# 1. Introduction

A person behaving strategically acquires and processes information in a purposeful manner with a view to furthering her own ends, recognizing those ends may conflict with the goals of other people, and taking account of how they react to her decisions. This chapter and its sequel abstract from the strategic aspects to amplify the premise that strategic play is based on rational choices, focusing on dynamic problems and situations where there is uncertainty. The examples in this chapter concern important choices individuals confronted throughout the lifecycle, and generic investment choices made by firm managers.

Section 2 describes different choice sets, the notion of a preference relation over a choice set, and explains the two axioms that define rationality. A person is rational when her preferences over her list of choice alternatives, that is her choice set, are complete and transitive. These two axioms are necessary and sufficient to rank the elements of her choice set, and if the set is countable, we can construct an index called a utility function which prioritizes her preferences. This transforms the behavior of a rational person into an optimization problem where the decision maker chooses the alternative that maximizes her utility function on the choice set.

The remaining sections apply two optimization tools used repeatedly throughout this textbook, backwards induction and marginal adjustment, to solve dynamic optimization problems. The third section analyzes stopping problems. We begin with labor force entry and exit, when to leave school and when to retire. Then we consider discrete investment decisions by firms, replacing equipment, and prioritizing investment projects to be undertaken sequentially.

Continuous choices, or marginal adjustment, is topic of Section 4. It begins with optimization problems in consumption and savings over the lifecycle, typically made against a backdrop of fluctuating income and expenditure demands from children. This has an analogue in business we explain, where a firm with regional production sites optimally allocates an aggregate production target across the sites to minimize its costs. Smoothing consumption over the lifecycle resembles conservation; to round out the section we analyze a resource extraction problem, that is how to deplete a fixed resource stock to maximize firm value, and an internal investment problem, how much investment to undertake when all investment must be internally financed from profits.

The applications we consider in the next two sections all have closed forms, but this is not a generic feature of dynamic optimization problems. Section 5 provides a more abstract formulation of dynamic optimization problems, using the previous examples to highlight the various components. We define the economic environment in terms of a set of state variables that describe the state of play at each period or move for any given history thus far. We define the set of possible choices the player has at her disposal. The state of play next period is determined by a probability distribution which itself depends on the current state, as well as the choices of the decision maker. The decision maker has preferences over the states she visits throughout her lifetime, and maximizes an objective function based upon them. This formulation leads us into a discussion for how to search for a solution in more general terms, and analyze deviations from optimal behavior. We describe two algorithms that can be used to numerically solve dynamic optimization problems, and explain why the behavior of experimental subjects can be tested for opitmality even if the solution is unknown and the algorithms are too expensive to implement.

These general principles are illustrated by three dynamic optimization problems discussed in Section 6. The problems do not have closed form solutions, but they can be solved numerically. The first two, inventory control and scheduling production runs, are in the field of operations management. The final example, in human capital accumulation, is about allocating time between the demands of job and home, forging a career versus giving birth and nurturing offspring. In out concluding summary section, we argue that human experimentation and artificial intelligence are complementary inputs in developing and understanding dynamic optimization problems.

# 2. Axioms for Rational Choice

This section expalins what choice sets are and defines rationality. The first two axioms of rational behavior are that preferences are complete and transitive. They form the basis for three principles we carry forward throughout the book: backwards induction, marginal balancing, and dominance.

## 2.1 Choice sets

Choice sets circumscribe what is feasible for a player confronting a problem. A finite choice set of alternatives can be labeled as $C \equiv \{c_1, \ldots, c_J\}$, where selecting $c_j$ means choosing the $j^{th}$ alternative for each $j \in \{1, \ldots, J\}$. For example if $c_1$ is "offer money back guarantee on product" and $c_2$ is "do not offer a guarantee" then choice set is the two element set $C \equiv \{c_1, c_2\}$. Another example is how many letters to write seeking aid from potential supporters for a special interest. For all intents and purposes this is a unbounded counting number, meaning $C \equiv \{1, 2, \ldots\}$. A third example is how much overtime a wage earning full time employee might accept over and above a standard forty week. In this case the choice set is the closed interval of real numbers between zero and say forty, that is $C \equiv [0, 40]$.

In the examples above $C$ is respectively a finite set, the set of counting numbers, and a closed interval of (uncountable) real numbers. In other examples the choice set is a probability space, meaning the elements are positive and sum or integrate to the whole (such as dividing a product market into disjoint regions that cover the whole territory), a vector of real numbers (the budget for each regional marketing director), or a function (the compensation of the regional manager as a function of the region's sales).

Choices made within a dynamic context can be considered in a similar way. An application considered in the fourth section concerns the allocation of wealth to consumption over the lifecycle when borrowing and savings are possible. Suppose consumption is undertaken in $T$ periods, and let $c_t$ denote consumption in period $t \in \{1,\ldots,T\}$. Denote by $w$ the wealth available, and let $r_t$ denote the interest accumulated on savings held for $t$ periods beginning at date $t = 1$. We now define the choice set as a consumption sequence $\{c_1,\ldots,c_T\}$ that contains only positive elements and satisfies the budget constraint, that wealth exceeds the present value of expenditure. In terms of the notation we have developed:

$$C = \left\{ \{c_1,\ldots,c_T\} \text{ such that } \sum_{t=1}^{T}(1+r_t)c_t \le w \text{ and } c_t \ge 0 \text{ for all } t \in \{1,\ldots,T\} \right\}$$

The same principles also apply to dynamic optimization problems under uncertainty, but we postpone our discussion until Section 6, where we review solution algorithms to such problems as well.

## 2.2 Complete and transitive preference orderings

Preference ordering are relations that rank elements pair-wise in a choice set, where ties are permitted. We use the symbol $\succeq$ to denote a preference ordering over a choice set $C$. Suppose $c_1$ and $c_2$ are both elements of $C$. Intuitively, $c_1 \succsim c_2$ means that choosing $c_1$ is at least as good as choosing $c_2$. Also $c_1 \succsim c_2$ and $c_2 \succsim c_1$ means the choices are as good as each other, while $c_1 \succsim c_2$ but $c_2 \not\succsim c_1$ means $c_1$ is strictly better than $c_2$. A preference ordering over a choice set $C$ is complete if every pair of choices $c_i \in C$ and $c_j \in C$ can be ranked by the preference ordering $\succeq$. Thus $C$ is complete if and only if $c_i \succsim c_j$ and/or $c_j \succsim c_i$ for all $c_i \in C$ and $c_j \in C$. Therefore a preference ordering is incomplete if there is a pair of alternatives such that $c_i \not\succsim c_j$ and $c_j \not\succsim c_i$. It is hard to interpret someone saying that swimming is not at least as enjoyable as running, and in the same breath, that running is not at least as enjoyable as swimming either.

The example about preferences over athletic activities suggests that clear headed people have complete preferences. But we might argue that the way choices can create confusion. Giving restaurant diners a menu in a foreign language they barely understand might elicit different choices than if the same menu were presented in their native tongue. Is the axiom of complete preferences violated if diners who not multilingual make different choices, depending on the language in which the menu is printed? The answer depends on whether diners only care about the food they eat, or whether they care about the menu as well.

More generally, suppose the $j^{th}$ choice yields a payoff of $c_j \in C$, which is the unique solution to a complicated mathematical problem denoted $P_j$ In the restaurant example, the mapping from $P$ to $C$ is a translation from the foreign language into one's own language. Presented with the solution set $\{c_1,\ldots,c_J\}$ for $J$ distinct problems, an experimental subject typically expresses a set of complete preferences. (If $c_j$ were real numbers representing monetary payoffs, then surely $c_i \succsim c_j$ if and only $c_i \geq c_j$.) But presented the problem set $P \equiv \{P_1, \ldots, P_J\}$, and told that picking $P_j \in P$ induces the solution to the problem as an outcome, the subject would not necessarily rank the choice set $\{P_1, \ldots, P_J\}$ the same way. Are $P$ and $C$ different choice sets even though making choice $j$ induces the same outcome for all $j \in \{1, \ldots, J\}$? We take the alternative view, that $P$ and $C$ are different representations of the same choice set. Consequently we regard this book as a guide to strategic play, because it helps readers play more rationally in strategic situations.

The second axiom of rationality concerns triplets in the choice set. Suppose $c_i,\ c_j$ and $c_k$ are all elements of $C.$ The transitivity axiom states that if $c_i \succeq c_j$ and $c_j \succeq c_k$ then $c_i \succeq c_k.$ When individuals and organizations violate transitivity, they are rife for exploitation. For example suppose three countries belong to the European Union, who annually rotate the leadership of the European Union. England prefers capitalism to social contract, Germany prefers social contract to agricultural subsidies, and France prefers agricultural subsidies to capitalism. Treated as a whole, the leadership preferences of the EU are intransitive. If leadership confers the power to tax the Community and pay officials in Brussels for rewriting social policy, then it is easy to imagine that successive leaders employ a bureaucracy that continously updates documents on social policy without converging.

## 2.3 Utility functions

In business the elements of the firm's choice set are typically valued by the additional wealth they confer upon the firm. When the firm's underlying preference ordering is based on profitability, and each choice can be evaluated, its preferences are complete and transitive. Thus for many businesses, rational choice amounts to maximizing wealth. More generally a utility function $u(c),$ defined from the choice set $C$ on to the real line, represents a preference ordering $\succeq$ if and only if $c_i \succeq c_j$ whenever $u(c_i) \geq u(c_j)$ for all $c_i$ and $c_j$ in $C.$ For example, in a business application $u(c)$ might represent the value of the firm when choice $c \in C$ is selected.

If preferences defined on a countable choice set are rational, that is complete and transitive, then a utility function representing those preferences invariably exists. The following algorithm constructs a utility function for this case. Starting at $c_1$ we pick any number on the open $(0,1)$ interval, called $u_1$. If both $c_1 \succeq c_2$ and $c_2 \succeq c_1$, then we set $u_2 = u_1.$ Otherwise we pick any $u_2 \in (0, u_1)$ if $c_1 \succeq c_2,$ and $u_2 \in (u_1, 1)$ if $c_2 \succeq c_1.$ Proceeding in this fashion, the utility for choice $c_j$, called $u_j,$ is compared pairwise with all lower ranked indexes. If both $c_j \succeq c_k$ and $c_j \succeq c_k$ for some $k < j,$ then we set $u_k = u_j.$ Otherwise

$$0 < \max\left\{u_k : c_j \succeq c_k \text{ and } k < j\right\} < u_j < \min\left\{u_k : c_k \succeq c_j \text{ and } k < j\right\} < 1$$

Since open intervals on the real line have interior points, choosing $u_j$ this way is possible for all $j \in \{1, 2, \ldots\}.$ The utility function defined this way maps $C$ to $(0,1).$

Utility functions are not uniquely defined. If $u(c)$ is a utility function for $\succeq$ on $C,$ then any monotonic transformation of $u(c)$ is also a utility function for the same prerefereces on the same choice set, because all the defining inequalities are preserved. For example expressing porfits in dollars or cents has no effect on decision making providing a uniform accounting convention is adopted to avoid confusion. Similarly in the case of countable choice sets, having defined a utility function on $(0,1)$ by $u_j$ as described above, the transformed utility functions $u_j' \equiv \log(u_j)$, or $u_j'' \equiv u_j^2$, for all $j \in \{1,2,\ldots\}$, serve equally for representing the preference ordering $\succeq$ on $C$.

# 3. Stopping and Starting

Many investment decisions can be cast as optimal stopping problems. An example of an optimal stopping problem is a software developer deciding when to stop developing a new product and release its best version to the market. The firm compares the benefits of a further marginal delay or continuation, that allows a little more product development, with the loss from not exercising an option now, lost or delayed sales from not introducing the product earlier and postponing the reallocation of software programmers to other projects. In human resource management optimal stopping is used to characterize entry into and exit from the labor force. When students enter the labor force their investment in general education drops precipitously, but the jobs they accept pay wages and confer working experience, a different type of human capital. When workers retire from the labor force they typically stop saving too. Periodically replacing durable equipment, and more generally, replacement investment, can also be modeled as a stopping problem. We explore these applications in this section, and then, using software release as an example, we consider the related problem of prioritizing investment projects that must be completed sequentially rather than undertaken simultaneously.

## 3.1 Retirement

Within the realm of human resource management, two of the most important decisions a person makes is when to enter the labor market (and stop schooling) and when to exit it (and retire). Framing the questions that way ignores some of the subtlety of the questions. Neither of these decisions are irrevocable. For example many students receive specialized training after completing their general training. Similarly retirees return to work on a part time basis, sometimes treating retirement as an opportunity to start a new career. For example how much part time and summer work to do to support continued schooling, whether to quit work temporarily to take a professional degree such a  law degree or an MBA, and whether to work fewer hours and thus taper off labor force participation, rather than transit from full time work to no work at all. Nevertheless the conditions we derive below yield insight into considerations that also apply to more complicated settings. The simplifications help us to focus on two basic questions, how much value extra schooling creates relative to its costs, and the value of continuing to work more years for a shorter but wealthier retirement phase.

Roughly speaking people enter the labor force when they leaves school, they remain in the labor force for many years, sometimes interrupted (by childbearing for example), and then quit the workforce. Compared with a century ago, males are entering the workforce later, the age at entry slowly increasing with greater college attendance. They are retiring at an earlier age, and living longer. More females are entering the workforce than before, but the entrants are also older and more educated. They take less time off work per birth, have fewer births, are retiring later in their lives, and are also living longer.

The longer people work, the less time they spend in retirement before dying, but the more resources they accumulate for retirement. Both factors increase the amount that can be consumed each year in retirement. Here a person trades off the value of consuming more in retirement with fewer years in life to enjoy their wealth. In this application we suppose that the annual sum of the pleasures from retirement plus pecuniary retirement benefits have a monetary value of $b(t)$ where $t$ is the person's age. We assume that $b(t)$ is an increasing function; as people age their retirement benefits increase, as does the effort of continuing to work. The current value of earnings from working at age $t$ is denoted by $y(t)$. We assume $y(t)$ is declining with age, or is a concave function with a maximum. We also assume it is optimal to work when young, so that say $y(25) > b(25)$. Letting $T$ denote age at death, and $R$ age at retirement, the person chooses $R$ to maximize

$$\int_{25}^{R} y(t)\exp(-rt)dt + \int_{R}^{T} b(t)\exp(-rt)dt$$

The assumptions above imply that the functions $y(t)$ and $b(t)$ cross once. Denoting the optimal age of retirement by $R^o$, it is straightforward to see that $R^o$ uniquely solves $y(R^o) = b(R^o)$. In this case the person works if and only if $y(t) > b(t)$. Although this is a dynamic model of optimal stopping, treating it as a static one yields the solution.

## 3.2 Schooling

At the end of the career profile is the decision to retire; at the beginning of the career cycle is the decision to quit school. How much schooling is optimal? More schooling delays the earnings phase, but raises the wage rate and employment opportunities during it. We now assume that study at school and university is a costly activity that delays the onset of lifetime earnings. The current value of lifetime earnings at the time a student decides to quit study and join the workforce is

$$h_0 + h_1\tau$$

where $h_0$ is the lifetime earnings of unskilled laborer, and $h_1$ is the rate at which lifetime earnings increase with more study. We also assume that both work and study have nonpecuniary features, and that the net benefit from study is $b$ per unit of time, after accounting for school fees per year. Therefore the present value of studying for $\tau$ amount of time, and then accepting a job is

$$\exp(-r\tau)(h_0 + h_1\tau)$$

The nonpecuniary benefits from studying until $\tau$ are

$$\int_0^\tau b\exp(-r\tau)dt = \frac{b}{r}[1 - \exp(-r\tau)]$$

The student chooses $\tau$ to maximize her objective function, which is the sum of the human capital component and the nonpecuniary benefit component. The first derivative of this sum with respect to $\tau$ is

$$\exp(-r\tau)[h_1 - r(h_0 + h_1\tau) + b]$$

We define $\tau^*$ as the value of $\tau$ which equates this expression to zero and thus uniquely solves first order condition:

$$b + h_1 - r[h_0 + h_1\tau^*] = 0$$

Each term has an intuitive meaning. Extending the schooling phase by a marginal amount yields nonpecuniary benefits of $b$. The second term, $r[h_0 + h_1\tau^*]$, is the interest lost by marginally delaying lifetime career earnings. The last term is the increment in lifetime earnings $h_1$. Solving for $\tau^*$ we obtain

$$\tau^* = \frac{b + h_1 - rh_0}{rh_1}$$

To fully solve this problem we check the boundary conditions for this problem. Let $\tau^0$ denote the solution to the optimization problem. For all values $\tau > \tau^*$ the objective function declines in $\tau$, because the first derivative is negative. Hence $\tau^0 \leq \tau^*$. The second derivative of the objective function can be expressed as

$$-r\exp(-r\tau)[h_1 - rh_1(\tau - \tau^*)]$$

Inspecting the first and second derivatives, we remark that the sum of lifetime earnings and net nonpecuniary benefits from schooling is a concave increasing function for all values of $\tau$ satisfying $0 \leq \tau < \tau^*$. Consequently it is optimal to acquire $\tau^*$ schooling if and only if the net benefit from the first unit of schooling is positive, or

$$h_1 - rh_0 + b > 0$$

We conclude $\tau^0 = \tau^*$ if the inequality above is satisfied and $\tau^0 = 0$ otherwise.

Conduct an experiment with this specification, and plot the derivative of the objective function. Do experimental subjects acquire too little or too much education? How much do they lose relative to the optimum as a fraction of the optimal value of education

## 3.3 Renewal

The periodic replacement of a durable good is another example of an optimal stopping problem. The decision maker chooses when to liquidate an aging asset and begin again. For example, suppose the productive asset deteriorates at a constant rate over time, denoted by $\delta$, yielding a service flow worth $q\exp(-\delta t)$ at age $t$. We also assume the scrap value of the old machine is $\gamma$, the firm buys a new machine for $\gamma_0$, and downtime at the plant from replacing the machine is $\rho$. Given a constant interest rate $r$, the net present value from buying new machine for $\gamma_0$ at date 0, running it from date $\rho$ to date $\tau + \rho$, and then selling it for scrap is:

$$-\gamma_0 + \int_{\rho}^{\rho+\tau} q \exp(-rt - \delta t)dt + \gamma \exp(-r\rho - r\tau)$$

These assumptions ensure that the optimal policy is to replace the asset at evenly spaced intervals, if at all. Intuitively, the problem facing the firm only depends on the age of the existing machinery, not the equipment replacement policy that has been pursued in the past. It is convenient to normalize the machine price $\gamma_0$ and the scrap value $\gamma$ in terms of the service units $q$, because in this way we can reduce the number of parameters characterizing the problem. Setting $q = 1$, the value of the firm from installing new machinery every $\tau + c$ periods is then:

$$\left\{\int_{\rho}^{\rho+\tau} \exp(-rt - \delta t)dt + \gamma \exp(-r\rho - r\tau) - \gamma_0\right\} \sum_{s=0}^{\infty} \exp[-rs(\tau + \rho)]$$

Upon performing the integration and summing the geometric series, this formula simplifies to

$$\left\{\gamma \exp(-r\rho - r\tau) - \gamma_0 + (r + \delta)^{-1}[1 - \exp(-r\tau - \delta\tau)] \exp(-r\rho)\right\} \{1 - \exp[-r(\tau + \rho)]\}^{-1}$$

A closed form solution for the optimal replacement vintage $\tau_0$ does not exist, but this problem can be solved numerically for different values of $(\rho, \gamma, \gamma_0, \delta, r)$.

## 3.4 Project priorities

The renewal problem derives its dynamics from the fact that the cost of maintaining old equipment must be balanced against the amortization costs of purchasing new replacement equipment, not just once, but as warranted in the future. Similar dynamic issues arise from ordering projects that cannot be undertaken simultaneously, but must be undertaken sequentially. This constraint might arise because of a scarce factor that might be used on all of them. For example a drilling rig might be used to explore one of several oil tracts at a time, each of which is a separate investment project. The expertise of a manager might be required by several projects, none of which can be undertaken without her active involvement.

Suppose an expert is required to supervise two investment projects, called $A$ and $B$, with net present values of $v_A$ and $v_B$ respectively. Project $A$ lasting $a$ periods and Project $B$ for $b$ periods. Both projects have positive present value, so both would be commenced immediately if that was feasible. However the demands of each project require the undivided attention of the expert. Which one should she undertake first?

Project $A$ should be undertaken before Project $B$ if the present value of doing $A$ and then $B$ exceeds the present value of doing $B$ and then $A$:

$$v_A + \beta^a v_B > v_B + \beta^b v_A$$

Subtracting $(\beta^a v_B + \beta^b v_A)$ from both sides of the inequality, and then dividing though by the product $(1 - \beta^a)(1 - \beta^b)$ gives:

$$\frac{v_A}{1-\beta^a} > \frac{v_B}{1-\beta^b}$$

Noting:

$$\frac{v_A}{1-\beta^a} = v_A(1+\beta^a+\beta^{2a}+\ldots)$$

it immediately follows that $A$ should be completed first if an infinite sequence of undertaking type $A$ projects, one after another, yields a higher net present value than undertaking an infinite sequence of type $B$ projects. Another interpretation of the opitmal sequencing rule from the fact that:

$$1-\beta^a = (1-\beta)\sum_{s=0}^{a-1}\beta^s$$

and similarly for $B.$ Substituting for $1-\beta^a$ and $1-\beta^b$ in the inequality determining opitmal precedence, $A$ should be undertaken before $B$ if

$$\frac{v_A}{\sum_{s=0}^{a-1}\beta^s} > \frac{v_B}{\sum_{s=0}^{b-1}\beta^s}$$

The numerator is the net present value of the project in dollars say, while the denominator is the present value of receiving one dollar each period over the life of the project, or the number of discounted time periods. Thus the quotient is the net present value per unit of discounted time.

This formula for prioritizing two projects that take a known amount of time can be extended in several directions. For example if the amount of time is uncertain, and $a$ and $b$ are random variables, then the formula becomes:

$$\frac{v_A}{E\left[\sum_{s=0}^{a-1}\beta^s\right]} > \frac{v_B}{E\left[\sum_{s=0}^{b-1}\beta^s\right]}$$

This optimal rule also when to switch from one project to another if this is feasible. Suppose $A$ can undertaken in $I$ stages, denote by $a_i$ the length of the $i^{th}$ stage, and let $w_{ai}$ denote the present value of completing the $i^{th}$ stage as a stadn alone project. That is

$$a = \sum_{s=0}^{a-1} a_i$$

and the net present value of partially completing $H \leq I$ stages is

$$v_{AH} = \sum_{i=1}^{H}\beta^{(a_1+\ldots+a_{i-1})}w_{ai}$$

with $v_A = v_{AI}$. Similarly we suppose $B$ can undertaken in $K$ stages, denoting by $b_k$ the length of the $k^{th}$ stage, and $w_{bk}$ the stand alone present value of completing the $k^{th}$ stage. The index for ranking indivisable projects $A$ and $B$ also applies here (a claim that can be verified by the induction principle), so $A$ should be started before $B$ if

$$\max_{H\leq I}\left\{\frac{\sum_{i=1}^{H}\beta^{(a_1+\ldots+a_{i-1})}w_{ai}}{\sum_{i=1}^{H}\beta^{(a_1+\ldots+a_{i-1})}}\right\} > \max_{J\leq K}\left\{\frac{\sum_{k=1}^{J}\beta^{(b_1+\ldots+b_{k-1})}w_{bk}}{\sum_{k=1}^{J}\beta^{(b_1+\ldots+b_{k-1})}}\right\}$$

Finally, having undertaken the first stage in $A$, and not yet started $B,$ the same logic can be used to determine whether a second stage of $A$ should be undertaken or not, by defining the $I-1$ remaining stages of $A$ as a new project. In this way the expert optimally allocates her time between competing project demands.

# 4. Level Adjustments

Optimization over continuous choice sets is characterized by marginal adjustments, and the net improvement by deviating slightly from an interior optimum is zero. Calculus is used to derive the solution.

## 4.1 Consumption over the lifecycle

A problem facing us all is how to allocate total personal consumption between periods, that is over the life cycle. It quickly becomes evident in early adulthood as young men and women decide how much debt to take on in our early years to finance their schooling, housing loans and other durable goods such as a cars. The basic elements of this problem can be captured in a simple model of individual optimization by a typical household. How should this household plan its consumption and savings, given the number of dependents, the associated educational expenditures,and the a lifecycle income profile?

Consider first a consumer who is allocating his wealth between his working life and his retirement, which for simplicity we model as two periods. We assume that utility received in each period he lives is concave taking the logarithmic form $\log(c_t)$ where $c_t$ denotes consumption in period $t \in \{1,2\}.$ The length of life is unpredictable in this example. With probability $\beta \in (0,1)$ the person lives both periods and then dies, but with probability $(1-\beta)$ the person dies after the first period, and his entire estate is confiscated by the government in the form of death duties. Hence the expect lifetime utility of the consumer is

$$\log(c_1) + \beta \log(c_2)$$

Consumption is financed by wages, denoted $w.$ Unspent income in the working period is invested in a retirement account which attracts an interest rate of $r.$ Consumption in the second period cannot exceed the amount set aside for retirement, after allowing for growth from interest. Thus the budget constraint of the consumer is

$$c_2 \leq (w - c_1)(1 + r)$$

which can be expressed as

$$c_1 + \frac{c_2}{1+r} \leq w$$

The consumer maximizes his expected utility by choosing $(c_1, c_2)$ subject to his budget constraint. It is easy to see that the budget constraint is binding, meaning the consumer will spend all his wealth. For suppose this was not the case. Accordingly we shall now use this result and

Let $(c_1^o, c_2^o)$ denote the solution to this problem. The optimal choices can be found by using calculus. The Lagrangian for the problem is

$$L = \log(c_1) + \beta\log(c_2) + \lambda\left(w - c_1 - \frac{c_2}{1+r}\right)$$

Since the marginal utility of consumption in either period is infinite at zero, and $\log(c)$ is concave, an interior optimum (in which both $c_1^o$ and $c_2^o$ are strictly positive) pertains. The two first order necessary conditions for require

$$1 = \lambda c_1^o$$

for the first period and

$$\beta(1+r) = \lambda c_2^o$$

for the second. We now have three equations to solve (the two first order conditions plus the budget constraint) in three unknown variables, namely $(c_1^o, c_2^o, \lambda)$. Multiplying the budget constraint by $\lambda$ and imposing equality we see that

$$1 + \beta = \lambda w$$

or

$$\lambda = \frac{1+\beta}{w}$$

Substituting the solution for $\lambda$ into the first order conditions for consumption it immediately follows that

$$c_1^o = \frac{w}{1+\beta}$$

and

$$c_2^o = \frac{\beta(1+r)w}{1+\beta}$$

The expected utility the worker receives from choosing consumption optimally is therefore

$$\log(c_1^o) + \beta\log(c_2^o) = \log\left(\frac{w}{1+\beta}\right) + \beta\log\left[\frac{\beta(1+r)w}{1+\beta}\right]$$

Some intuition for understanding this problem comes from taking the quotient of the two first order equations to derive a fundamental relationship that emerges from optimization between preferences and prices. In the log utility case the ratio of quantities consumed in each period varies proportionately with the probability of survival and the interest rate:

$$\frac{c_2^o}{c_1^o} = \beta(1+r)$$

A sense of the quantitative significance of these results comes from assigning numerical values to the parameters $\beta$ and $r$. The shorter the life span, and the lower the rate of return on savings, the less provision is made for the future.

It is easy to extend this two period model with $\log(c)$ utility to any multi-period setting with differentiable concave utility function $u(c)$ satisfying $u'(0) = \infty$. Suppose a consumer with a wealth of $w$ must allocate her consumption over $T$ periods. Let $c_t$ denote her consumption in period $t$, and suppose the $t$ interest rate from date $0$ to date $t$ is $r_t$. For example if the one period interest rate is the constant $r$ then

$$1 + r_t = (1 + r)^t$$

The consumer's budget constraint is then

$$\sum_{t=0}^{T} \frac{c_t}{1 + r_t} \leq w$$

We suppose that consuming $c$ in period $t$ gives her benefits of $\beta^t u(c)$, where $\beta \in (0, 1)$ is a subjective discount factor that reflects her preference for consuming sooner rather than later. She maximizes her lifetime utility

$$\sum_{t=0}^{T} \beta^t u(c_t)$$

subject to her budget constraint.

The Lagrangian for this problem is defined as

$$L = \sum_{t=0}^{T} \beta^t u(c_t) + \lambda \left[ w - \sum_{t=0}^{\infty} \frac{c_t}{1 + r_t} \right]$$

Differentiating with respect to $c_t$ we obtain the first order condition

$$\beta^t u'(c_t^o) = \frac{\lambda}{1 + r_t}$$

which implies that

$$\frac{\beta^{t-s} u'(c_t^o)}{u'(c_s^o)} = \frac{1 + r_s}{1 + r_t}$$

This equation says that the marginal rate of substitution between consuming in one period versus another equals the relative price in the two periods. If the interest rate is constant then

$$\frac{u'(c_t^o)}{u'(c_{t+1}^o)} = \beta(1 + r)$$

It immediately follows that if $\beta < (1 + r)^{-1}$ and $t > s$ then $u'(c_t) > u'(c_s)$, implying $c_s > c_t$. In words, if the utility from future consumption is discounted more than savings grows with interest, then consumption declines throughout the periods.

This solutions to this problem has a closed form for several parametrizations. For example if the consumer has log utility:

$$u(c_t) = \log(c_t)$$

Then

$$u'(c_t) = 1/c_t$$

Substituting this expression into the first order condition we obtain

$$c_t = \frac{(1 + r_t)\beta^t}{\lambda}$$

The life cycle budget constraint implies

$$w = \sum_{t=0}^{T} \frac{c_t}{(1+r_t)} = \sum_{t=0}^{T} \frac{\beta^t}{\lambda}$$

Therefore

$$\lambda = \frac{1}{w} \sum_{s=0}^{T} \beta^s$$

Substituting the solution for $\lambda$ back into the first order condition for consumption we obtain the optimal consumption as a function of the interest rate and wealth.

$$c_t = \frac{w\beta^t(1+r_t)}{\sum_{s=0}^{T} \beta^s}$$

This yields a utility of

$$v_T(w) \equiv \sum_{t=0}^{T} \beta^t \log\left( \frac{w\beta^t(1+r_t)}{\sum_{s=0}^{T} \beta^s} \right)$$

Experiment. Lifecycle consumption. In the U.S. the average household has about two children, and about half of them go to college. Wages rise with work experience, recent experience counting the most, and are quadratic in age, first increasing and then decreasing. At the other end of the life cycle we must decide how much saving to undertake for retirement.

## 4.2 Procurement

The framework we have used to study intertemporal consumption decisions can readily adapted to analyze spacial location problems. Suppose a cooperative food retailer buys its product from various farms scattered throughout the region, and must determine the number, size and location of its suppliers. There are a countable number of potential farm suppliers, which we index by the distance from the central market $t$. The cost of growing output $x$ at location $t$ is the quadratic function

$$\alpha_{0t} + \alpha_{1t}x + \alpha_{2t}x^2$$

where $\mu_{0t} < 0$ represents a fixed cost that must be incurred maintain the farm, $\mu_{1t} > 0$ is a linear coefficient in food output, and $\mu_{2t} < 0$ is the coefficient on the quadratic term representing declining marginal productivity that sets in as less fertile areas of the farm are cultivated. Aside from cultivation costs, there are transportation costs, denoted by $\gamma_t$, and spoilage, which we denote by $\delta_t$. Let $1\{x_t > 0\}$ denote the indicator function for having a farm joint the cooperative or not. The coop procurement manager chooses the sequence $\{x_t\}_{t=0}^{\infty}$ that minimizes its total costs

$$\sum_{t=0}^{\infty}(1\{x_t > 0\}\alpha_0 + \alpha_1 x_t + \alpha_2 x_t^2 + \gamma_t x_t)$$

subject to demand from the marketing department's requirement that quantity $x$ be displayed on the shelves for sale

$$\sum_{t=0}^{\infty}(1-\delta_t)x_t \geqslant x$$

The formulation of this problem is almost identical to the consumer problem we have just studied. Let $1\{x_t > 0\}$ denote the indicator function that takes on a value of unity if $x_t > 0$ and zero if $x_t = 0$. The Lagrangian for the optimization problem is to choose a sequence $\{x_t\}_{t=1}^{\infty}$

$$L = \sum_{t=0}^{\infty}[1\{x_t > 0\}\alpha_0 + \alpha_1 x_t + \alpha_2 x_t^2 + \gamma_t x_t - \lambda_0(x_t - \delta_t x_t - x) + \lambda_t x_t]$$

where $\lambda_0$ is the Lagrange multiplier, or shadow price of relaxing the demand constraint imposed by marketing, and the $\{\lambda_t\}_{t=1}^{\infty}$is the sequence of Lagrange multipliers ensuring non-negativity from each farm supplier.

The first order conditions for this problem are:

$$\alpha_1 + 2\alpha_2 x_t + \gamma_t - \lambda_0(1 - \delta_t) + \lambda_t = 0$$

which generates the interior solution

$$x_t^o = \frac{\lambda_0(1 - \delta_t)}{2\alpha_2} - \gamma_t - \alpha_1$$

Using the fact that it is not optimal to buy more produce than the amount that satisfies the marketing departments requirements we solve for $\lambda_0$ to obtain

$$\sum_{t=0}^{\infty}(1 - \delta_t)\left[\frac{\lambda_0(1 - \delta_t)}{2\alpha_2} - \gamma_t - \alpha_1\right] = x$$

and solve for $x_t$.

In this experiment we consider a manufacturer which must decide to how many suppliers to draw from.

## 4.3 Resource extraction

Depletable resources such as oil and, to a lesser extent, minerals are typically used in production or consumed soon after they have been extracted. They are less costly to store in their natural state than artificially, extraction costs are delayed if the resource is not immediately required, and transportation costs between the extraction site and the final consumer destination market are lower if there is no need to temporarily store quantities of the resource in another location. Consequently firms roughly determine the current consumption of natural resources and the amount preserved for future periods through their mining and oil drilling activities. Suppose there is a fixed stock $q$ of the exhaustible resource that is depleted by extraction and consumption. As a function of its current price in period $t$, which we denote by $p_t$, the quantity demanded is $p_t^{\eta}$, where $\eta$, a negative constant, is the elasticity of demand. Hence revenue in period $t$ is $p_t^{1+\eta}$ and the overall resource constraint facing the firm is captured by the inequality

$$\sum_{t=0}^{\infty} p_t^{\eta} \leq q$$

For simplicity we shall initially assume there are no extraction costs of the resource. Thus the value of the firm is the net present value of revenue

$$\sum_{t=0}^{\infty}(1+r_t)p_t^{1+\eta}$$

where $r_t$ is the interest on a $t$ period bond bought at date $t = 0$ when the firm makes its drilling plans. We also require that the amount extract each period is nonnegative or that $q_t \geq 0$ for all periods $t$. Thus the owner of the resource chooses a sequence of quantities $\{q_t\}_{t=1}^{\infty}$ to maximize the Lagrangian

$$L = \sum_{t=0}^{\infty}(1+r_t)p_t^{1+\eta} + \lambda_0\left(q - \sum_{t=0}^{\infty} p_t^{\eta}\right) + \sum_{t=1}^{\infty} \lambda_t p_t$$

where $\lambda$ is the Lagrange multiplier on the overall resource constraint, and the $\lambda_t$ sequence are multipliers on the non-negativity constraints for prices and quantities. The first order condition is

$$(1+\eta)(1+r_t)p_t^{\eta} + \lambda_0 \eta p_t^{\eta-1} + \lambda_t = 0$$

Since the marginal revenue at infinitesimal In every period $s$ of the complementary slackness constraints require

$$(1+\eta)(1+r_t)p_t^{\eta} + \lambda_0 \eta p_t^{\eta-1} = 0$$

for each $t$. To interpret this formula, first note that $p(q_t) + p'(q_t)q_t$ the term in inside the brackets on the left side is the marginal revenue from selling resource in period $t$ denominated in current value. The $\beta^t$ term is a discount factor that gives the present value of the unit. On the right side is $\lambda$, which may be interpreted as the shadow value of a new discovery. If $\lambda_t > 0$, then $q_t = 0$ and the marginal revenue of the first sale is less than $\lambda$, the shadow price of extraction. But if $\lambda_t = 0$, then the present value of marginal revenue to the monopolist in period $t$ is equated with the present value of mining one less unit out in another period where extraction is positive.

The resource extraction problem also supports a closed from solution for several parameterizations. Initially assume the inverse demand function takes the form

$$p(q_t) = \gamma q_t^{\delta}$$

which is to say that

$$q_t = \gamma^{\frac{1}{\delta}} p_t^{\frac{1}{\delta}}$$

This is called a constant elasticity demand of curve, because the percentage change in quantity demanded in response to a one percent change in price is the constant

$$\frac{p_t}{q_t}\frac{\partial q_t}{\partial p_t} = \frac{1}{\delta}\gamma^{\frac{1}{\delta}}$$

Substituting $\gamma q_t^{\delta}$, and its derivative, $\gamma\delta q_t^{\delta-1}$ into the first order condition, we obtain

$$\beta^t\left[\gamma q_t^{\delta} + \gamma\delta q_t^{\delta-1}\right] + \lambda_t = \lambda$$

Collecting terms and using the fact that $\lambda_t \geq 0$ yields

$$\beta^t \gamma(1+\delta) q_t^{\delta} \leq \lambda$$

The resource is extracted in one period it must be extracted in every period. Therefore

$$\beta^t \gamma(1+\delta) q_t^{\delta} = \lambda$$

As in the consumption saving examples we can obtain a closed form solution to this problem by first solving for $\lambda$, and then substituting the formula derived back into the first order condition to obtain the optimal $q_t^0$ choices form the first order condition above

$$q_t = \left[ \frac{\lambda}{\beta^t \gamma(1+\delta)} \right]^{1/\delta}$$

$$\lambda = \sum_{t=1}^{\infty} (\bar{q} - q_t)$$

Experiment. In this experiment we include a linear cost of extraction

## 4.4 Internal growth

The dynamic optimization problems we have analyzed above exhibit a linear technology that links constrains current consumption relative to future consumption. In the extraction case the trade of is one for one, and in the consumptions savings model the trade off is determined by the interest rate. In this next application the trade off is nonlinear: in this case the greater the provision made for future production through capital accumulation by a firm, the less productive is an extra unit of investment. We consider an infinitely lived firm that expands by drawing upon its internal resources rather than borrowing funds at a fixed interest rate. Some explanation is necessary to motive why the real investment polices of firms might be related to their liquidity provision. After all when there are well functioning financial markets and little disagreement about the market prospects, then households, firms and grow by seeking support from venture capitalists, shareholders and issuing other forms of debt instruments.

One explanation for this premise is technological. We might suppose that raw materials for investment (say a new variety of bird seed) can only be siphoned off from output, which would otherwise be sold. Another more common explanation for internal growth is explored more fully in later chapters is based on how well markets function when there is incomplete information. Entrepreneurs seeking to carve out niches and create new markets find it more difficult to borrow against their future visions than firms making investments with known returns. On the one hand they have to convince their potential creditors that their ideas are well founded to show that they are realistic and not fraudulent. On the other hand if large sums of credit are required to fund a successful enterprise if even larger profits are anticipated in the future. Unless the entrepreneur has some other way of protecting the rent from his idea the more he shares it with others he exposes it, and hence reduces its value. For these reasons most firms fund their growth internally.

We consider the life cycle of product of a firm which produces revenue from its plant of size $k_t$ at period $t$. Given capital of $k_t$ in period $t$, it produces material of $k_t^\gamma$ which can be retailed as a consumption good, denoted by $x_t$, or transformed into capital for next period $t+1$. The internal financing constraint is then

$$k_t^\gamma = x_t + k_{t+1}$$

The period $t$ current revenue from selling the output $x_t$ is $\log(x_t)$. Given an initial start up capital of $k_0$ the firm chooses plant size $\{k_t\}_{t=1}^T$, or equivalently retail output $\{x_t\}_{t=1}^T$ to maximize its present value

$$\sum_{t=0}^T (1+r)^{-t} \log(x_t)$$

subject to the internal financing constraint.

To solve this problem we substitute for $x_t$ in the firm's objective function using the internal financing constraint to obtain

$$\sum_{t=0}^T \beta^t \log(k_t^\gamma - k_{t+1})$$

In current value terms for period $t$, the marginal loss from an extra investment unit from foregone sales is $(k_t^\gamma - k_{t+1})^{-1}$, while the marginal value from converting the investment unit to sales next period is the product of marginal revenue from sales next period $(1+r)^{-1}(k_{t+1}^\gamma - k_{t+2})^{-1}$ and the increase in production attributed to the extra unit of investment $\gamma k_{t+1}^{\gamma-1}$. The first order condition for an interior solution equates these two values:

$$(k_t^\gamma - k_{t+1})\gamma\beta k_{t+1}^{\gamma-1} = (k_{t+1}^\gamma - k_{t+2})$$

This second order difference equation can be simplified with a change in variables. Defining $K_{t+1} \equiv k_{t+1}/k_t^\gamma$ as the ratio of new capital on current output the first order condition can be rewrtten as:

$$\gamma\beta(1-K_t) = (1-K_{t+1})K_t$$

A solution to this first order difference equation is:

$$K_t = \gamma\beta \frac{1-(\gamma\beta)^{T-t+1}}{1-(\gamma\beta)^{T-t+2}}$$

which can be confirmed by substituting the proposed solution for $(1-K_t)$ and $K_t$ into the difference equation to obtain:

$$\gamma\beta(1-K_t) = \gamma\beta\frac{1-\gamma\beta}{1-(\gamma\beta)^{T-t+2}} = (1-K_{t+1})K_t$$

Morevoer, the firm should not undertake any investment in the last period because the marginal revenue from sales is positive at every level of production, whereas there is no scrap value from retaining capital beyond period $T$. Therefore $k_{T+1} = K_{T+1} = 0$. Since the solution to the difference equation also satisfies the boundary condition that

$K_{T+1} = 0$, we conclude by an induction that for all $T < \infty$, the uniquely defined optimal path of capital accumulation is:

$$k_{t+1} = \gamma\beta \frac{1-(\gamma\beta)^{T-t}}{1-(\gamma\beta)^{T-t+1}} k_t^{\gamma}$$

# 5. Recursive Methods

Many of the problems in this chapter have the same underlying structure. A single player sequentially pursues some objectives that are defined by a criterion function. The game proceeds in discrete time periods. Each period the state of play is given by the value taken by a finite dimensional vector of state variables. This value also determines the choice set for the player in that period. After the player moves, a transition probability, or a mapping, then stochastically, or deterministically, generates the state for the next period. Payoffs accruing to the player accumulate throughout the duration of the game, measuring his performance. In the next two sections we explore this underlying recursive structure with an eye to uncovering some general principles that can be applied to a broad class of dynamic optimization problems.

## 5.1 Notation

We label the periods by $t \in \{1,\ldots T\}$, , where $T$ might be finite, or possibly infinite. When we say $T$ is infinite, we mean that $T$ is a random variable with a distribution that might depend on the state of play, with a probability distribution determines the final period. The state of play at each time $t$ during the game is denoted by $s_t$, which is an element of the time invariant set $S$, where $S$ might be a finite set, a simplex, or a Euclidean space. We denote the choice set at $t$ as $C(s_t)$. This notation means that the options available to the decision maker at $t$ vary from period to period, but only through the realized state $s_t$. At that state of play in the game, the decision maker chooses some $c_t \in C(s_t)$.

The transition from one period to the next determines how the value of the state changes. The transition probability is a conditional probability distribution which generates next period's state variables $s_{t+1}$ conditional on the state variables, and the choice of the player in period $t$. That is $s_{t+1}$, is a random (or deterministic) variable generated by the distribution conditional on $(s_t, c_t)$. We denote this conditional distribution (or law of motion equation) by $F(s_{t+1}|s_t, c_t)$.

The choices and the probability transtions fully determine how the state variables evolve throughout the game. All that remains to fully state this optimization problem is to define the objective of the decision maker. We assume that the decision maker's utility is additively separable over periods and takes the form

$$\sum_{t=0}^{T} \beta^t u(s_t, c_t)$$

We refer to $u(s_t, c_t)$ as the period utility function, again invariant over time, and $\beta$ as teh discount factor. In settings where there is uncertainty we assume that the player maximizes his expected utility throughout the game, a hypothesis we examine more closely in the next chapter. This implies that if the decision maker were to use the same decision rule each period $\tau \geq t,$ say $c(s_\tau),$ his expected utility for the game's remaining duration would be found by taking the conditional expectation

$$E\left[\sum_{\tau=t}^{T} \beta^\tau u(c(s_\tau), s_t) | s_t\right]$$

A typical (additive) term of the summand is computed by substituting the decision rule into the successive probabilty transitions and evaluating the multiple integral

$$E[\beta^\tau u(c(s_\tau), s_t)|s_t] = \beta^\tau \int \ldots \int u(c(s_\tau), s_t) dF[(s_\tau)|s_{\tau-1}, c(s_{\tau-1})] \ldots dF[(s_{t+1})|s_t, c(s_t)]$$

To summarize, a Markov decision problem is defined by the horizon $T,$ state space $S$, the choice set $C(s)$, the transition probability $F(s'|s,c)$, the periodic utility function $u(c,s)$, and the discount factor $\beta.$ To acquire a working familiarity with this notation and the results which follow, we reconsider (discrete time versions of) several applications studied in the previous sections. They all have one state variable.

In the second retirement problem we analyzed in Section 3 income from work, denoted $y_t,$ depends on age alone, but retirement income depends on the age at which a person quits the workforce. Following our notational convention we denote the state variable, age at retirement, by $s_t$ where $s_t = 0$ if the person is still working. At the end of every period during his working life, the person chooses whether to retire next period or not. Let $c_t = 1$ if the person quits at the end of period $t,$ with $c_t = 0$ otherwise. Having quit the workforce, there are no more decisions to make. Therefore the choice set $C(s_t)$ is defined by $C(0) = \{0,1\}$ and $C(s_t) = \{0\}$ for all strictly positive values of $s_t$. The transition for $s_t$ is deterministic, with $s_{t+1} = (1 - c_t)s_t + c_t(t+1).$ Whether the person receives a wage or a pension depends on whether he has already retired or not, so the periodic utility in this problem is defined as

$$u(s_t, c_t) = 1\{s_t \neq 0\}b(s_t) + 1\{s_t = 0\}y_t$$

where $b(s_t)$ denotes the pension received by someone who retires at age $s_t$.

The age of the current machine, now denoted by $s_t,$ is the state variable in the renewal problem. If the machine is older than $\rho$ periods, then the firm can replace it by setting $c_t = 1,$ or retain it by setting $c_t = 0.$ During the installation phase the firm doesn't have a choices. Thus $C(s_t) = \{0,1\}$ for $s_t > \rho$ and $C(s_t) = \{0\}$ for $s_t \leq \rho.$ The transition law for the state variable is dictated by the replacement times. Specifically $s_{t+1} = 1 + (1 - c_t)s_t.$ In this discrete version of the model, we let $\delta$ measure machine efficiency retention, so current utility to the firm is

$$u(s_t, c_t) = 1\{s_t > \rho\}(1 - c_t)\delta^{s_t} + c_t(\gamma - \gamma_0)$$

Unspent lifetime wealth is the state variable in the lifecycle consumption problem, so we denote it by $s_t$. Consumption in period $t$ is a poisitve real number bounded above by this amount, so $C(s_t) = [0, s_t].$ The transition law for $s_t$ recursively define this state variable as $s_t = (1+r)(s_{t-1} - c_{t-1}),$ with initial condition $s_t = w.$ If periodic utility is logarithmic then $u(s_t, c_t) = \log(c_t)$.

Similarly, the resource stock at each point in time is the state variable in the extraction problem, and the quantity extracted is bounded in the same way as in the lifetime consumption problem. In this example

$$s_t = q - \sum_{\tau=0}^{t-1} q_\tau = s_{t-1} - q_{t-1}$$

with $s_0 = q$ and $C(s_t) = [0, s_t]$. Setting $c_t = q_t$ the period utility functiion can be expressed as $u(s_t, c_t) = p(c_t)c_t$.

The state variable for the internal growth problem is the current capital stock, so $s_t = k_t$, while the choice variable is next period's capital stock, or $c_t = k_{t+1}$ with $C(s_t) = [0, s_t^\gamma]$, implying the transition law is $s_{t+1} = c_t$. Finally the periodic utility is $u(s_t, c_t) = \log(s_t^\gamma - c_t)$.

## 5.2 Approximating the solution

We begin our general discussion with the case in which there are a finite number of choice each period, coming from the same set $X$. We also assume that the state space $S$ is a finite set. Each period the subject recieves a current utility $u(x, s)$ and there is a probability distribution that governs the transition from one period to the next, depending on the current state and the action taken. We assume that current utlilty is discounted by a factor $\beta \in (0, 1)$. It now follows that the problem has a finite maximium since

$$\sum_{t=0}^{\infty} \beta^t u(x_t, s_t) \leq \sum_{t=0}^{\infty} \beta^t \max_{(x,s)\in X\times S} u(x, s) = (1-\beta)^{-1} u_{\max} < \infty$$

where

$$u_{\max} = \max_{(x,s)\in X\times S} u(x, s) < \infty$$

Let $v(s_0)$ denote the value of the problem when the initial state is $s_0$. Note that

$$v_T(s_0) + \beta^T (1-\beta)^{-1} u_{\max} \geq v_\infty(s_0) \geq v_T(s_0)$$

Since $\beta^T(1-\beta)^{-1}u_{\max}$ converges to zero as $T$ diverges it follows that

$$\lim_{T\to\infty} v_T(s_0) = v_\infty(s_0)$$

Denote by $\|v_i\|$ the maximum operator over the $K$ elements of the vector $v_i = (v_{i1}, \ldots, v_{iK})$, defined as:

$$\|v_i\| \equiv \max_{j\in\{1,\ldots,K\}} |v_{ij}|$$

For each state $j \in \{1, \ldots, K\}$ We define the mapping $g_j(v_i)$, and the value $x_{ij}$ respectively as

$$g_j(v_i) \equiv \max_{x\in X} \left\{ u_j(x) + \beta \sum_{k=1}^{J} p_{ij}(x) v_{ik} \right\} \equiv u_j(x_{ij}) + \beta \sum_{k=1}^{K} p_{ij}(x_{ij}) v_{ik}$$

Thus $x_{ij}$ solves the maximization problem when the current state is $j$ and the continuation value is $v_i = (v_{i1}, \dots, v_{iK})$. Letting $g(v_i)$ denote the $K$ dimensional vector of values $g(v_i) = (g_1(v_i), \dots g_K(v_i))$, it immediately follows from the recursive representation of $v_\infty$ we established above that $v_\infty = g(v_\infty)$. Also let $h$ denote the state that maximizes the absolute difference between $g_j(v_1)$ and $g_j(v_2)$ over all the states $j \in \{1, \dots, J\}$. That is:

$$h = \arg\max_{j \in \{1,\dots,J\}} |g_j(v_1) - g_j(v_2)|$$

Finally we let $x_h$ denote the optimal choice for the problem that yields the most value at state $h$. In terms of our notation, $x_h = x_{1h}$ if $g_h(v_1) \geq g_h(v_2)$, otherwise $x_h = x_{2h}$. We are now ready to derive the central result of this section:

$$\begin{aligned}
\|g(v_1) - g(v_2)\| &= |g_h(v_1) - g_h(v_2)| \\
&\leq \left| u_h(x_h) + \beta \sum_{k=1}^{K} p_{hk}(x_h) v_{1k} - u_h(x_h) - \beta \sum_{k=1}^{K} p_{hk}(x_h) v_{2k} \right| \\
&= \beta \left| \sum_{k=1}^{K} p_{hk}(x_h)[v_{1k} - v_{2k}] \right| \\
&\leq \beta \|v_1 - v_2\| \sum_{k=1}^{K} p_{hk}(x_h) \\
&= \beta \|v_1 - v_2\|
\end{aligned}$$

The top line follows from the definitions of the norm $\|\cdot\|$ and state $h$, because the latter maximizes the absolute difference between $g_j(v_1)$ and $g_j(v_1)$ over $j \in \{1, \dots, K\}$. The second line uses the definitions of $x_{1h}$, $x_{2h}$ and $x_h$, exploiting the fact that

$$u_h(x_h) + \beta \sum_{k=1}^{K} p_{hk}(x_h) v_{ik} \leq g_h(v_i)$$

which is met with equality if $x_{ij} = x_h$ for $i \in \{1, 2\}$. Cancelling the terms $u_h(x_h)$ yields the third line. From the $K$ inequalities:

$$v_{1k} - v_{2k} \leq \|v_1 - v_2\|$$

implied by the norm $\|\cdot\|$ we obtain the fourth line, while the bottom line is a consequence of the $K$ probabilities summing to one. To summarize:

$$\|g(v_1) - g(v_2)\| \leq \beta \|v_1 - v_2\|$$

Using this inequality, and recalling that $v_\infty = g(v_\infty)$, the triangle inequality implies:

$$\|v_i - v_\infty\| \leq \|v_i - g(v_i)\| + \|g(v_i) - g(v_\infty)\| \leq \|v_i - g(v_i)\| + \beta \|v_i - v_\infty\|$$

for any $v_i = (v_{i1}, \dots, v_{iK})$. Rearranging the the expresions on either end, the main result of this subsection is thus obtained:

$$\|v_i - v_\infty\| \leq (1 - \beta)^{-1} \|v_i - g(v_i)\|$$

In words, the maximal difference between the solution $v_\infty$ and the approximation $v_i$ over the $K$ states is bounded above by an expression that depends on the approximation alone. This upper bound is found by iterating the approximation once, to obtain $g(v_i)$ from $v_i$, calculating the maximal update across the states, which is defined as $\|v_i - g(v_i)\|$, and then treating the maximal update as a perpetuity discounted at rate $\beta$.

As a technical remark we note that the derivation of the main result can be easily extended to larger state spaces by changing the maximum operator to any sup norm, and summation to integration (either using first principles from the definition of Lebesque measure, or by simply replacing sums with integrals over conditional probability distributions).

This result can guide the search for the solution to relatively complicated dynamic optimzation problems. Taking any initial guess $v_i$, one checks its proximity to the unkown $v_\infty$, and as necessary, improves upon the guess by using one of the althogthms described below.

The Approximation Theorem demonstrates the sequence of successve approximation using backwards induction steps is Cauchy. Indeed the approximation error is bounded from above by the product of a known constant and the distance between the approximation and a mapping found by performing one backwards induction step on the approximation. This theorem is of fundamental importance because the investigator is freed from the burden of justifying how he arrived at the approximation.

Dynamic programming problems become intractable for problems with more than a handful of state variables. Intuitively, the nodes on a decision tree increase more than proportinately with the number of states, so when there are a large number of states this renders much less useful standard procedures for solving the dynamic optimization problems. Formal methods that exhaustively update each value of the state are bogged down by the curse of dimensionality. The impact of the approximation theorem is that ad hoc procedures, specifically tailored to the problem's specifications, and experimentation, are not handicapped by the inherent difficulties of properly documenting them. This is because the integrity of the discovery process is essentially immaterial, the quality of the approximation is judged without reference to it.

Note that we are not arguing against the use of algorithms, so much as a recommending their flexible implemention, supplemented with incentivized human intelligence. Rather than exclusively deploying artificial intelligence to solve a dynamic optimization problem, we propose using experimental subjects to sequentially solve it as well as they can, by simply playing it out several times for a reward. By analogy, the top ranked chess players are competitive with the best chess software programs, notwithstanding the strong commercial incentives to develop the latter and very little to support the former. In both chess and dynamic optimization, there is an easily measured objective way of determining superior performance.

## 5.3 Algorithms

We now review two algorithms for solving infinite horizon stationary dynamic problems that satisfy the conditions of the approximation theorem. Value function iteration updates the value function using backwards induction methods. Policy function iteration updates the value function by applying the updated decision rule to the utliity function directly. Both methods exploit the contraction mapping property that since convergence is global, arbitrarily close approximations to the value function can be found by starting at any initial function in teh appropriate space of value functions.

To implement value function iteration, we start with any bounded initial function in the same space as the utility function, denoted by $w_1(s)$. For example we might start with $w_1(s) = 0$, or the value function for the last period of a finite horizon problem with $w_1(s) = v_1(s)$ where:

$$v_1(s) = E\left[\max_c u(c,s)\right]$$

or the maximium value attained from treating the one period problem as an annuity received each period, which is $(1-\beta)^{-1}v_1(s)$. If the utility function $u(c,s)$ is positive for all $(c,s)$, then the value function $v(s)$ exceeds the first two choices of the initial function, but not the third. That is $0 < v_1(s) < v(s)$, but except in trivial cases where there is no investment, no ranking characterizes $(1-\beta)^{-1}v_1(s)$ and $v(s)$. The $(i+1)^{th}$ iterate, denoted $w_{i+1}(s)$, is successively defined by:

$$w_{i+1}(s) \equiv \max_c \left\{u(c,s) + \int w_i(s)dF(s|c,s)\right\}$$

for each $s \in S$. Appealing to the principle of backwards induction, if $w_1(s) = v_1(s)$, then $w_i(s) \equiv v_i(s)$ is the value function for a finite horizon problem $T = i$ periods long. Therefore starting with $v_1(s)$ has the added advantage of showing how quickly the analogous finite horizon problems converge to $v(s)$. In many applications, however, there is less concern with such theoretical niceties, than with achieving convergence in as few steps as possible. In that case we should naturally try a conjecture an initial function that is closer to $v(s)$, such as $(1-\beta)^{-1}v_1(s)$.

Policy function iteration is based on updating the decision rule. Beginning with any decision rule, denoted $c_1(s)$, we form the lifetime value attained using this rule, which is:

$$r_1(s_0) = E\left[\sum_{t=0}^{\infty}\beta^t u(c_1(s_t),s_t)|s_0\right]$$

In practice a large finite value for is substituted for $T = \infty$. The next step is analogous to a backwards recursion. We now update the rule from $c_1(s)$ to $c_2(s)$ by choosing $c_2(s)$ to maximize

$$\max_c \left\{u(c,s) + \int r_1(s')dF(s'|c,s)\right\}$$

for each $s \in S$. Iteration is defined by substituting $c_i(s)$ for $c_1(s)$, and $r_i(s_0)$ for $r_1(s_0)$ in the expressions above to obtain the $(i+1)^{th}$ iterate in the maximization step. Because $c(s)$ is optimal for the decision maker, but $c_i(s)$ is suboptimal, $r_i(s) \leq v(s)$. Furthermore one can prove that convergence is uniform and monotonic with $r_i(s) \leq r_{i+1}(s)$ for all $i \in \{1,2,\dots\}$

Comparing value function iteration with policy function iteration, the key difference is that while both procedures exploit the backwards induction operator, policy function iteration has an additional step of updating the new rule throughout the remainder of the game. Consequently policy function is more involved than value function iteration. But depending on the initial choices of $w_1(s)$ and $c_1(s)$, policy iteration converges more quickly than value function iteration. For example supposing $c_1(s)$ solves $v_i(s)$ for some $i \in \{1,2,\ldots\}$, then one can show that $v_{i+j}(s) \le r_{j+1}(s) \le v(s)$ for all $j \in \{0,1,\ldots\}$.

# 6. Applications

In the three applications below we apply the tools devloped in the previous section.

## 6.1 Inventory control

In the market for perishable goods and personal services, balancing demand flow against supply orders and deliveries affects the profitability of the enterprize. We begin with a simple problem of restocking items when demand is perfectly predictable. Suppose demand for teh item is constant each period, but the store faces nonlinear reordering costs. If it orders $c_t \in \{1,2,\ldots\}$ in period $t$, it pays a fixed cost of $\alpha_0$ for the order plus unit costs of $\alpha_1$. Scaling the quantity units by the amount demanded each period, and denominating unit values by the item's sale price, the net profits in period $t$ are:

$$u_t = 1 - 1\{x_t > 0\}\alpha_0 - \alpha_1 x_t$$

The store maximizes it value by miniminizing the present value of total supply costs subject to the constraint that it demand is met each period. Let $s_t$ denote store inventory at period $t$. It follows the law of motion:

$$s_{t+1} = s_t - 1 + c_{t+1}$$

Thus the store minimizes

$$\sum_{t=0}^{\infty} \beta^t [1\{x_t > 0\}\alpha_0 + \alpha_1 x_t]$$

subject to the constraint that $s_t \ge 0$.

These assumptions ensure that the optimal policy is to order the same quantity of the item at evenly spaced intervals. Intuitively, the problem facing the firm only depends on the current inventory, not the inventory policy that has been pursued in the past. It is also easy to see that the firm whould not order anything until its inventores are exhaisted. These observations considerably simplify the task of deriving the optimal inventory policy. Instead of choosing $c_t$ each period as a function of $s_t$, the firm simply orders an optimally determined $s$ every $s$ periods to minimize:

$$\sum_{t=0}^{\infty} \beta^{st}(\alpha_0 + \alpha_1 s) = (1 - \beta^s)^{-1}(\alpha_0 + \alpha_1 s)$$

The first order condition for this problem is

$$\alpha_1(1 - \beta^s)^{-1} + (\alpha_0 + \alpha_1 s)(1 - \beta^s)^{-2}\beta^s \log\beta = 0$$

Simplifying we obtain

$$(1-\beta^s)\alpha_1 + (\alpha_0 + \alpha_1 s)\beta^s \log\beta = 0$$

the for demanders in the face of uncertain demand, demand flow technology $d_t$ is demand each period with $d_t \leq s_t$ which is current shelf inventory. Thus stocks depreciate at rate $\eta$ implying teh law of motion for food inventory is

$$s_{t+1} = (1-\eta)(s_t - d_t) + x_{t+1}$$

The grocery sequentially chooses $x_t$ as a function of $s_t$ to maximize the expected present value of sales less costs,

$$E\left[\sum_{t=0}^{\infty} \beta^t (pd_t - c_t)\right]$$

## 6.2 Scheduling production runs

On time delivery versus finished product inventory, downtime due to retooling different product lines, reputation for prompt service. Suppose there is a backlog of orders for two types of production $(x_{0t}, x_{1t})$. There is an order flow for each type of good, which must be processed in the order its was received. Orders processed early incur no penalty but carry an inventory cost. Orders processed late incur a cost that increases with tardiness.

Denote by $(d_{0t}, d_{1t})$ the new orders demanded in period $t$ and $(s_{0t}, s_{1t})$ the stock of outstanding orders. Managers decide how many of the new orders to accept. Let $(g_{0t}, g_{1t})$ denote the number of new orders management accepts in period $t,$ where $g_{it} \leq d_{it}$ for each $t \in \{1,2,\ldots\}$ and $i \in \{0,1\}.$ Demand is generated by a probability distribution given by $F(d_{it}).$ Production in period $t$ is $(y_{1t}, y_{2t}).$ In period $t$ the plant incurs costs $f(s_{0t}, s_{1t})$ from departures from just-in-time delivery and also setup costs of $k$ if it shifts from one incurred Every time we switch from one type to the other the plant incurs a setup cost in terms of lost time. A setup reduces the amount of production by $\alpha$ as a fraction of total available run time. The state variables for this problem at time $t$ are $(s_{0t}, s_{1t}, h_t)$ where $h_t$ indicates the status of the factory floor:

$$h_t = \begin{cases} 0 \text{ if the factory is set up for producing } x_0 \text{ at time } t \\ 1 \text{ if the factory is set up for producing } x_1 \text{ at time } t \end{cases}$$

Production at time $t$ is therefore

$$x_{0t} = h_t d_t (1-\alpha_0) + (1-h_t)(1-d_t)$$

$$x_{1t} = h_t (1-d_t) + (1-h_t) d_t (1-\alpha_1)$$

while the demand flow is

$$s_{0t} = s_{0,t-1} + d_t - x_{0t}$$

and the current revenue is, where $x_{0t} \leq d_{0t} + s_{0t},$

$$p_{0t}x_{0t} - [\gamma 1\{s_{0t} > 0\} + \delta 1\{s_{0t} < 0\}] s_{0t}^2$$

If occurs then the plant can process one unit of output Longer runs increase productions reduce the backlog but also incIn teh means for production runs

$$E\left\{\sum_{t=0}^{\infty}\beta^{t}[p_{0t}x_{0t}-[\gamma 1\{s_{0t}>0\}+\delta 1\{s_{0t}<0\}]s_{0t}^{2}]\right\}$$

## 6.3 Child rearing versus career pursuit

In our last dynamic optimization problem we analyze a problem of how to allocate time between projects that are simultaneously demand attention. Fertility and female labor supply The careers of females are complicated by their plans for families. Leaving the workforce, even temporarily, depletes a worker's human capital, shifting to the left the wage distribution she draws from upon her return to work. That is why career concerns and family planning are intertwined. Gender differences in labor supply Women make up about 46% of the workforce, males 53%. Nevertheless certain occupations are very "gender biased": For example cashiers (75% female),Textile sewing machine operators (80% female), Engineers (80% males) and Corporation managers (90% male). These choices may explain Wage differences between men and women. When they enter the workforce female wages are roughly the same as male wages, that is controlling for education, race and ethnicity. But they soon fall behind. By age 50 females are earning about two thirds of what males earn. Trends in fertility. Over the last century, the birth rate has halved. On average U.S women now bare just under 2 children in their lifetime. Incidence of childlessness in the U.S. has more than doubled in the last 20 years to about 20 percent. One reason for the trends As the price of energy declined, market and home production became mechanized. This in turn closed the wage gap between men and women (since women are physically weaker than men). Encouraged women to invest more in market skills and less in domestic skills. Made children relatively more expensive, which led to declining fertility and home childcare. Made manual labor cheaper, leading to a more educated workforce and later entry. Increased the standard of living, inducing men to retire earlier. Another reason for the trend Advances in contraceptive technology are also partly responsible: Fewer births and hence less interruptions, and greater planning about birth timing, increased female labor force participation. Anticipating having fewer children, women now invest more in market based skills. Both explanations are consistent with the fact that the college attendance and graduation rates by females has increased faster than males, and that this further increases their participation and commitment to the work force.

Interruptions in labor force experience. I have already remarked that job turnover dampens the wage experience trajectory. However withdrawing from labor force for several years hurts wages more than a sequence of bad job matches! Furthermore if a person anticipates that she will withdraw from the labor force, she will: Initially search less for a good job and Invest less in human capital on the job. The first factor reduces her starting wage; the second increases her starting wage, but reduces its growth rate. Therefore even a woman who later decides not to have kids, experiences lower wage growth. The next experiment in this section focuses on the effects of labor force interruptions on wages from having children. As before there is a lifetime budget constraint. Children demand household time and goods to raise. Each child yields his/her parents a lifetime upwards shift in utility. Having children earlier starts the benefits and the costs sooner. Wages depend on experience: In reality wages are not fixed but depend on experience in the workforce. Suppose for example, we allow wages to a function of past experience, following the process

How then would you choose life cycle labor supply and plan your family? A dynamic optimization problem. The final application we consider in this chapter returns to the human capital problem, the balancing of pursuing a career and raising a family. We suppose that experience on the job raises future wages, but that the benefits of this experience depreciates over time. Let $I$ denote the amount of time spent at work, and benefits from a larger family where

$$k_t = 1 + \delta k_{t-1} - \sum_{s=1}^{\infty} b_t$$

Suppose a person has a utility function over his consumption and labor supply that takes the form

$$\sum_{t=1}^{\infty} \beta^t [\alpha b_t - \lambda(c_t) + w(k_t)]$$

where $t$ is the period or year, $\beta$ is a subjective discount factor, $c_t$ is consumption in period $t$, $l_t$ is labor supply in period $t$, $b_t$ indicates a birth at that time, $c_t$ is child minding time. As in our previous examples of life cycle behavior we assume there is a lifetime budget constraint on the household that takes the form

$$\sum_{t=1}^{\infty} \beta^t [\alpha b_t - \lambda(c_t) + w(k_t)]$$

where $w_t$ is the current wage rate, $e_t$ is lifetime expenditure per child, and $r_t$ is the interest rate. As in a previous example, we allow wages to a function of past experience, following the process

$$\sum_{t=1}^{\infty} \beta^t [\alpha b_t - \lambda(c_t) + w(k_t)]$$

Several studies (and anecdotal observation) show that young children require less childcare than older ones:

# 7. Summary

This chapter is an overview of dynamic optimization problems, using problems that many individuals personally confront. Roughly speaking a human being passes through five phases between birth and death, namely childcare ( for about 5 years by parents, relatives, friends and preschool), schooling (for about 10 to 20 years, at elementary and high schools, followed for many by further study at university), working in labor force (with intermittent interruptions by females to give birth to and nurture children, which may last from about 5 years to for between over 50 years), and retirement. This chapter has taken examples from each of these phases.

in which the choice sets could be discrete or continuous. It develops techniques for solving them and presents several examples that demonstrate how pervasive they are in professional and domestic matters. We have shown that dynamic optimization problems share many features in common with a much broader class of optimization problems, and in many cases can be solved using similar tools, such as the first order condition, augmented by Lagrangian and Kuhn Tucker multipliers, in the case of constrained optimization. In this spirit we provided closed form solutions for many of the applications we presented. In this application of allocating resources over time, we consider consumption choices over the lifecycle. Nevertheless an important property of many dynamic optimization problem is the additive separability in the objective function across successive periods, a key simplification that allows the analyst to use recursive methods for obtaining numerical solutions even when a closed form solution does not exist. We explained two numerical algorithms for solving such problems, and gave an approximation theorem for testing how close the algorithms, and also experimenting human subjects, come to achieving the optimal choices. Finally we argued that whether humans or computers are cheaper to solve dynamic optimization problems that have no closed form is an open question that depends on a variety of factors that we discuss.

This leads into a discussion of how the behavior of experimental subjects can be tested for opitmality even when the solution is unknown, and describe how experimental subjects may assist in solving dynamic optimization problems through repeated trials.Through a sequence of examples we pose the following question. Simply stated, how complex must a dynamic optimization problem be to lead experimental subjects into serious error? This question is somewhat related to problem solving in artificial intelligence and operations research, where algorithms are rated on how quickly they converge to a fixed point, and how well the fixed point approximates the solution of the given problem. Here we are concerned with how experimental methods compare with, say applied mathematics and computer programmes.