v)}\,dv\right], \tag{56}$$

where m is a constant of integration, chosen so that $\int_0^\infty \phi_\infty(x)dx = 1$.

As an example, suppose that $f(x)$ is the logistic function, that is,

$$f(x) = \alpha\,x\,(1 - x/K),$$

where K is the "carrying capacity" of the resource stock. Also, suppose that there is no harvesting, that is, $q(x) = 0$, and that $\sigma(x) = \sigma x$. Then equation (56) yields the following steady-state density that applies when $\sigma^2 < 2\alpha$:

$$\phi_\infty(x) = (2\alpha/\sigma^2 K)^{2\alpha/\sigma^2 - 1}\,x^{2\alpha/\sigma^2 - 2}\,e^{-2\alpha x/\sigma^2 K}\,/\,\Gamma(2\alpha/\sigma^2 - 1), \tag{57}$$

where Γ denotes the gamma function. From this we can determine that the expected value of x in steady-state equilibrium is

$$\mathcal{E}(x_\infty) = K\left(1 - \frac{\sigma^2}{2\alpha}\right). \tag{58}$$

Note that stochastic fluctuations reduce the steady-state expected value of x, and as σ^2 approaches 2α, $\mathcal{E}(x_\infty)$ approaches zero. Also, if $\sigma^2 \geq 2\alpha$, stochastic fluctuations will drive the resource stock to extinction, that is, ϕ_∞ collapses and $x(t) \to 0$ with probability 1. [For a more detailed discussion of this and related models of renewable resources, as well as a derivation of the optimal stochastic harvesting rule $q^*(x)$, see Pindyck (1984).]

[13] See, for example, Beddington and May (1977) and Goel and Richter-Dyn (1974). For a good overview of renewable resource economics in a deterministic context, see Clark (1976).

[14] Merton (1975) also provides a derivation of this equation, and shows how it can be applied to a neoclassical model of growth with a stochastically evolving population.

Chapter 4

Dynamic Optimization under Uncertainty

TIME PLAYS a particularly important role for investment decisions. The payoffs to a firm's investment made today accrue as a stream over the future, and are affected by uncertainty as well as by other decisions that the firm or its rivals will make later. The firm must look ahead to all these developments when making its current decision. As we emphasized in Chapter 2, one aspect of this future is an opportunity to make the same decision later; therefore the option of postponement should be included in today's menu of choices. The mathematical techniques we employ to model investment decisions must be capable of handling all these considerations.

In this chapter we develop two such techniques: dynamic programming and contingent claims analysis. They are in fact closely related to each other, and lead to identical results in many applications. However, they make different assumptions about financial markets, and the discount rates that firms use to value future cash flows.

Dynamic programming is a very general tool for dynamic optimization, and is particularly useful in treating uncertainty. It breaks a whole sequence of decisions into just two components: the immediate decision, and a valuation function that encapsulates the consequences of all subsequent decisions, starting with the position that results from the immediate decision. If the planning horizon is finite, the very last decision at its end has nothing following it, and can therefore be found using standard static optimization methods. This solution then provides the valuation function appropriate to the penultimate decision. That, in turn, serves for the decision two stages from the end, and so on. One can work backwards all the way to the initial condition. This sequence of computations might seem difficult, but advances in computing hardware and software have made it quite feasible, and in this book we will obtain solutions to several problems of this kind. If the planning horizon is infinite, what might seem like an even more difficult calculation is simplified by its recursive nature: each decision leads to another problem that looks exactly like the original one. This not only facilitates numerical computation, but also often makes it possible to obtain a theoretical characterization of the solution, and sometimes an analytical solution itself.

Contingent claims analysis builds on ideas from financial economics. Begin by observing that an investment project is defined by a stream of costs and benefits that vary through time and depend on the unfolding of uncertain events. The firm or individual that owns the right to an investment opportunity, or to the stream of operating profits from a completed project, owns an *asset* that has a value. A modern economy has markets for quite a rich menu of assets of all kinds. If our investment project or opportunity happens to be one of these traded assets, it will have a known market price. However, even if it is not directly traded, one can compute an implicit value for it by relating it to other assets that are traded.

All one needs is some combination or portfolio of traded assets that will exactly replicate the pattern of returns from our investment project, at every future date and in every future uncertain eventuality. (The composition of this portfolio need not be fixed; it could change as the prices of the component assets change.) Then the value of the investment project must equal the total value of that portfolio, because any discrepancy would present an arbitrage opportunity: a sure profit by buying the cheaper of the two assets or combinations, and selling the more valuable one. Implicit in this calculation is the requirement that the firm should use its investment opportunity in the most efficient way, again because if it did not, an arbitrager could buy the investment opportunity and make a positive profit. Once we know the value of the investment opportunity, we can find the best form, size, and timing of investment that achieves this value, and thus determine the optimal investment policy.

In the first section of this chapter we develop the basic ideas of dynamic programming. For continuity we start with the same two-period example that was the mainstay of Chapter 2, and then extend it to longer time spans, in discrete periods and in continuous time. There we consider the two special forms of uncertainty that were introduced in Chapter 3, namely Ito and Poisson processes. Section 2 develops the general ideas of contingent claims to

uncertain future payoffs, and the determination of their prices by arbitrage. Again the focus is on Ito and Poisson processes. Section 3 explains the relationships between the two approaches. We also develop this relationship to obtain a simple method for discounting and valuing payoffs with varying degrees of risk.

Throughout, we emphasize the intuition behind the methods rather than formal proofs. A sketch of the technical details is placed in an appendix.

1 Dynamic Programming

In this section we introduce the basic ideas of dynamic programming. We start with the two-period example from Chapter 2, thus providing a simple concrete setting for the ideas and continuity with the previous analysis. Then we extend the ideas and develop the general theory of multiperiod decision strategies. Finally, we let time be continuous, and represent the underlying uncertainty with either Ito or Poisson processes. That is the setting for most of the applications that will appear in later chapters.

1.A The Two-Period Example

An extended example in Chapter 2 introduced the ideas of optimal timing of investment decisions and of option values. We began by comparing the present values that result from immediate investment and from waiting. Dynamic programming is in essence a systematic method of making such comparisons for more general dynamic decisions. In order to illustrate this, let us proceed to generalize our example.

As the first step, in this section we rework the simple two-period example from Chapter 2 in a more general way. Before, we chose specific numerical values for most of the parameters in the example; now we will let them have arbitrary values. Let I denote the sunk cost of investment in the factory that then produces one widget per period forever, and let r be the interest rate. Suppose the price of a widget in the current period 0 is P_0. From period 1 onward, it will be $(1+u)\,P_0$ with probability q, and $(1-d)\,P_0$ with probability $(1-q)$.

First, suppose that the investment opportunity is available only in period 0; if the firm decides not to invest in period 0, it cannot change its mind in period 1. Let V_0 denote the expected present value of the revenues the firm gets if it invests. Weighting the two alternative possibilities for widget prices by their respective probabilities, discounting, and adding, we have

$$V_0 = P_0 + [q\,(1+u)\,P_0 + (1-q)\,(1-d)\,P_0]\left[\frac{1}{1+r} + \frac{1}{(1+r)^2} + \cdots\right]$$
$$= P_0 + [1 + q(u+d) - d]\;P_0\,\frac{1/(1+r)}{1-1/(1+r)}$$
$$= P_0\,[\,1 + r + q(u+d) - d\,]/r.$$

(Note that we need $r > 0$ for convergence of the sum.) If $V_0 > I$, the investment is made and the firm gets $V_0 - I$; if $V_0 < I$, the investment is not made and the firm gets 0; if $V_0 = I$, the firm is indifferent between investing and not investing and gets zero in either case. Let Ω_0 denote the net payoff of the project to the firm, if it is forced in period 0 to decide whether to invest, on a now-or-never basis. Thus we have shown that

$$\Omega_0 = \max\,[\,V_0 - I, 0\,]. \tag{1}$$

Now consider the actual situation, where the investment opportunity remains available in future periods. Here the period-0 decision involves a different tradeoff: invest now, or wait and do what is best when period 1 arrives. To assess this, the firm must look ahead to its own actions in different future eventualities. From period 1 onward the conditions will not change, so there is no point postponing any profitable projects beyond period 1. Hence we need look ahead only as far as period 1.

Suppose the firm does not invest in period 0, but instead waits. In period 1 the price will be

$$P_1 = \begin{cases} (1+u)\,P_0 & \text{with probability } q, \\ (1-d)\,P_0 & \text{with probability } 1-q. \end{cases}$$

It will stay at this level for periods 2, 3, ... The present value of this stream of revenues, discounted back to period 1, is

$$V_1 = P_1 + P_1/(1+r) + P_1/(1+r)^2 + \cdots$$
$$= P_1\,(1+r)/r.$$

For each of the two possibilities (the price going up or down between periods 0 and 1), the firm will invest if $V_1 > I$, realizing a net payoff

$$F_1 = \max\,[\,V_1 - I\,, 0].$$

This outcome of future optimal decisions is sometimes called the *continuation value*. From the perspective of period 0, the period-1 price P_1, and therefore the values V_1 and F_1, are all random variables. Let $\mathcal{E}_0$ denote the expectation (probability-weighted average) calculated using the information available at period 0. Then we have

$$\mathcal{E}_0[F_1] = q\;\max[\,(1+u)\,P_0\,(1+r)/r - I\,, 0] \\ + (1-q)\;\max\,[\,(1-d)\,P_0\,(1+r)/r - I\,, 0]. \tag{2}$$

This could be called the expected continuation value, or just the continuation value, with the expectation being understood.

Now return to the decision at period 0. The firm has two choices. If it invests immediately, it gets the expected present value of the revenues minus the cost of investment, $V_0 - I$. If it does not, it gets the continuation value $\mathcal{E}_0[F_1]$ derived above, but that starts in period 1 and must be discounted by the factor $1/(1+r)$ to express it in period-0 units. The optimal choice is obviously the one that yields the larger value. Therefore the net present value of the whole investment opportunity optimally deployed, which we denote by F_0, is

$$F_0 = \max\left\{\,V_0 - I,\;\frac{1}{1+r}\;\mathcal{E}_0[F_1]\,\right\}. \tag{3}$$

The firm's optimal decision is the one that maximizes this net present value.

This captures the essential idea of dynamic programming. We split the whole sequence of decisions into two parts: the immediate choice, and the remaining decisions, all of whose effects are summarized in the continuation value. To find the optimal sequence of decisions we work backward. At the last relevant decision point we can make the best choice and thereby find the continuation value (F_1 in our example). Then at the decision point before that one, we know the expected continuation value and therefore can optimize the current choice. In our example there were just two periods and that was the end of the story. When there are more periods, the same procedure applies repeatedly.

The decision where the investment opportunity remains available at period 1 is less constrained than the one where it must be made on a now-or-never basis in period 0. Equation (1) shows the net payoff Ω_0 for this latter case; since that situation terminates the decision process at time 0, let us call it the *termination value* at time 0. Now we have the net worth F_0 of the less constrained decision problem from equation (3). The difference $(F_0 - \Omega_0)$ is just the value of the extra freedom, namely the *option to postpone* the decision.

In Chapter 2 we calculated the value of the investment opportunity, F_0, and the termination value, Ω_0, for some specific cases of this general model, where the parameters P_0, q, etc., were given numerical values. Readers can now refer back to those examples and place them in the context of the general theory. Here we point out one feature of those results, by reference to Figure 2.4, which shows these values as functions of the initial price, P_0. When P_0 exceeds the critical level of 249, the firm finds it optimal to invest at once. Then the option to postpone is worthless, and F_0 coincides with Ω_0, which equals $V_0 - I$ in this range of the price. When $P_0 < 249$, it is optimal to wait; then the graph of F_0 lies above that of Ω_0. A similar property holds as other parameters are varied in other figures in Chapter 2. The idea is that the critical point where immediate investment just becomes optimal is found where the lines representing the value of the full opportunity, F_0, and the termination value, Ω_0, meet.

To get a better idea of the factors that affect the value of the option to postpone, let us examine more closely the sources of the differences between F_0 and Ω_0. First, by postponing the decision the firm gives up the period-0 revenue P_0. This difference favors immediate action. Second, postponing the decision also means postponing the cost of investment; this favors waiting since the interest rate is positive. (More generally, the cost of investment could itself be changing over time, and that would bring new considerations; for example, if the firm expects capital equipment to get cheaper over time, that is an additional reason for waiting.) Third, and most important, waiting allows a separate optimization in each of the contingencies of a price rise and a price fall, whereas immediate action must be based on only the average of the two. This ability to tailor action to contingency, specifically to refrain from investment if the price goes down, gives value to the extra freedom to wait.[1]

1.B Many Periods

We now generalize the two-period example above. Our applications in subsequent chapters are mostly to situations where time is continuous and the uncertainty takes the form of Wiener processes or more general diffusion processes for the state variables. However, in this subsection we develop the theory of dynamic programming in a setting where uncertainty is modelled using discrete-time Markov processes. Some general properties are easier to demonstrate in this format. Also, the setting of the rest of the book is a limiting case. Diffusion processes are Markov processes, and as we saw in Chapter 3,

[1] In technical terms, the maximum is a convex function, so by Jensen's Inequality the average of the separate maxima in equation (2) is greater than the maximum of the corresponding averages.

they can be regarded as limits of random walks in discrete time as the length of each time period and of each step of the walk both become small in a suitable way.

With our application to investment in mind, we will refer to the decisions of a firm, but the theory is of course perfectly general. The firm's current status as it affects its operation and expansion opportunities is described by a state variable x. For simplicity of exposition we take this to be a scalar (real number), but the theory extends readily to vector states of any dimension. At any date or period t, the current value of this variable x_t is known, but future values x_{t+1}, x_{t+2}, ... are random variables. We suppose that the process is Markov, that is, all the information relevant to the determination of the probability distribution of future values is summarized in the current state x_t.

At each period t, some choices are available to the firm, and we represent them by the control variable(s) u. In the above example where the only choice was whether to invest at once or wait, we could let u be a scalar binary variable, whose value 0 represents waiting and 1 represents investing at once. In other applications, for example, if the scale of investment is a matter of choice, u can be a continuous variable. If the firm has choices in addition to those bearing on investment, for example, hiring labor at time t, then u can be a vector. The value u_t of the control at time t must be chosen using only the information that is available at that time, namely x_t.

The state and the control at time t affect the firm's immediate profit flow, which we denote by $\pi_t(x_t, u_t)$. Here the relevant control variable u_t might be the quantity of labor hired or raw materials purchased. The x_t and u_t of period t also affect the probability distribution of future states. Here u_t can be the amount of investment or R&D, or even a decision to abandon the enterprise. Let $\Phi_t(x_{t+1} \mid x_t, u_t)$ denote the cumulative probability distribution function of the state next period conditional upon the current information (state and control variables).

The discount factor between any two periods is $1/(1+\rho)$, where ρ is the discount rate. The aim is to choose the sequence of controls $\{u_t\}$ over time so as to maximize the expected net present value of the payoffs. Sometimes we will force the decision process to end at some period T, with a final payoff that depends on the state reached; we denote this *termination payoff* function by $\Omega_T(x_T)$.

We are ready to apply the basic dynamic programming technique. Remember the idea is to split the decision sequence into two parts, the immediate period and the whole continuation beyond that. Suppose the current date is t and the state is x_t. Let us denote by $F_t(x_t)$ the outcome—the expected net present value of all of the firm's cash flows—when the firm makes all decisions optimally from this point onwards.

When the firm chooses the control variables u_t, it gets an immediate profit flow $\pi_t(x_t, u_t)$. At the next period $(t+1)$, the state will be x_{t+1}. Optimal decisions thereafter will yield, in the notation we have established, $F_{t+1}(x_{t+1})$. This is random from the perspective of period t, so we must take its expected value $\mathcal{E}_t[F_{t+1}(x_{t+1})]$. That is what we called the continuation value.[2] Discounting back to period t, the sum of the immediate payoff and the continuation value is

$$\pi_t(x_t, u_t) + \frac{1}{1+\rho}\ \mathcal{E}_t[F_{t+1}(x_{t+1})].$$

The firm will choose u_t to maximize this, and the result will be just the value $F_t(x_t)$. Thus

$$F_t(x_t) = \max_{u_t} \left\{ \pi_t(x_t, u_t) + \frac{1}{1+\rho}\ \mathcal{E}_t[F_{t+1}(x_{t+1})] \right\}. \qquad (4)$$

The idea behind this decomposition is formally stated in Bellman's Principle of Optimality: *An optimal policy has the property that, whatever the initial action, the remaining choices constitute an optimal policy with respect to the subproblem starting at the state that results from the initial actions.* Here the optimality of the remaining choices u_{t+1}, u_{t+2}, etc., is subsumed in the continuation value, so only the immediate control u_t remains to be chosen optimally.

The result of this decomposition, namely equation (4), is called the *Bellman equation*, or the *fundamental equation of optimality*. To reiterate, the first term on the right-hand side is the immediate profit, the second term constitutes the continuation value, and the optimum action this period is the one that maximizes the sum of these two components.

In the two-period example, immediate investment gave $V_0 - I$, waiting had no period-0 payout but only a discounted continuation value $\mathcal{E}_0[F_1]/(1+r)$, and the optimal binary choice between these alternatives yielded the larger of these two. Thus our earlier equation (3) is a special case of the general Bellman equation (4).

[2]The expectation notation is generally clear enough. However, to make it precise that the information at time t includes the state and the control at that time, we state it formally once for reference:

$$\mathcal{E}_t[F_{t+1}(x_{t+1})] = \int F_{t+1}(x_{t+1})\ d\Phi_t(x_{t+1} \mid x_t, u_t),$$

where the range of integration is that over which x_{t+1} is distributed, namely, the support of $\Phi_t(x_{t+1} \mid x_t, u_t)$.

If the many-period problem has a fixed finite time horizon T, we can similarly start at the end and work backward. At the end of the horizon the firm gets a termination payoff $\Omega_T(x_T)$. Then the period before,

$$F_{T-1}(x_{T-1}) = \max_{u_{T-1}} \left\{ \pi(x_{T-1}, u_{T-1}) + \frac{1}{1+\rho}\ \mathcal{E}_{T-1}[\Omega_T(x_T)] \right\}.$$

Thus we know the value function at $T-1$. That in turn allows us to solve the maximization problem for u_{T-2}, leading to the value function $F_{T-2}(x_{T-2})$, and so on. At one time this was thought to be too complex a procedure to be practicable, and all kinds of indirect methods were devised. However, advances in computing have made the backward calculation remarkably usable, and several of our numerical simulations in later chapters use it. Later in this chapter we will offer an example of it.

1.C Infinite Horizon

If there is no fixed finite time horizon for the decision problem, there is no known final value function from which we can work backward. Instead, the problem gets a recursive structure that facilitates theoretical analysis as well as numerical computation. The crucial simplification that an infinite horizon brings to equation (4) is independence from time t as such. Of course the current state x_t matters, but the calendar date t by itself has no effect. This works provided the flow profit function π, the transition probability distribution function Φ, and the discount rate ρ are themselves all independent of the actual label of the date, a condition that is satisfied or assumed in many economic applications.

In this setting, the problem one period hence looks exactly like the problem now, except of course for the new starting state. Therefore the value *function* is common to all periods, although of course it will be evaluated at different points x_t. Therefore we write the function as $F(x_t)$ without any time label on the function symbol. The Bellman equation for any t becomes

$$F(x_t) = \max_{u_t} \left\{ \pi(x_t, u_t) + \frac{1}{1+\rho}\ \mathcal{E}_t[F(x_{t+1})] \right\}.$$

Since x_t and x_{t+1} could be any of the possible states, write them in general form as x and x'. Then, for all x, we get

$$F(x) = \max_{u} \left\{ \pi(x, u) + \frac{1}{1+\rho}\ \mathcal{E}[F(x') \mid x, u] \right\}, \qquad (5)$$

where we have now denoted the expectation as conditioned on the knowledge of the current period's x and u. This is the Bellman equation for the infinitely repeating, or recursive, dynamic programming problem.

Now that we have no fixed terminal date from which to work backward, we seem to have lost an explicit or constructive way to find the value function F. And without knowing the function F, we cannot find the optimal control u by solving the maximization problem on the right-hand side of the Bellman equation. Thus we need assurance that a solution actually exists, and a way to find it. Luckily, neither question is very difficult.

The recursive Bellman equation (5) can be thought of as a whole list of equations, one for each possible value of x, with a whole list of unknowns, namely all the values $F(x)$. If x took on only a finite number of discrete values x_i, this would be a simultaneous system with exactly as many equations as the number of unknowns $F(x_i)$. More generally, we can regard (5) as a *functional equation*, with the *whole function* F as its unknown.

Despite superficial appearances, this equation is not linear. The optimal choice of u depends on all the values $F(x')$ that appear, weighted by the appropriate probabilities, in the expectation on the right-hand side. When this optimal control is substituted back, the result can be nonlinear in the $F(x')$ values.

In general we do not know whether nonlinear functional equations have solutions, let alone unique ones. Fortunately, the recursive Bellman equation has a very special structure that allows one to prove existence and uniqueness of a solution function $F(x)$ under conditions typical of our economic applications. This is a technical side issue for our more practical concerns. We include a brief sketch of the ideas in Appendix A to this chapter, and interested readers can find excellent treatments in more theoretical books, for example, Stokey and Lucas with Prescott (1989, Chapters 4,9). But the technical argument does have an indirect payoff: it is essentially a practical solution method.

This takes the form of an iterative procedure. Start with *any* guess for the true value function, say $F^{(1)}(x)$. Use it on the right-hand side of equation (5) and find the corresponding optimal choice rule u^1, which can now be expressed as a function of x alone. Substituting it back, the right-hand side becomes a new function of x; call it $F^{(2)}(x)$. Now use it as the next guess of the true value function, and repeat the procedure. Then the successive guesses $F^{(3)}(x)$, $F^{(4)}(x)$, etc., will converge to the true function. Convergence is guaranteed no matter how bad the initial guess, but of course with a good initial guess the procedure will reach the desired accuracy of the approximation in fewer steps.

The key lies in the factor $1/(1+\rho)$ on the right-hand side. This being less than 1, it scales down, or contracts, any errors in the guess from one step to the next. As long as the profit flows are bounded, any errors in the choice of u cannot blow up. Gradually, only the correct solution is left.

This procedure is very easy to understand, program, and compute. It can take a long time on the computer, especially if the discount rate ρ is very small so that each step does only a little scaling down. However, that is no longer a prohibitive consideration when individuals can leave their own personal computers running for days without tying up scarce mainframe resources. Therefore this method is increasingly used in many applications, and even in econometric work. We will give a numerical example of this method, too, later in the chapter.

1.D Optimal Stopping

One particular class of dynamic programming problems is very important for our applications. Here the choice in any period is binary. One alternative corresponds to stopping the process to take the termination payoff, and the other entails continuation for one period, when another similar binary choice will be available. In the investment model of Chapter 2 and the first section of this chapter, stopping corresponds to making the investment, and continuation corresponds to waiting. Here continuation does not generate any profit flow within the period. However, in other contexts there may be some such flow. For example, for a firm in bad economic conditions that is contemplating shutdown, continuation may give a profit flow (positive or negative), and termination may yield some scrap value of the plant and equipment, minus any severance payments the firm is required to make to its workers and other costs of site restoration, breaking contracts, etc.

Let $\pi(x)$ denote the flow profit, and $\Omega(x)$ the termination payoff. Then the Bellman equation becomes

$$F(x) = \max \left\{ \Omega(x)\, ,\ \pi(x) + \frac{1}{1+\rho}\ \mathcal{E}[F(x') \mid x] \right\}. \qquad (6)$$

For some range of values of x, the maximum on the right-hand side of this will be achieved by termination, and for other values of x it will be achieved through continuation. In general this division could be arbitrary; intervals where termination is optimal could alternate with ones where continuation is optimal. However, most economic applications will have more structure. There will be a single cutoff x^*, with termination optimal on one side and continuation on the other. For example, in the investment problem of Chapter 2 we had a critical level of the initial price, $P_0 = 249$, such that in period 0, investment was optimal to its right but waiting was optimal to its left. All of our applications will have a similar property, and in each case it will be intuitively clear which action is optimal on which side of the threshold or cutoff point. To complete the reader's understanding of this result, we should explain the general conditions that lead to it. We state these intuitively here, and explain them somewhat more formally in Appendix B.

For sake of definiteness we concentrate on the case where continuation is optimal for $x > x^*$ and stopping is optimal for $x < x^*$. Let us examine the forces that will make continuation more attractive relative to termination for higher values of x. First, the immediate profit from continuation should become larger relative to the termination payoff. Since the former is a flow and the latter is a stock, we need to express them in comparable terms. The precise condition turns out to be that

$$\pi(x) - \frac{\rho}{1+\rho}\,\Omega(x) \qquad (7)$$

should be increasing as x increases. Second, any current advantage should not be likely to be reversed in the near future. For this, we need positive serial correlation, or persistence, in the stochastic process of evolution of x. To be more precise, if this period's x rises, the conditional distribution $\Phi(x'|x)$ of next period's values x' should give greater weight to larger values, that is, it should shift everywhere to the right. (In technical language, this is the concept of "first-order stochastic dominance.") These two conditions together are sufficient to ensure the desired result.

If the expression in (7) is decreasing as x increases, then continuation will be optimal to the left and termination to the right of x^*. Note that the second condition stays unchanged; we do not switch it to require negative persistence of the stochastic process.

We repeat that both of these properties will always be satisfied for our applications. The first is easily verified in each instance; the second is true for random walks, Brownian motion, mean-reverting autoregressive processes, and indeed in almost all economic applications we can think of.

For ease of notation in conveying the concept, we did not allow the profit or terminal payoff to depend on time t as such, but that extension presents no difficulty. The threshold simply becomes a function of time, $x^*(t)$. This will be the case in much of the rest of the chapter and in many of our applications.

1.E Continuous Time

Now return to the general control problem of Section 1.B, but suppose each time period is of length Δt. Ultimately we are interested in the limit where

Δt goes to zero and time is continuous. We write $\pi(x, u, t)$ for the *rate* of the profit flow, so that the actual profit over the time period of length Δt is $\pi(x, u, t)\,\Delta t$. Similarly, let ρ be the discount rate per unit time, so the total discounting over an interval of length Δt is by the factor $1/(1 + \rho\,\Delta t)$.

The Bellman equation (5) now becomes

$$F(x,t) = \max_u \left\{ \pi(x,u,t)\,\Delta t + (1+\rho\,\Delta t)^{-1}\ \mathcal{E}[F(x', t+\Delta t)|x,u] \right\}.$$

Multiply by $(1 + \rho\,\Delta t)$ and rearrange to write

$$\begin{aligned}\rho\,\Delta t\, F(x,t) &= \max_u \left\{ \pi(x,u,t)\,\Delta t\,(1+\rho\,\Delta t) + \mathcal{E}[F(x', t+\Delta t) - F(x,t)] \right\} \\ &= \max_u \left\{ \pi(x,u,t)\,\Delta t\,(1+\rho\,\Delta t) + \mathcal{E}[\Delta F] \right\}.\end{aligned}$$

Divide by Δt and let it go to zero. We get

$$\rho\, F(x,t) = \max_u \left\{ \pi(x,u,t) + \frac{1}{dt}\,\mathcal{E}[dF] \right\}, \qquad (8)$$

where $(1/dt)\,\mathcal{E}[dF]$ is the limit of $\mathcal{E}[\Delta F]/\Delta t$. We must remember that the expectation is conditioned on the current x and u, and we must remember to include the influence of changes in both x and t when we calculate the change in $F(x, t)$ over the interval dt.

This form of the Bellman equation makes explicit the idea that the entitlement to the flow of profits is an asset, and that $F(x, t)$ is its value. On the left-hand side we have the normal return per unit time that a decision maker, using ρ as the discount rate, would require for holding this asset. On the right-hand side, the first term is the immediate payout or dividend from the asset, while the second term is its expected rate of capital gain (loss if negative). Thus the right-hand side is the expected total return per unit time from holding the asset. The equality becomes a no-arbitrage or equilibrium condition, expressing the investor's willingness to hold the asset. The maximization with respect to u means that the current operation of the asset is being managed optimally, of course bearing in mind not only the immediate payout but also the consequences for future values.

The limit on the right-hand side depends on the expectation corresponding to the random x' a time Δt later. There are two classes of stochastic processes in continuous time that allow such limits in a form conducive to further analysis and solution of the function $F(x, t)$ in the continuation region. Luckily they are particularly useful for many economic applications. In fact they are just the Ito and Poisson processes we discussed in Chapter 3. We will develop the theory of dynamic programming in their specific contexts in the next two subsections.

The above analysis is local to the short time interval $(t, t+dt)$, and the resulting equation holds for any t. We can complete the analysis by choosing a finite time horizon T and imposing a terminal payoff, or letting the horizon be infinite and using the recursive structure, or some other way. In any of these, rigorous mathematical proofs of existence and uniqueness of solutions become quite hard in continuous time. Since the details are immaterial for our applications, we omit them and refer the reader to Fleming and Rishel (1975) or Krylov (1980).

Our mathematics has been simplified in another respect. We have treated the limit to continuous time in a very casual and heuristic way, and will continue to do so. However, it is fair to warn the reader that some quite tricky issues are hidden, and must be handled carefully in more rigorous treatments. In discrete time, we stipulated that the action u_t taken in the current period t could depend on the knowledge of the current state x_t, but not on the random future state x_{t+1}. In continuous time the two coalesce. We have to be careful not to allow choices to depend on information about the future, even about "the next instant." Otherwise we would be acting with the benefit of hindsight, and could make infinite profits. Technically this can be avoided by requiring the uncertainty to be "continuous from the right" in time while the strategies are "continuous from the left." Then any jumps in the stochastic processes occur *at an instant*, while the actions cannot change until *just after the instant*. For a discussion and rigorous analysis, see Duffie (1988, pp. 139–40).

1.F Ito Processes

The first continuous-time stochastic process that yields a simple form for (8) is the Ito process we discussed in Chapter 3. Equation (11) of that chapter defined the formula for its increment, which we recapitulate here, but now allowing the drift and the diffusion parameters to depend on the control variable as well as the state variable:

$$dx = a(x, u, t)\, dt + b(x, u, t)\, dz, \tag{9}$$

where dz is the increment of a standard Wiener process. As before, we write the profit flow as $\pi(x, u, t)$ and the value of the firm (asset) as $F(x, t)$.

Let x be the known starting position at time t, and $x' = x + dx$ the random position at the end of a small interval of time Δt. Ito's Lemma for such a process was stated in equation (25) of Chapter 3. Applying it to the

value function F, we have

$$\begin{aligned}\mathcal{E}[F(x+\Delta x, t+\Delta t)|x, u]& \\ = F(x,t) + \big[\, F_t(x,t) + a(x,u,t)\, F_x(x,t)& \\ \tfrac{1}{2}\, b^2(x,u,t)\, F_{xx}(x,t)\,\big]\, \Delta t + o(\Delta t)&,\end{aligned}$$

where $o(\Delta t)$ represents terms that go to zero faster than Δt. Then the "return equilibrium" condition (8) becomes

$$\begin{aligned}\rho\, F(x,t) = \max_u \big\{\, \pi(x,u,t) + F_t(x,t) + a(x,u,t)\, F_x(x,t)& \\ + \tfrac{1}{2}\, b^2(x,u,t)\, F_{xx}(x,t)\,\big\}&.\end{aligned} \tag{10}$$

We can express the optimal u as a function of $F_t(x,t)$, $F_x(x,t)$, $F_{xx}(x,t)$ as well as x, t, and the various parameters that govern the functional form of π, a, and b. Substituting this expression for the optimal u back into the right-hand side of equation (10), we get a partial differential equation of the second order, with F as the dependent variable and x and t as the independent variables. In general this equation is very complicated. However, in many applications we can develop ways to solve it analytically or numerically.

The solution methods are generally analogous to those for discrete time. If there is a fixed time limit T when a termination payoff $\Omega(x_T, T)$ is enforced, then the equation has a boundary condition

$$F(x, T) = \Omega(x, T) \quad \text{for all } x.$$

We can start at time T and work our way backward to find $F(x,t)$ for all earlier times. In fact, in practice we have to choose a discrete grid of values of x and t on which to calculate the solution. We will offer two examples of this procedure, one later in this chapter when we directly solve the underlying dynamic programming problem, and one in Chapter 10 where we solve the partial differential equation itself.

If the time horizon is infinite and the functions π, a, and b do not depend explicitly on time, then neither does the value function depend on time, and equation (10) becomes an ordinary differential equation with x as its only independent variable:

$$\rho\, F(x) = \max_u \{\, \pi(x,u) + a(x,u)\, F'(x) + \tfrac{1}{2}\, b^2(x,u)\, F''(x)\,\}. \tag{11}$$

Note that we have followed standard calculus notation and used primes to denote total derivatives of a function of one independent variable, and subscripts to denote partial derivatives of a function of several independent variables.