Optimal Control Theory in Economics:
A Short Development of the Hamiltonian Method and
A Discussion of the Ramsey Model

An Honors Thesis (Honors 499)

by

Nathan Berggoetz

Thesis Advisor
Lee C. Spector

Ball State University
Muncie, Indiana

May 2007

Gradation Date: May 2007
Optimal Control Theory in Economics: 
A Short Development of the Hamiltonian Method and 
A Discussion of the Ramsey Model

An Honors Thesis (Honors 499)

by

Nathan Berggoetz

Thesis Advisor
Lee C. Spector

Ball State University
Muncie, Indiana

May 2007

Graduation Date: May 2007
Abstract

This paper explores the economic facets of optimal control theory. The discussion includes the development of the Hamiltonian method, discrete optimal control theory applied to basic consumption analysis, a transition to continuous optimal control problems, and a complete discussion of Dorfman's work with the Ramsey Growth Model.

Acknowledgements

I would like to thank Dr. Lee Spector for his input and guidance during this project. I would also like to thank him for all the knowledge he has given me during my four years at Ball State that has allowed me to understand and write about the complicated topics within this paper.
Think big picture. Consider being able to give every person exactly what they want. The implications of knowing a rule to get every person what they want. The complicated problems of the world could melt away. By no means can this paper tackle the world's problems on an individual or global basis, but there is a curious idea behind these lofty thoughts. Can we think of economic problems in a way that generates rules for giving people what they want? Perhaps a noble goal would be to give everyone the most enjoyment possible, given we could effectively measure enjoyment. We seek to explore this idea from a strict economic and mathematical point of view. The first goal is to develop the mathematics behind optimal control theory. Optimal control theory will serve as the basis to arrive at economic rules for reaching desired goals, such as giving people the most enjoyment possible. As we proceed through the mathematical material, we will accompany each step with an economic example for clarity and applicability. The final step will be to present Dorfman's landmark article where he explores Ramsey's model to maximize utility or enjoyment.

The hope of this paper is to present a sufficient base for the reader to investigate more complicated economic maximization problems. The first step is to consider what we mean by maximizing a function. There are two basic forms of maximization. The first, a typical calculus problem, finds where a function reaches its maximum. The second, the focus of this paper, maximizes a function over an entire interval.

In order to understand the fundamental difference between maximizing a function over an interval of time and maximizing a function locally, we examine each from a graphical point of view. First, consider the classical calculus problem of finding the maximum of a function, \( f(t) \), at a point in time, \( t \). A graphical representation of \( f \) is below in figure 1. The key in this situation
Berggoetz 4

is that \( f(t) \) is only a function of \( t \) and has no constraints except being defined on a close interval, which is a trivial fact in this case. Calculus tells us that \( f(t) \), since it only depends on \( t \), will be maximized when the derivative of \( f(t) \) with respect to \( t \) is zero. This is represented graphically by locating where \( f(t) \) has a horizontal tangent line (see figure 1 at \( t^* \)). There are also second order conditions that must hold to verify that a function reaches a maximum for a certain \( x \). These will be discussed later during our look at Dorfman’s work with the Ramsey Model.

To analyze the idea of maximizing a function over time, consider the idea of having several functions, each with the same initial and terminal conditions. This means that we have several functions that describe different paths between two points (see figure 2). When we say a function \( f \) is maximized over an interval, we have implicitly said that if every function value for every \( t \) in a given interval were summed, this sum would be the largest possible value. In our example we have several functions from which to choose. We then “maximize” by choosing the function that has the most area under its curve. In figure 2 this appears to be line B.

**A Quick Example of Optimal Control Theory in Economics**

Now let us examine this idea within the context of an economic example. Let \( \psi \) be the flow of consumption, \( L \) is the population, and \( U(\psi/L) \) the per capita utility from consumption. Our desire is to maximize utility from consumption. Therefore, we wish to maximize
where \( L(t) = L_0 e^{nt} \) is exponential population growth. The utility function is multiplied by the population to account for wanting to maximize aggregate utility, not the utility of one person. The factor \( e^{-\theta t} \) represents the time preference of each consumption choice. The integral is used because if we want each value of \( U \) to be as large as possible, we want the sum of these \( U \) values to be as large as possible, implying the use of an integral.

The function we are looking for is called the control function. In this example \( \psi \) is the control variable because the flow of consumption, i.e. the rate at which a person consumes, is a factor chosen at each point in time. We want to choose \( \psi \) so that (1) is maximized. However, in maximizing (1), the economy faces a constraint. Let \( X(t) \) be the aggregate capital stock. We call \( X \) the state variable because it shows that the “state” of consumption at each point in time is based on the remaining capital stock. The key here is that we can deduce how \( X \) changes intuitively and use this fact to find \( \psi \). If \( \Phi(X, L) \) is the production function in this example, then \( X \) changes by the difference between production and flow of consumption, i.e. \( \dot{X} = \Phi(X, L) - \psi \). This is the economies constraint. Thus, the problem is to

\[
\max \int_0^\infty U(\psi(t), L(t))L(t)e^{-\theta t} \, dt \quad \text{subject to} \quad \dot{X} = \Phi(X, L) - \psi .
\]

This is the well known Ramsey Model (Weitzman).

We can now see that the difference between maximizing at a point and over time is the result. Maximizing at a point results in, as the name suggests, a point. Where as maximizing over time results in the control function. In our example, this optimal control function outlines from time 0 until the terminal condition what the level of consumption needs to be at every point.
in time. The goal here is to maximize (1) by choosing an optimal level of consumption for each time $t$. As a result, we have an optimal control problem. This paper will examine the Ramsey Model again, although in a slightly different form and in more detail, when reviewing Dorfman’s article on optimal control, but first, we need to develop the definitions and tools necessary to work with optimal control problems in a discrete case.

**Discrete Optimal Control Theory**

Our discussion of optimal control begins with a basic example from consumption analysis. This example will give economic context to the development of optimal control theory, as well as help the reader understand the reasoning behind optimal control theory. Suppose there is an individual that lives for three periods. Assume that he knows his income for each period and starts with assets from saving of $A_k$. We assume income is known for each period because this model does not include any labor leisure decisions. His goal is to maximize his utility over his three period lifetime which is dependent on his consumption each period. His utility function is $U(C) = C^k$, where $C$ is consumption and $k = 0, 1, 2$. Note that no discounting factor is included because we assume for simplicity that consumption this period yields the same satisfaction as consumption next period. A more complicated model would include discounting. His budget constraint is that he can consume, at most, his income from the current period and savings from previous periods and that any unconsumed income is put into his assets from savings. Thus, we know how his assets from saving, $A$, are determined, namely,

$$A_{k+1} = Y_k - C_k + A_k.$$  

Here, $Y$ is income and savings do not earn interest during his lifetime. Hence, his goal is to maximize his total utility over the three periods subject to how his assets from saving are determined or to
maximize $\sum_{k=0}^{2} C_k^\alpha$, subject to $A_{k+1} = Y_k - C_k + A_k$, $A_0 = 0$, $A_3 = a^*$, and $k = 0, 1, 2$.  \hfill (2)

In the initial period our individual has zero assets, $A_0 = 0$, and in the terminal period wishes to leave, as a bequest, an asset amount $a^*$ after period three, or $A_3 = a^*$. Again, the goal for (2) is to find a rule, in the form of a function, that specifies what each period’s consumption should be so that the assets from savings start at zero and end at $a^*$, all while maximizing the utility function $U(C) = C_k^\alpha$.

Below, functions (3) and (4) represent a general formulation of a discrete optimal control problem. In the example (2), (3) would be $\sum_{k=0}^{2} C_k^\alpha$ and (4) would be a manipulated form of $A_{k+1} = Y_k - C_k + A_k$, or $A_{k+1} - A_k = Y_k - C_k$.

$$\text{Maximize } V(x_k, u_k) = \sum_{k=0}^{n} f(x_k, u_k), \hfill (3)$$
$$\text{subject to } x_{k+1} - x_k = g(x_k, u_k). \hfill (4)$$

Our general optimal control problem can be reformulated in terms of the Lagrangean (Fryer 144). We define the Lagrange Function, $L$, to be $L(x_k, u_k) = \sum_{k=0}^{n} f(x_k, u_k) - \sum_{k=0}^{n} \lambda_k (x_{k+1} - x_k - g(x_k, u_k))$ and we seek to find when the following conditions hold in order to maximize (3),

$$\frac{\partial L}{\partial x_k} = 0, \quad \frac{\partial L}{\partial u_k} = 0, \quad \text{and } \frac{\partial L}{\partial \lambda_k} = x_{k+1} - x_k - g(x_k, u_k) = 0 \text{ for all } k \text{ (Fryer 145).} \hfill (5)$$

Note that the result of finding the solutions to the Lagrange equations is not a point but a recursive function involving $x$ and $u$. The conditions in (5) can be thought of intuitively as similar to the first order conditions for finding the maximum a function $r(x)$. These results will become apparent as we develop the Hamiltonian equations.

Following Fryer’s method, if the function $L$ is expanded we see that
\[ L = f_k(x_k, u_k) - \lambda_0 (x_k - x_0 - g_0(x_0, u_0)) + f'_k(x_k, u_k) - \lambda_1 (x_k - x_1 - g_1(x, u)) + \ldots + f_{k-1}(x_{k-1}, u_{k-1}) - \lambda_{k-1}(x_k - x_{k-1} - g_{k-1}(x_{k-1}, u_{k-1})) + f'_k(x_k, u_k) - \lambda_k (x_{k+1} - x_k - g_k(x_k, u_k)) + \ldots + f_n(x_n, u_n) - \lambda_n (x_{n+1} - x_n - g_n(x_n, u_n)). \] (6)

In our consumption example, the expanded Lagrange function yields

\[ L = C_0 - \lambda_0 (A_1 - A_0 - Y_0 - C_0) + C'_1 - \lambda_1 (A_2 - A_1 - Y_1 - C_1) + C'_2 - \lambda_2 (A_3 - A_2 - Y_2 - C_2). \] (7)

Since the goal is to find where the conditions in (5) hold, we can look at (6) row by row. Important to note here is that the \( x_k \) term appears in the \( k-1 \)st and \( k \)th row expansion of \( L \), but the \( u_k \) term only appears in the \( k \)th row. Therefore, for the \( k \)th case

\[ \frac{\partial L}{\partial x_k} = \frac{\partial f'_k}{\partial x_k} + \lambda_k \frac{\partial g_k}{\partial x_k} - \lambda_{k-1} + \lambda_k = 0 \text{ or } \frac{\partial f_k}{\partial x_k} + \lambda_k \frac{\partial g_k}{\partial x_k} = \lambda_{k-1} - \lambda_k \quad \text{and,} \] (8)

\[ \frac{\partial L}{\partial u_k} = \frac{\partial f'_k}{\partial u_k} + \lambda_k \frac{\partial g_k}{\partial u_k} = 0 \quad \text{(Fryer 145).} \] (9)

We can now define a function \( H \), the Hamiltonian, in order to write (8) and (9) more compactly. Let \( H_k = f'_k + \lambda_k g_k \). Taking the partial derivative of \( H \) with respect to \( x \), we see that

\[ \frac{\partial H_k}{\partial x_k} = \frac{\partial f'_k}{\partial x_k} + \lambda_k \frac{\partial g_k}{\partial x_k}, \text{ and from (8), } \frac{\partial H_k}{\partial x_k} = \lambda_{k-1} - \lambda_k. \] Also the partial derivative of \( H \) with respect to \( u \) is

\[ \frac{\partial H_k}{\partial u_k} = \frac{\partial f'_k}{\partial u_k} + \lambda_k \frac{\partial g_k}{\partial u_k}, \text{ which from (9) is equivalent to } \frac{\partial H_k}{\partial u_k} = \frac{\partial L_k}{\partial u_k}. \] Taking the partial derivative of \( H \) with respect to \( \lambda \), \( \frac{\partial H}{\partial \lambda_k} = g(x_k, u_k) = x_{k+1} - x_k \), yields the final condition in (5) (Fryer 145). The system of equations we seek to solve is then:
The first equation in (10) is referred to as the maximum principle. Note that the resultant functions \( x, u, \) and \( \lambda \) may not always give a maximum result for (3) because (10) is only a necessary condition for the maximization of (3) (Fryer 144). More strenuous constraints, such as bounds on \( x \) or \( u \), can be put into any maximization problem. The necessary tools to deal with these constraints will be introduced as needed.

We now have the tools to solve (2). First we need to extract a \( g_k \) function from our constraint. If we subtract \( A_k \) from both sides of the constraint we obtain \( A_{k+1} - A_k = Y_k - C_k \), which is in the same form as (4), our desired result. Hence, \( g_k = Y_k - C_k \). Note that from our general formulation, \( A = x = \) state variable and \( C = u = \) control variable. Although \( Y \) is also a variable in the problem, the assumption that it is known in advance, i.e. a constant for each period, makes it mathematically negligible. Thus, the Hamiltonian for the problem is found to be

\[
H_k = C_k + \lambda_k (Y_k - C_k). \tag{11}
\]

We are now in a position to develop the equations in (10). They are found to be

\[
\begin{align*}
  i) \quad \frac{\partial H_k}{\partial u_k} &= 0 & k=0,\ldots,n \\
  ii) \quad \frac{\partial H_k}{\partial x_k} &= \lambda_{k-1} - \lambda_k \quad k=1,\ldots,n \\
  iii) \quad \frac{\partial H_k}{\partial A_k} &= x_{k+1} - x_k \quad k=0,\ldots,n.
\end{align*}
\]
From the first two equations in (12), we see that $\lambda_0 = \lambda_1 = \lambda_2$ and $C_k = \left(\frac{\lambda_0}{\lambda}\right)^{\frac{1}{\alpha - 1}} = \beta$, $k = 0, 1, 2$.

Note, $\beta$ is a constant because $\alpha$ and $\lambda$ are constant for all periods, implying that consumption is the same for all periods. Expanding the third equation in (12) and substituting $C_k = \beta$ gives

\begin{align*}
\text{i)} & A_b = 0 \\
\text{ii)} & A_i = Y_i - \beta + A_k \\
\text{iii)} & A_k = Y_k - \beta + A_i \\
\text{iv)} & A_i = Y_i - \beta + A_k = a^* ,
\end{align*}

(13)

a system of three equations and three unknowns. The most important variable to solve for is $\beta$ since it determines our consumption for all periods. If we back substitute iv into iii and then iii into ii, we get an equation in terms of just $\beta$, namely $a^* - Y_2 - Y_1 + 2\beta = Y_0 - \beta$. Thus,

$$
\beta = \frac{Y_0 + Y_1 + Y_2 - a^*}{3} .
$$

To find the $A_k$'s, forward substitute ii into iii, and iii into iv. This gives $A_2 = Y_1 - Y_0$ and $A_3 = Y_2 + Y_1 - Y_0 - \beta$. The result for $\beta$ makes economic sense because it states that if we total the income we will receive over the three periods of the model and subtract the amount of assets desired to leave as a bequest, we get the amount of money we have available for consumption. We then divide this amount up over the three periods equally because there is no time preference factor or discount rate in the model. The interesting result here is that since

we are supposed to know our income over the life of the model, we plan consumption accordingly. This idea supports the permanent income hypothesis, which states that an individual's consumption over their lifetime is a fixed constant, while income varies. The result of (2) says that utility is maximized when
consumption is constant over the three periods and income is exogenously determined varying over the three periods. Figure 3 shows the permanent income hypothesis graphically. The graph is split into three regions, representing the three periods of our maximization problem. As the graph of the permanent income hypothesis shows, consumption is constant while income varies. In region 0, we are spending more than we are earning. In region 1, we are earning more than we are spending. In region 2, we are again spending more in consumption than we are earning in income.

**Transitioning to Continuous Time**

There are some issues with using discrete models in optimization problems. Note that when the model was formulated, great care was taken in specifying which periods were involved in the constraint. This was important in order to frame the problem in a way that allowed the Hamiltonian method to work. As a result, we will formulate the remaining models in continuous time to avoid this problem. The advantages of both methods are put succinctly by Bertola, Foellmi, and Zweimuller who state that

The advantage of continuous time formulation is that it frequently yields simple analytic solutions, and it is not necessary to specify whether stocks are measured at the beginning or the end of the period. The advantage of a discrete time model is that empirical aspects of the role of uncertainty are discussed more easily in a discrete time framework (6).

We now turn to continuous optimal control problems. There are several methods to arrive at the system of equations for the continuous case. We have chosen to follow Fryer's method. Other formulations can be found in Shone or Dorfman. The discussion here focuses on Fryer's intuitive presentation of an arrival at the continuous optimal control problem forgoing a mathematically rigorous explanation in the interest of simplicity and clarity. Consider an
example where we want to plan our optimization model over a fixed time interval of length $T$. If we solved the discrete model in the form of (3) and (4) finding all the points $x_k$ and $u_k$ for $k=0,1,...,n+1$, giving an optimal result for (3), we can graph examples of these points as in figures 4 and 5.

![Figure 4](image1.png) ![Figure 5](image2.png)

Between each point is an equal distance, call it $h$, where $h = T/(n+1)$. Allowing the number of periods in our planning horizon to approach infinity, i.e., increasing $n$, while holding $T$ fixed, forces $h$ to approach zero. This implies that more and more points will be on the above graphs and as more points are added they will be closer and closer together until the lines connecting the currently displayed points are actually a series of points themselves. In the limit of $h$ approaching zero, $x_k$ and $u_k$ become continuous functions of time $t$, $x(t)$, $u(t)$. Similar arguments verify that $f(x_k, u_k)$ and $g(x_k, u_k)$ become continuous functions of the form $f(x(t), u(t), t)$ and $g(x(t), u(t), t)$, respectively (Fryer 153). The next step is to multiply the remaining discrete statements by $h$ and manipulate them in such a way that taking the limit as $h$ approaches zero yields a continuous function.

Multiplying (3) and the right hand side of (4) by $h$ yields
Statement (14) is nothing more than a left endpoint approximation of an integral of a continuous function \( f \), and we have already assumed that our discrete \( f \) is approaching a continuous version (Stewart 512). Here we are summing the area of \( n \) rectangles under the function \( f \) with base \( h \) and height \( f(x(t), u(t)) \), where \( x(t) \) and \( u(t) \) are evaluated at \( t = k = 0, 1, \ldots, n \). As \( h \) approaches zero the summation in the limit becomes \( \sum_{k=0}^{n} f(x_k, u_k)h \) and \( (14) \)

\[
\sum_{k=0}^{n} f(x_k, u_k)h \quad \text{and} \quad (14)
\]

\[
x_{k+1} - x_k = g(x_k, u_k)h. \quad (15)
\]

Dividing both sides of (15) by \( h \) yields

\[
\frac{x_{k+1} - x_k}{h} = g(x_k, u_k). \quad (16)
\]

Note that the definition of a derivative \( \frac{dx}{dt} \) is

\[
\frac{dx}{dt} = \lim_{h \to 0} \frac{x(t+h) - x(t)}{h} \quad (Stewart\ 156).
\]

If we know that \( x_k \to x(t) \) from above, then (16) yields \( \frac{x(t+h) - x(t)}{h} \). Hence (16) becomes

\[
\frac{dx}{dt} = g(x, u, t),
\]

where \( \dot{x} \) denotes the derivative of \( x \) with respect to \( t \).

The discrete Hamiltonian, \( H_k = f_k(x_k, u_k) + \lambda_k g_k(x_k, u_k) \), now becomes a continuous function,

\[
H(x, u, t) = f(x, u, t) + \lambda(t) g(x, u, t), \quad (17)
\]

where \( \lambda(t) \) is now a continuous function as well. Converting (10) into a continuous problem yields (10i) to be \( \frac{\partial H}{\partial u(t)} = 0 \). Equation (10ii) is in the same form as we argued for (16) when -1 is
factored out. Thus $-\left(\frac{\lambda_k - \lambda_{k-1}}{h}\right)_{h \to 0} - \frac{d\lambda}{dt} = -\frac{\partial H}{\partial x(t)}$. Finally, (10iii) becomes $\frac{\partial H}{\partial \lambda(t)}$ by the same argument as in (16) (Fryer 155). Hence our new continuous conditions to solve for an optimal time path are

$$
\begin{align*}

&i) \frac{\partial H}{\partial u} = 0 \\
&ii) \frac{\partial H}{\partial x} = -\dot{\lambda} \\
&iii) \frac{\partial H}{\partial \lambda} = x. \\
\end{align*}
$$

Next we reexamine the consumption analysis problem of (2), only in the continuous case. Arguing as we did for the general transition from a discrete problem to a continuous one, (2) becomes

$$\begin{align*}
\text{maximize} \int_{0}^{T} C(t)^\alpha \, dt \quad \text{subject to} \quad &A = Y(t) - C(t), \, A(0) = 0, \, A(T) = a^*. \\
\end{align*}
$$

There are several differences between (2) and (19). Notice that each function is continuous with respect to time $t$, and that we have adjusted the number of intervals to a ambiguous time frame of length $T$. This time interval can be finite, for example where $T = 3$ as in (2), or $T$ can be infinite.

The Hamiltonian for (19) is

$$H(C(t), \, A(t), \, Y(t), \, t) = C(t)^\alpha - \lambda(t)(Y(t) - C(t)), \quad (20)$$

and, as before, we know that $C(t) = u(t) = \text{control variable}$ and $A(t) = x(t) = \text{state variable}$ in our general formulation. Thus the system (18) becomes
\[ i) \frac{\partial H}{\partial C(t)} = \alpha C(t)^{\alpha - 1} - \lambda(t) = 0 \Rightarrow \alpha C(t)^{\alpha - 1} = \lambda(t) \]

\[ ii) \frac{\partial H}{\partial A(t)} = 0 = -\lambda \Rightarrow \lambda(t) = K_0, \text{ where } K_0 \text{ is a constant} \quad (21) \]

\[ iii) \frac{\partial H}{\partial \lambda(t)} = Y(t) - C(t) = \dot{\lambda}. \]

Note that \( K_0 \) is a constant of integration from solving equation (21ii). From (21) we can immediately solve for \( C(t) \). Substituting (21ii) into (21i) and solving for \( C(t) \) yields,

\[ \alpha C(t)^{\alpha - 1} = K_0 \Rightarrow C(t) = \left( \frac{K_0}{\alpha} \right)^{\frac{1}{\alpha - 1}} = \beta. \quad (22) \]

Thus, our maximization consumption rule, \( C(t) \), in (22) gives the same result as our discrete version where consumption is set at a constant level, \( \beta \), for all periods. A constant consumption level is expected as this model has not put any preference on consuming now or a period in the future. We have been purposefully ambiguous about the form of the function for income, \( Y(t) \), because of the difficulty \( Y(t) \) could present in solving (21iii). Most likely, the desired form of \( Y(t) \) would exhibit behavior similar to the income function in Figure 3. This shape of a function would undoubtedly cause problems in solving (21iii). As this paper is not on solving differential equations we leave this analysis out. Now that we have the necessary tools to deal with continuous optimal control problems we revisit Ramsey’s Model through Dorfman’s article *An Economic Interpretation of Optimal Control Theory*.

**Interpreting the Ramsey Model with Dorfman**

The method here will be to present Dorfman’s article, following his approaches and presenting his insights. Note that, where necessary, additional explanations are given to foster understanding. Also added to Dorfman’s work is a proof from Chiang to show that the Hamiltonian equations as in (18) give a maximum result for the Ramsey Model.
Suppose that we want to know how to allocate the capital in a one-sector economy over time, with the aim of getting the most out of the capital. One perspective as to how to get the most from the capital is to maximize the utility we gain from the consumption that the capital affords us. Ramsey formed this problem by asking, “How much of its income should a nation save?” (543). Dorfman formulates the same problem by asking what the “socially optimal path of capital accumulation for a one-sector economy” should be (824). Our first assumption is that the population grows exponentially at a fixed rate, $n$. Therefore, the population $N(t)$ at anytime $t$ is,

$$N(t) = N_0 e^{nt}.$$  \hspace{1cm} (23)

We assume that $N_0$ is one, measured in hundreds of millions of people. Consumption is represented by $C(t)$, and $u(c)$ is a measure of the utility from per capita consumption, i.e., $c = C/N$. We assume that $u(c)$ has diminishing marginal utility such that $u'(c) > 0$ and $u''(c) < 0$ (Dorfman 824).

The total utility from all persons is then $e^{nt}u(c)$. If we assume there is a time preference for consumption now versus later in life, then denote $\rho$ as the measure of this time preference (Dorfman 824). We treat $\rho$ as a discount factor where the present value of 1 unit of consumption $t$ years from now would be $e^{-\rho t}$. Thus total discounted utility from all persons is $e^{(n-\rho)t}u(c)$. As stated before our objective could be to maximize the above discounted utility. Hence we proceed as Dorfman does, and seek to

$$\text{maximize } W = \int_0^T e^{(n-\rho)t}u(c)dt.$$  \hspace{1cm} (24)

The planning period interval $T$ can be finite or infinite. We will discuss the implications of each later. The next step is to consider what factors go into consumption.
Consumption is controlled by output and in turn output is controlled by capital (Dorfman 824). Let \( K(t) \) be the amount of capital at time \( t \). We again consider capital on a per capita basis where \( k = K/N \). Assume that the production function for the model has constant returns to scale and diminishing marginal product. Hence our production function is of the form

\[
Y(t) = N(t) f(k(t)) \quad \text{where} \quad f'(k) > 0 \quad \text{and} \quad f''(k) < 0. \quad (25)
\]

Note that the change in capital over time is equal to net investment, or output minus consumption minus deteriorated capital (Dorfman 824). Hence,

\[
\dot{K} = Y - Nc - \delta K, \quad (26)
\]

where \( \delta \) is the deterioration rate of capital. The next step is to manipulate (26) to eliminate \( K \) and \( N \), putting everything in terms of \( k, c, \) and \( n \). This will allow us to find a constraint to go along with (24).

First substitute our definition of \( Y \) into (26). Then notice that \( K = Nk \) and substitute for \( K \) in (26). Thus,

\[
\dot{K} = Nf(k) - Nc - \delta Nk \\
= N( f(k) - c - \delta k). \quad (27)
\]

The last step is to consider that \( k = \frac{d}{dt} \frac{K}{N} \), and applying the quotient rule for derivatives yields,

\[
\dot{k} = \frac{\dot{K}N - K \dot{N}}{N^2} = \frac{\dot{K}}{N} - \frac{K \dot{N}}{N^2} \\
= k \left( \frac{\dot{K}}{\dot{N}} - \frac{N}{K} \right) = k \left( \frac{\dot{K}}{Nk} - \frac{n}{Nk} \right). \quad (28)
\]

Substituting (27) into (28) gives,
\[
\dot{k} = k \left( \frac{N(f(k) - c - \delta k)}{Nk} - n \right)
= k \left( \frac{f(k) - c - \delta k}{k} - n \right) = f(k) - c - (n + \delta)k.
\]

(29)

We can now state a complete optimal control problem in the form of (24) subject to (29) (Dorfman 825). Note that the goal is to maximize the utility from consumption over the period T subject to how the capital labor ratio changes. The result will be a rule for what consumption, the control variable, should be at any point in time \( t \), and a rule for what the resulting capital labor ratio, the state variable, should be at any time \( t \).

The next step is to generate the Hamiltonian from the optimal control problem in (24) and (29). The Hamiltonian is

\[ H = e^{(n-\rho)t}u(c) - \lambda(t)(f(k) - c - (n + \delta)k). \]

(30)

The system we then seek to solve is

\[
\begin{align*}
\dot{\lambda} &= e^{(n-\rho)t}u'(c) - \lambda = 0 \\
\dot{f}'(k) &= n + 8 - \frac{\partial H}{\partial c} = \lambda(f'(k) - (n + \delta)) \Rightarrow f'(k) = n + \delta - \frac{\lambda}{\dot{\lambda}} \quad (31) \\
\dot{k} &= \frac{\partial H}{\partial \lambda} = f(k) - c - (n + \delta)k.
\end{align*}
\]

The \( t \)'s are suppressed in (31) for convenience. Dorfman gives an economic meaning to the first equation in (31). Namely, he defines \( \lambda(t) \) as the value of a unit of capital at time \( t \). Thus \( \lambda(t) \) is equal to the marginal utility from consumption adjusted for population growth and social time preference (Dorfman 820). Dorfman develops his reasoning for defining \( \lambda(t) \) as such in his development of the Hamiltonian equations. We have omitted this complete discussion as it is hindering instead of illuminating to those new to optimal control and accept the definition.
Before we analyze this problem anymore, we need to verify the conditions in (31) in fact yield a maximum result. As mentioned earlier the Hamiltonian method is only a necessary condition for a maximum result. We follow Chiang's rather simple approach to check that solving (31) gives a maximum result. In calculus, if we want to find the maximum of a function \( f(x) \), we find where \( f'(x) = 0 \), then check if \( f''(x) < 0 \) for all \( x \). This would imply that \( f(x) \) is always concave, and hence, \( f'(x) = 0 \) must reveal the \( x \) that maximizes \( f(x) \) (Stewart 277). Thus, Chiang argues that since \( \frac{\partial^2 H}{\partial c^2} = e^{(n-p)t}u''(c) < 0 \) for all \( c \) by our assumption that there is diminishing marginal utility, we must have a maximum (256).

Since solving (31) will yield a maximum, we proceed with our analysis. First, we need to eliminate the term \( \frac{\dot{\lambda}}{\lambda} \) from (31ii). We know what the function \( \lambda \) is from the first expression (31i).

Therefore, taking the derivative of \( \lambda \) from (31i) with respect to \( t \) yields

\[
\dot{\lambda} \frac{\lambda}{\lambda} = \frac{d\lambda}{dt} = (n - \rho)e^{(n-p)t}u'(c) + e^{(n-p)t}u''(c) \frac{dc}{dt}.
\] (32)

Now dividing equation (32) by the function for \( \lambda \) from (31i) yields

\[
\frac{\dot{\lambda}}{\lambda} = \frac{(n - \rho)e^{(n-p)t}u'(c)}{e^{(n-p)t}u'(c)} + \frac{e^{(n-p)t}u''(c) \frac{dc}{dt}}{e^{(n-p)t}u'(c)}
\] (33)

Substituting (33) into (31ii) eliminates \( \frac{\dot{\lambda}}{\lambda} \), giving that

\[
f''(k) = n + \delta - n + \rho - \frac{u''(c) \frac{dc}{dt}}{u'(c) \frac{dt}{dt}}
\] (34)
Dorfman interprets the last equation to be such that

Along the optimum path of accumulation the marginal contribution of a unit of capital to output during any short interval of time must be just sufficient to cover the three components of the social cost of processing that unit of capital, namely, the social rate of time-preference, the rate of physical deterioration of capital, and the additional psychic cost of saving a unit at the beginning of the interval rather than at the end (825).

The psychic cost of saving is a tricky idea. Essentially, the last term in (34) is the percentage rate of change over time of the psychic cost of saving. We know that \( u'(c) \) is the marginal utility of consumption and in equilibrium must be equal to the marginal cost. The marginal cost is measured as an individual’s gratification, which as Hoxie notes, is a psychic phenomenon immeasurable in real terms (212). Thus \( u'(c) \) is the disgratification caused from processing a unit of capital instead of saving and \( \frac{u^*(c)}{u'(c)} \frac{dc}{dt} \) must be the percentage change over time.

In a sense we have a “rule” in (34) for maximizing our utility, but there is a catch. When solving differential equations the initial conditions, in this case \( k(0) \) and \( c(0) \), can greatly effect the behavior of the system. Thus we would have to pick the point \( (k(0), c(0)) \) that would give the correct behavior to the solution path for the system of differential equations. Therefore, we need to generate a system of differential equations in terms of \( k \) and \( c \), the state and control variables, to solve.

Solving (34) for \( \frac{dc}{dt} = c \) yields

\[
\dot{c} = \frac{u'(c)}{u^*(c)} \left( \rho + \delta - f'(k) \right).
\]  

(35)
We now have a system of differential equations in (31) and (35) to solve for the time path of per capita capital and consumption. If we choose a production, \( f(k) \), and a utility, \( u(c) \), function, most likely each would be nonlinear making (31) and (35) nearly impossible to solve. Thus we leave (31) and (35) in general form and analyze the system's behavior using a phase diagram (Dorfman 825). First, we find the steady state for the system which occurs when there are no changes in capital or consumption, i.e., when \( k = c = 0 \). Thus,

\[
\begin{align*}
\dot{k} &= 0 \Rightarrow c = f(k) - (n + \delta)k \\
\dot{c} &= 0 \Rightarrow f'(k) = \rho + \delta \Rightarrow \text{constant.}
\end{align*}
\]

Consider the graph of \( k = 0 \) in the k-c plane in figure 6.

The graph of \( k = 0 \) has this shape because of the constant returns to scale and diminishing marginal productivity assumptions for \( f(k) \). Pick any point A such that A is below \( k = 0 \).
Evaluating $\dot{k}$ at $(k_i, c_i)$ yields the expected, $\dot{k} = 0$. If we evaluate $\dot{k}$ at A where $c_2 < c_1$, then $k(k_i, c_i) > k(k_1, c_1) = 0$. Thus for any point A below $k = 0$, $\dot{k} > 0$. Thus capital k is increasing and we draw the arrow in figure 6 to the east. If follows by similar argument that for a point B above $\dot{k} = 0$ that $\dot{k} < 0$ implying that k is decreasing. Hence, we draw the arrow to the west in figure 6. Next, consider the graph of $\dot{c} = 0$ in figure 7. Pick a point A such that A is less than $c = 0$. Evaluating $\dot{c}$ at $(k_i, c_i)$ gives $\dot{c} = 0$. For a point A, where $k_2 < k_1$, we can see that $\dot{c}(k_2, c_1) > \dot{c}(k_1, c_1) = 0$. Thus, c is increasing for points less than $\dot{c} = 0$. We draw an upward arrow as a result in figure 7. By similar argument we can see that a point B greater than $\dot{c} = 0$ must exhibit the opposite characteristics. Hence, we draw the downward arrow in figure 7.

Superimposing these two graphs reveals the behavior of points in each of the four quadrants created by the two graphs. Figure 8 shows the complete representation of these directional behaviors. This sort of behavior is classified as a saddle path fixed point because solution paths generally start to approach the fixed point for any $(k_0, c_0)$ but eventually diverge away forming the “saddles” seen in figure 9 depicting several stylized solution paths. Notice that in figure 9 each curve exhibits the behavior outlined by the arrows in
figure 8 for each quadrant. The goal now is to pick an initial starting point \((k_0, c_0)\) such that we end up with a desired level of capital and consumption (Dorfman 826). This choice will become obvious as we examine different \((k_0, c_0)\).

Presumably, we would start with some low level of capital, say \(k_0\), as we are considering the early developmental stages of a country. Now we must choose a \(c_0\). Most choices of \(c_0\) will give a stylized result similar to path A or C in figure 9. In A, as Dorfman notes, the country would adopt a policy such that initially consumption increases until we reach the critical \(k\) where \(c = 0\), at which point consumption would start decreasing. During all this, capital would continue to grow until consumption reaches zero. Most likely, this country would be draining all consumption into capital to reach some goal for capital accumulation. The result of this is the starvation of the population (Dorfman 826). This is not a viable path for capital and consumption. On a solution path such as C, a country would adopt an overindulgent mentality. Capital increases initially until the critical value where \(k = 0\) is reached, at which point all capital would be drained off into consumption. This is the overuse of resources in order to consume a great deal (Dorfman 826). Again, this is not a viable path for capital and consumption. In either case there is a point where either consumption is zero and capital accumulation must stop because the population is dead, or capital is zero and consumption must stop since there are no more inputs to create consumable goods (Dorfman 826).

The question that remains is if there is a choice of \((k_0, c_0)\) that leads to the fixed point and not the exhaustion of capital or consumption. Unfortunately, as with many problems in differential equations finding the actual answer to this question is difficult or impossible. In this case we must settle with stating that we know a solution path, call it \(\{s\}\), starting at \((k_0, c_0)\) and
approaching the fixed point where \( \dot{k} = \dot{c} = 0 \) must exist. We know this path exists because every saddle path fixed point has a curve that separates regions of different behavior called the separatrix. The separatrix path approaches the fixed point in the limit (King 227). In our case the separatrix denotes the line that separates the regions of behavior where consumption is driven to zero and where capital is driven to zero. This path, by definition, must exist and approach the fixed point where \( \dot{k} = \dot{c} = 0 \) in the limit. The separatrix represents our desired solution and must be the optimal path since all other paths lead to the exhaustion of one of the variables. We denote a stylized version of this path by B in figure 9.

Returning to the original problem of this discussion, have we gained any information about what the best path for capital accumulation should be to maximize utility from consumption? We have a basic rule in the form of (34) stating that the marginal product of capital must be covered by three components of social cost. The problem then becomes choosing an initial level of consumption that will lead to a solution path that approaches the fixed point. We know such a choice exists. Considering our original problem asked us to find a function to determine our consumption over time, and now we are only left with finding the correct starting level of consumption, this is considerable progress. Also note that picking linear production and utility functions would yield a solvable system of differential equations. The only problem is economically justifying that these functions should be linear. We are also now able to ask questions about whether or not all countries will end up in the same place or growing at the same rate at the fixed point, given their economies operate as outlined by the Ramsey Model. There is certainly the possibility for this to occur, but it would require that all parameters in the model, population growth, capital depreciation factor, time preference, etc, be the same.
across countries. Even given that all parameters are the same, we are still left with the question of where to start, and as we have seen this is not an easy question to answer.

If we did end up on the stable path in our model, we would approach the fixed point. Dorfman notes that once the fixed point is reached consumption, capital, and the population must all grow at the same rate (826). Some may argue that this is not a desirable result since it affords us no long term growth in per capita capital and as a result per capita income. This assertion contradicts the original formulation of our model though. The goal throughout has been to maximize utility. Thus if we are on the optimal path, regardless of no per capita growth, we have afforded all included in the model the maximum utility at all points in time. Thus, there is no need to allow for long term growth as we have not valued growth as the long term goal of the economy. We have instead focused on maximizing utility.

Perhaps a more interesting question to ask is what happens if we consider a country starting with a high level of capital, say $k_1$ as in figure 9. This starting point is perfectly feasible considering a country could experience a positive economic shock from perhaps a technological boom sending capital and as a result consumption very high. Policy makers would then seek to find a way to deal with these high levels and bring the country towards stability. The case could be made that we are currently in this state considering the advances made from carbon based energy. It has afforded countless amounts of consumption. The question remains, are we consuming the correct amounts to reach the stable fixed point. Our model shows us such a path exists since their must be a separatrix on this side of the graph as well.

**Summary**

This paper has attempted to deal with the very complex issue of economic planning, through the use of optimal control theory in maximizing an economic goal. The fundamental
tools developed here where basic discrete optimal control problems, in the form of consumption analysis. We also developed a framework to move from the discrete to continuous case problems to avoid time period issues is setting up the model. This paper also develops the Hamiltonian method of solving optimal control problems. Applying this method to the Ramsey Model shows not only applicability but the clear notion that optimal control is the key element in analyzing economic growth. In fact optimal control has affected economic growth theory enough to change its name from originally being capital theory to its current name growth theory (Dorfman 817). Applying optimal control theory to the Ramsey Model fundamentally transformed the questions we asked about the economy; from what should our consumption be over the life of the model to what level of consumption should we start out at. This change in questioning is substantial in the progress it offers but still leaves the Ramsey Model looking for answers.
Berggoetz 27

Works Cited


Chiang, Alpha C. *Elements of Dynamic Optimization.* Prospect Heights, IL: Waveland P, Inc,
2000. 253-263.


Fryer, Mj, and Jv Greenman. *Optimization Theory Applications in OR and Economics.*

(1905): 210-230.


543-559.

Shone, Ronald. *Economic Dynamics Phase Diagrams and Their Economic Applications.* 2nd ed.
