1994

Permanent income hypothesis and the cost of adjustment

Gerald F. Parise
Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/rtd
Part of the Economic Theory Commons

Recommended Citation

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
Permanent income hypothesis and the cost of adjustment

Parise, Gerald Francis, Ph.D.

Iowa State University, 1994
Permanent income hypothesis

and the cost of adjustment

by

Gerald F. Parise

A Dissertation Submitted to the

Graduate Faculty in Partial Fulfillment of the

Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Department: Economics
Major: Economics

Approved:
Signature was redacted for privacy.

In Charge of Major Work
Signature was redacted for privacy.

For the Major Department
Signature was redacted for privacy.

For the Graduate College

Iowa State University
Ames, Iowa

1994
TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION 1
CHAPTER 2: PERMANENT INCOME HYPOTHESIS 9
CHAPTER 3: DATA 42
CHAPTER 4: COSTLY ADJUSTMENT 73
CHAPTER 5: LEARNING 138
CHAPTER 6: CONCLUSION 205
BIBLIOGRAPHY 215
APPENDIX A 221
APPENDIX B 224
APPENDIX C 227
APPENDIX D 230
ACKNOWLEDGMENTS 232
CHAPTER 1: INTRODUCTION

Consumption, which accounts for more than 60 percent of the U.S. gross national product, is an area of macroeconomics that has always received much attention. Understanding this important sector would offer greater insights into the evolution and direction of the economy in general. All this attention, however, has not always translated into greater understanding. Despite the development of numerous models and prodigious study, the field of consumption theory remains very much open.

From the 1950s to today, consumption theory has seemed intent on shaking off the naivety commonly associated with the Keynesian view of consumption. Contrary to the Keynesian view, where consumption is simply a function of current income, theory would be based on an optimization framework where rational agents would attempt to maximize utility. Being rational, agents would formulate a lifetime consumption path subject to the constraint imposed by total lifetime resources. Current income would not be a good predictor of current consumption because such a limited notion of income forms but a small part of total lifetime resources. The agents from a Keynesian world would appear quite primitive compared to the agents in the sophisticated, forward-looking models of today (Hadjimatheou 1987, p. 4).

Two popular attempts at incorporating microfoundations when determining consumption are the life-cycle hypothesis and the permanent income hypothesis (PIH). By placing the agent in an optimizing framework, both models show that consumption is more than simply a function of income. Advantages of these derivations include precise theoretical rationale for determining consumption and expanding the information set upon which consumption may depend. According to Modigliani and Brumberg (1955),

according to this theory there need not be any close and simple relation between consumption in a given short period and income in the same period. The rate of consumption in any given period is a facet of a plan which extends
over the balance of the individual's life, while the income accruing within the period is but one element which contributes to the shaping of such a plan. This lesson seems to have been largely lost in much of the empirically-oriented discussion of recent years, in the course of which an overwhelming stress has been placed on the role of current income . . . almost to the exclusion of any other variable. (pp. 391-392)

This paper is cast entirely in terms of the PIH. This framework is chosen because it models consumption in a very reasonable way. Rather than simply using current income, agents utilize the expected future stream of discounted income and current wealth in determining a consumption path, an idea that appears intuitively plausible. Deviations between current consumption and this notion of permanent income, are attributed to surprises, or a transitory component.

Using this framework and the assumptions of rational expectations, perfect markets, perfect information, fixed real interest rates, and quadratic utility, Hall (1978) derived the important result that consumption follows a random walk. Given the assumptions of the model, this result makes intuitive sense because consumption in any period will reflect all the information to which the agent is privy. The change in consumption must thus occur as a surprise because it represents something that cannot be foreseen. Current income, of any given lag, would thus provide no information in determining current consumption because all relevant information is embodied in the previous period's level of consumption.

This notion of consumption, around which much contemporary work is centered, appears as the antipode of the simple Keynesian consumption function. By using a rational expectations/PIH Hall showed that current consumption depends on nothing beyond its lagged level. Arriving at such a cogent result, however, comes with a very high cost, a cost that must be paid in terms of strong assumptions and diminished flexibility. These assumptions may be considered as being both explicit and implicit. As mentioned, the explicit assumptions are those relating to notions of perfect capital markets, perfect information, fixed interest rates, and quadratic utility. Among the various implicit assumptions used by the rational expectations/PIH are that adjustment to new information and past errors is very swift (Hadjimatheou 1987, p. 8). As will be seen later in this paper, weakening these assumptions
will produces results where information beyond lagged consumption has a significant effect upon current consumption.

Because the assumptions of the PIH are central in determining how consumption evolves, much work has been undertaken to study consumption when these conditions are altered. Analysis of two assumptions, that of perfect capital markets and that of rational expectations, form the basis of this paper. Under the assumption of perfect capital markets are notions of the agent’s ability to borrow as much as is desired, absence of transactions costs and discrimination, equality among lending and borrowing rates, and perfect and free information. Only with these conditions can the agent smooth consumption as predicted by the PIH. What happens when some or all of these conditions fail to hold is discussed in Chapters 3 through 5, where an alternative form of the PIH is considered. The alternative presented in these chapters incorporates the notion of imperfect capital markets by analyzing the costs that these situations may impose upon agents as they attempt to formulate a consumption plan. Formulating the consumption decision in this manner is sufficiently flexible so as to have the PIH as a special case.

The second assumption addressed in this paper is that of rational expectations. Analysis of this assumption is analyzed in Chapter 5, where the informational requirements of the rational expectations hypothesis are discussed and an alternative method by which expectations may be determined is offered. This alternative permits agents to learn about the system as the sample progresses. Expectations generated in this manner are shown to be very flexible because the expectation-generating mechanism itself will be able to change over time, thereby permitting some degree of structural change to influence the determination of consumption.

In analyzing consumption under the conditions just presented, this paper is organized as follows. Chapter 2, entitled “The Permanent Income Hypothesis,” presents a brief review of the PIH and a time-series representation that will prove useful in estimation. Following the presentation of the PIH, a number of papers that test the PIH are reviewed. Most of these papers reveal empirical shortcomings associated with the PIH as it is currently formulated.
Three results become evident from this work. First, is that current consumption depends on more information than simply the previous level of consumption. Second, consumption is excessively sensitive to changes in income, that is, income influences consumption more than is warranted under the PIH. Third, consumption is too smooth under the PIH, implying that the observed variance in consumption is less than that predicted under some specifications of the permanent income hypothesis.

Rejection of the PIH begs the question of what caused such rejection. A section Chapter 2 provides one possible explanation, which centers on the assumption of perfect capital markets. A number of papers are reviewed in addressing this question. The basic conclusion that can be reached from these studies is that the assumption of perfect capital markets is excessively strong and that once these markets are allowed to be imperfect, consumption will depend on more information than simply its lagged level.

In much of the literature, the incorporation of market imperfections has proceeded in two ways. The first explicitly specifies the imperfection that is assumed to be confronted by the consumer, and the second relies on the amorphous notion that one segment of the population is liquidity constrained. Both approaches encounter problems. The former raises the question of how agents can identify all possible manifestations of market imperfections and properly incorporate such notions into the determination of consumption. The latter approach, while ignoring the complications of specifying all imperfections that may affect the agent, must provide an explanation as to how agents are categorized, something which is not done. Additionally, the liquidity-constrained approach makes the implicit assumption that the unconstrained agent is completely unaffected by market imperfections.

With the apparent rejection of the PIH and the inappropriate methods in which market imperfections have been incorporated when determining consumption, an alternative approach to modeling consumption is considered. The foundations for this alternative are established in Chapter 3, which discusses variables that may be used in modeling consumption. We start by offering a reformulation of permanent income that will allow access to more information in estimation. In Chapter 2, permanent income relies upon disposable income and capital income.
Using the assumption of Ricardian equivalence, Chapter 3 shows that permanent income can be expressed as a function of gross labor income, capital income net government bonds, and government expenditures. This alternative view of permanent income is used in all our subsequent work.

Following the reformulation of permanent income, definitions of all the variables used in studying consumption are presented. Various statistical properties such as the order of integration and potential cointegration are also considered, so as to ascertain the proper manner in which these data may be modeled. The last section of Chapter 3 presents basic tests of the PIH using the data described in that chapter. Rejecting the PIH with our data establishes the rationale for the alternative model of consumption formulated in Chapter 4.

Given that consumption does not follow the PIH, as concluded in the review of the literature in Chapter 2 and with tests using the data presented in Chapter 3, an alternative approach to modeling consumption is needed. Chapter 4 provides one possible suggestion by modeling capital market imperfections. The suggested alternative avoids the criticisms associated with previous attempts at incorporating market failures. In particular, Chapter 4 models the costs associated with market imperfections.

The method presented in Chapter 4, which forms the basis of all the work that follows, expresses imperfections encountered by agents in terms of the costs that such phenomena may generate. Two types of costs that may affect an agent's consumption decision are envisioned. The first models the costs associated with being unable to obtain desired levels of consumption because of possible market failures. These costs may arise because of the breaking of habits or legal contracts, or through possible search costs, as the agent attempts to reformulate a consumption plan given the inability of achieving desired levels. Deviations between actual and desired consumption will encourage the agent to eliminate such a gap; however, the act of altering consumption levels initiates a second cost, examples of which include liquidity constraints, transactions costs, or possibly losses incurred when assets are liquidated. Imperfections may thus introduce rigidities into the system that prohibit
instantaneous change. Consumption in this framework will be that level which balances these two costs.

Modeling imperfections in terms of the costs that these phenomena may generate provides the advantage of not specifying a priori the exact nature of the imperfections that may confront the agent. In our complex economy, it may be exceedingly difficult for the agent to be cognizant of all possible imperfections and the proper method of incorporating these effects into the determination of consumption. By focusing on the costs that these imperfections may generate, the informational assumption imposed upon the agent is diminished. Costly adjustment is also more flexible compared to the liquidity constraint approach discussed previously because all notions of market imperfections are permitted to affect the agent, implying that no arbitrary division of the population is required.

To model this idea of costly adjustment, we start by considering a quadratic cost model that may be used in determining the optimal level of consumption. Both the cost associated with deviations between actual and desired consumption and the cost related to altering consumption levels can be utilized in arriving at the optimal consumption path. Desired consumption in this model is assumed to be represented by the level of consumption predicted by the PIH. Modeling costs in this framework will show that the PIH is nested within the costly adjustment model, thereby allowing for a test of the PIH within a more general setting. Additionally, an alternative view of consumption similar to the one found in Cushing (1992), which presents a costly adjustment model, that is shown to be nested within this more general framework. These special cases are important because they allow for tests of the costly adjustment model against two explicit alternatives.

Following the derivation of the model, an auxiliary system is specified so that necessary expectations can be generated and joint estimation performed once formidable cross-equation restrictions are imposed. Estimation suggests that both special cases can be easily rejected, implying the importance of incorporating the two costs discussed above in determining the optimal consumption path. These results, however, rely upon strong assumptions about how the agent formulates expectations. As an alternative, we offer a
second model that imposes a lower level of knowledge upon the agent. In particular, a simplistic rule is utilized in determining the optimal consumption. However, despite the naiveté introduced into this model, the notion of costly adjustment still cannot be rejected. An important implication of this work is the possibility that government policies may influence current consumption levels.

This second model serves as an introduction to the work presented in Chapter 5, where the assumption of rational expectations is explicitly discussed. Estimation in Chapter 4 relies heavily upon this assumption because it allows for the handling of unknown expectations in a tractable and theoretically appealing manner. However, this assumption also imposes strong conditions upon the agent, primary of which is a correct understanding of the complete system. Knowledge of the system is obtained by simply assuming that the agent has been operating in the system indefinitely. Chapter 5 provides an alternative method in which expectations may be formed, a method that allows the agent to learn about the system as the sample progresses. This work may be viewed as an attempt to reduce the amount of information presupposed upon the agent and seems a natural extension of Chapter 4, which acknowledges that the consumer possesses less than perfect knowledge of the economic system.

Implementation of a learning model in Chapter 5 provides an interesting alternative framework in which consumption determination may be modeled. By reducing the amount of a priori knowledge imposed upon the agent, a more flexible environment is obtained. Learning, as implemented in Chapter 5, permits model parameters to alter over time, reflecting any change that may have occurred in the economy over the sample. Possibly important information may be conveyed in these time-varying parameters, both in terms of the coefficients themselves and also with regard to altering expectations. Tests of the general model within this learning framework will explicitly take into account possible changes in the economy and the manner in which expectations are formed.

An important result of Chapter 5 is that a number of coefficients display large variations over the sample, a situation relevant to the results of Chapter 4, which relies heavily
upon the assumption of rational expectations. With the possibility of time-varying parameters, expectations generated under such an assumption may be incorrect. This possibility casts some doubt on the conclusions reached in that chapter, while revealing the importance of permitting the agent to learn about these variations as the sample progresses. Chapter 5 concludes with one possible test that allows us to comment upon the validity of the general model, given a time-varying environment.

The final chapter offers conclusions from all the work presented. Central to this chapter are results and implications from testing the general model. As discussed, acceptance of the general model implies a certain structure with which consumption should be modeled. Estimation results and implications are also reviewed and discussed in the case of the time-varying/learning model.
CHAPTER 2. THE PERMANENT INCOME HYPOTHESIS

Introduction

Over the years, the permanent income hypothesis (PIH) has provided valuable insights into modeling consumption behavior. Originally proposed by Friedman (1957), the PIH offered an appealing alternative to the simplistic Keynesian notion that consumption is some function of current income. Beyond being ad hoc, this Keynesian notion was considered incorrect because it ignored the optimization nature of the consumption decision. Friedman’s insight was that consumption will be determined to reflect the agent’s lifetime expected resources. Current income forms but a small part of this larger quantity, which a truly optimizing individual uses in formulating current consumption. An obvious implication of this hypothesis is that the consumption equation utilizing only current and lagged income as regressors would be misspecified, because these notions only measure the true regressor, permanent income, with error. An explanation is thus offered for the empirical shortcomings, as reviewed by Friedman (1957), associated with the application of the Keynesian consumption function.

Tying consumption to permanent income seems very plausible, because the agent will use notions of expected income in determining current consumption, behavior that appears quite likely, given rationality. Current income is of diminished importance because it does not provide information as to whether some desired consumption path can be sustained into the future. However, this reliance on expected income and wealth rests very strongly upon the assumptions of the model. Consumption levels can be sustained relative to diminished current levels of income only if the agent is freely able to borrow sufficient amounts based upon expected future earnings. The agent’s inability to achieve this, destroys the smoothing aspect
of the PIH because the agent may be forced to consume some proportion of current income. Thus, while Friedman's notion of consumption determination may be considered more sophisticated relative to the Keynesian view, implementation rests heavily upon model assumptions. As will be shown, violation of these assumptions casts even the rational agent back into a Keynesian world.

The assumptions upon which the PIH is constructed seem to be the key in ascertaining consumption determination. True optimizing behavior is only possible given the model assumptions, whereas the ad hoc Keynesian notion may be seen to arise when such assumptions fail to hold. If it is assumed that this optimizing view of consumption determination is correct, the view taken by this paper, tests of the PIH are in essence statements dealing with the appropriateness of the underlying assumptions. This paper accepts the idea that consumption is determined by some optimization procedure because it seems intuitively plausible. What is questioned, however, is whether the agent can carry out the results of this optimization exercise. Formulating and implementing some desired level of consumption are considered to be very different.

In this paper, it is the dichotomy of the formulation and implementation of a consumption plan that is studied. Thus, while we assume that the agent operates in an optimizing environment, implementation of consumption plans based on such maximization may be hampered by reality. In particular, the assumption made by the PIH that is of primary concern is that of perfect capital markets. By being unable to borrow sufficient amounts and by encountering market rigidities that may be caused by imperfect markets, the agent may be unable to achieve desired consumption levels, levels predicted by the PIH.

To best study consumption determination in the PIH framework this chapter is organized as follows. The following section, entitled "The Permanent Income Hypothesis," provides a brief summary of Friedman's original work. Whereas the PIH offers a theoretical rationale for the work to follow, it is the rational expectations/time-series representation of the PIH that is used in studying the properties and validity of the model. Explicit modeling of the
notion of permanent income, based on disposable income and current wealth, is considered in the following section which provides a time-series representation of the PIH.

Testing the validity of the time-series representation of the PIH forms the basis of the subsequent section, entitled "Testing the PIH." Within this section, a number of studies will reveal whether or not the PIH is consistent with economic reality. The papers reviewed study various implications of the PIH in an effort to determine whether the hypothesis is correct or not. Most of these papers suggest that the PIH is not consistent with the data, though it must be emphasized that these results relate to the time-series representation of the hypothesis.

Empirical shortcomings of the theory raise the question "Why?". To address this question, a section entitled "Imperfections" discusses the possibility of market imperfections. Papers studying this question are reviewed and comments are presented dealing with the manner in which researchers have approached this question. These papers reveal that researchers have incorporated market imperfections into the modeling of consumption in two ways. The first method uses an explicit representation of perceived imperfections, while the second assumes that one segment of the economy is liquidity constrained.

The chapter concludes with the section entitled "Criticisms" which critiques the two methods by which market imperfections have been incorporated into the determination of consumption. As noted, both approaches possess shortcomings related to the assumptions imposed upon the agent and the economy. With these shortcomings, an alternative approach is briefly discussed. This alternative forms the basis of Chapters 4 and 5.

The Permanent Income Hypothesis

The development of the PIH began with the classic work of Friedman (1957). In his introduction an espousal is made for a new view of consumption determination. As Friedman states, a revision to the Keynesian notion is needed because
each set of budget studies separately yields a marginal propensity decidedly lower than the average propensity. Finally, the savings ratio in the period after World War II was sharply lower than the ratio that would have been consistent with findings on the relation between income and savings in the interwar period. This experience dramatically underlined the inadequacy on a consumption function or savings solely to current income. (p. 4)

The PIH was formulated as a possible explanation for the state of affairs in which the notions of permanent income and permanent consumption assume special importance. It was Friedman's contention that permanent income, and not current disposable income, that is relevant in determining consumption. A simple regression of consumption on a constant and disposable income must then be handled as an errors-in-variable problem because disposable income measures permanent income, but only with error (Stock 1988). Friedman's model is comprised of three elements:

\[(2.1) \quad y = y_p + y_t,\]
\[(2.2) \quad c = c_p + c_t,\]
\[(2.3) \quad c_p = k(i, w, u)y_p.\]

Equation (2.1) defines income as being composed of permanent income, \(y_p\), and transitory income, \(y_t\), where permanent income is defined as

reflecting the effect of those factors that the unit regards as determining its capital value or wealth: the non human wealth it owns; the personal attributes of the earners in the unit, such as their training, ability, personality; the attributes of the economic activity of the earners, such as the occupation followed, the location of the economic activity, and so on. (Friedman 1957, p. 21)

Transitory income reflects chance occurrences, whereas permanent consumption, \(c_p\), in equation (2.2) was purposely left vague by Friedman ("... best left to be determined by the data themself..." (p. 21). Transitory consumption, \(c_t\), is defined similarly to \(y_t\), as reflecting chance occurrences. Equation (2.3) establishes the relationship between \(c_p\) and \(y_p\), where the term \(k(i, w, u)\) is a function of the interest rate (\(i\)), the ratio of nonhuman wealth to income (\(w\)), and consumer tastes (\(u\)) (p. 26). It is additionally assumed that \(E(y_t, y_p) = E(c_t, c_p) = E(y_t, c_t) = 0\).
Analysis following Friedman's original work has incorporated the notion of rational expectations in defining permanent income. Permanent income in such a model can be characterized as

\[ y_p = \frac{r}{1+r} \left[ w_t + \sum_{s=0}^{\infty} \left( \frac{1}{1+r} \right)^s E_t y_{t+s} \right], \]

where \( w_t \) represents real wealth at the beginning of period t; \( y_t \) is income to be paid at the end of the period; the real interest rate, \( r \), is constant (Flavin 1981, p. 977); and \( E_t \) denotes the expectation conditional on information in period t. In this context, permanent income is interpreted as the constant resource flow that can be sustained for the remainder of the individual's time horizon (p. 977). Speight (1990) terms such a modification as the rational expectations/permanent income hypothesis (REPIH); however, reference will be made simply to the PIH in all the work that follows.

The following section presents a possible implementation of the PIH. The section serves to make clear the derivations and assumptions that comprise the PIH, while the derived time-series representation will provide insights into the modeling of consumption and the basis for which tests of the PIH may be performed.

Time-Series Representation

Assumptions needed in deriving the intended representation are as follows:

1. An infinitely lived representative individual uses a quadratic utility function to quantify the pleasure of consumption. Such a utility function is expressed as:

\[ u(c_t) = u_0 + u_1 c_t - \frac{u_2}{2} c_t^2. \]
The following properties are ascribed to preferences and the utility function noted in the first assumption. These assumptions are from Speight (1990, pp. 86-89).

(a) Independence. Preferences are a function of one's own consumption, i.e., utility is not a function of the consumption of others.

(b) Two-stage budgeting. This assumption allows the agent to partition the total set of possible consumptions into groups, with expenditures being allocated across groups of goods at the first stage and within groups at the second. Justification depends on the notions of the composite commodity theorem and weak intertemporal separability. The composite commodity theorem presumes that "... consumption can be treated as a 'Hicks aggregate,' whereby when the relative prices associated with a set of goods move in parallel, that set of goods can be treated as a single good" (Speight 1990, p. 87). Weak intertemporal separability assumes that we are able to break up an intertemporal utility function into a series of sub-utility functions.

(c) Strong (additive) separability. The utility function is taken to be comprised of sub-utility functions, which can be combined additively.

(3) Single perfect financial asset (Speight 1990, p. 89). This property assumes that there exists a single financial asset, $A$, available in positive and negative amounts measured at the end of each period and receiving interest payments at the beginning of each period. Further, there is a single interest rate, $r$, which applies to both borrowing and lending. The consumer can borrow as much as is desired, because it is assumed that the capital market for this asset is perfect. There are no transactions costs, implying that consumption levels may be changed instantaneously.

Using these three assumptions, a time-series representation of the PIH can be presented. Following Sargent (1989), the consumer is assumed to choose consumption so as to maximize $E_0 \sum B^t u(c_t), 0 < B < 1$ subject to the restriction that wealth evolves as $A_{t+1} = R[A_t + y_t - c_t]$, where $y_t$ is defined as total labor income net of taxes, $A_t$ is defined
as wealth, \( r \) denotes the real (constant) rate of interest, and \( R = 1 + r \). Summing this equation for \( A_{t+i} \) over all time results in the intertemporal budget constraint:

\[
\sum_{i=0}^{\infty} R^{-i} E_t c_{t+i} = A_t + \sum_{i=0}^{\infty} R^{-i} E_t y_{t+i}.
\]

This expression shows that the present discounted value of consumption equals the present discounted value of income and current assets. Assuming a quadratic utility function, that is,

\[
u(c_t) = u_0 + u_1 c_t - \frac{u_2}{2} c^2_t, \quad \text{with } u_0, u_1, u_2 > 0,
\]

and maximizing such a function given the budget constraint, the optimal level of consumption can be found to be

(2.4) \( c_t = \left( \frac{-\alpha}{R - 1} \right) + \left( 1 - \frac{1}{BR^t} \right) \left[ \sum_{i=0}^{\infty} \left( \frac{1}{R} \right)^i E_t y_{t+i} + A_t \right].\)

Where \( E_t \) denotes the expectation operator conditional with respect to information up to period \( t \), \( B \) is a discount factor, and \( \alpha = \frac{u_1 (1 - (BR)^{-1})}{u_2} \). Assuming that \( BR = 1 \), as in Sargent (1989) and Flavin (1981), Equation (2.4) can be expressed as

\[c_t = \left( 1 - \frac{1}{R} \right) \left[ \sum_{i=0}^{\infty} \left( \frac{1}{R} \right)^i E_t y_{t+i} + A_t \right].\]

By defining permanent income as:

(2.5) \( y_{p,t} = \left( \frac{R - 1}{R} \right) \left[ \sum_{i=0}^{\infty} \left( \frac{1}{R} \right)^i E_t y_{t+i} + A_t \right].\)

we can write current consumption as

(2.6) \( c_t = y_{p,t}.\)

In defining permanent income in this manner, \( \frac{R - 1}{R} A_t \) may be considered as a measure of capital income (Campbell and Deaton 1989, p. 358). Permanent income, as expressed by Equation (2.5), may be thought of as representing a sustainable resource flow conditional on
current information (Speight 1990, p. 99). In each period the agent will consume $\frac{R - 1}{R}$ of total discounted expected income and current wealth, which is permanent income.

It should be noted that this definition of permanent income used above is slightly different compared to that used by Flavin (1981). Differences arise from the definition of variables since Flavin assumes that $y_t$ and $c_t$ are measured at the end of the period, while we assume that these quantities are measured at the beginning of the period. Nothing of substance changes by using these definitions because, as can easily be shown, our notion of permanent income used above fulfills the property that $E_t y_{t+1} = y_t$. This is a required property because permanent income can be thought of as the constant resource flow, conditional on expectations in period $t$, which can be sustained for the remainder of the individual’s time horizon (Flavin 1981, p. 977). At period $t$, the agent’s best guess for future levels of permanent income is thus the current period’s level.

Equation (2.6) can now be manipulated to yield a form that has come to assume much importance in modern consumption theory. Differencing Equation (2.6), while adding and subtracting the term $E_t y_t$ shows that

$$\Delta c_t = E_{t-1} \Delta y_t + (E_t - E_{t-1}) y_t. \tag{2.7}$$

To evaluate (2.7) the two components on the right-hand side must be determined. The first term, $E_{t-1} \Delta y_t$, can be determined by using the work of Flavin (1981). In her paper, Flavin shows that $E_t y_{t+1} = (1 + R) y_t - Rc_t$ which, using Equation (2.7), shows that $E_t y_{t+1} = y_t$. Thus, given that the PIH is correct, $E_{t-1} \Delta y_t = 0$. Under the assumption of rational expectations, the second term in Equation (2.7) is a disturbance term that is serially uncorrelated and has a mean zero, implying that the equation can be rewritten as

$$\Delta c_t = \omega_t \quad \text{where} \quad \omega_t = (E_t - E_{t-1}) y_t. \tag{2.8}$$

Equation (2.8) shows that consumption follows a random walk. For such a property to hold, all the assumptions implicit in the PIH must be fulfilled and there must be no transitory component to consumption (Flavin 1981, page 978). Equation (2.8) makes a strong statement
about the evolution of consumption. This representation implies that all information up till period t-1 is embodied in \( c_{t-1} \). Once this variable is taken into account, no other variable dated prior to time \( t \) will have any relevancy in modeling current consumption. This is the so-called orthogonality property. The testing of which is discussed later. We now consider a number of papers that provide tests of the time-series representation of the PIH.

### Testing the PIH

Over the years a myriad of papers have been, and are being, generated which have tested various implications of the PIH. Rather than attempt yet another review of this literature, only a sampling is presented here. Complete reviews of the literature are found in Deaton (1991) and Speight (1990).

Hall (1978) brings together Friedman’s initial work and the notion of rational expectations to form the basis of modern consumption theory. Using the assumptions of quadratic utility, a constant real rate of interest, no transitory consumption, and perfect capital markets, Hall derives his famous result that consumption follows a random walk, which is Equation (2.8).

To test the random walk hypothesis, Hall regressed current nondurable consumption, defined as nondurable goods and services and denoted as \( C_t \), on its own lag, \( c_{t-1} \), and the two lags of disposable income, \( y_{t-1} \) and \( y_{t-2} \). Under the PIH, the coefficients for \( y_{t-1} \) and \( y_{t-2} \) should be jointly insignificant because once lagged consumption is included in the regression, the agent has incorporated all relevant information, implying that \( \Delta C_t \) should be orthogonal to all additional information in period \( (t-1) \). Thus, further lags of consumption, lagged values of income or any other variable known at time \( t-1 \) should enter with an insignificant coefficient. Estimation by Hall fails to reject the null hypothesis that the coefficients for \( y_{t-1} \) and \( y_{t-2} \) are
zero; however, he does find that lagged stock prices appear with a statistically significant coefficients when included in the regression (p. 984).

An alternative test of the PIH is provided by Flavin (1981). Flavin starts with the premise that because future income is not known, consumption plans must be made contingent upon some set of expectations of future income. The PIH discussed above provides one way in which these expectations of future income and current wealth can be combined to provide an estimate of consumption. In particular, Flavin assumes that consumption evolves as

\[ c_t = y_{pt} + u_t, \]

where \( y_{pt} \) denotes permanent income that may, for example, be represented by Equation (2.5), \( c_t \) denotes nondurable goods, and \( u_t \) represents a transitory component of consumption. As mentioned previously, Flavin's definition of permanent income is slightly different. In her paper permanent income is expressed as

\[ y_p = R \left[ w_t + \sum_{s=0}^{\infty} \left( \frac{1}{1 + R} \right)^{s+1} E_t y_{t+s} \right], \]

where \( w_t \) denotes beginning period wealth, \( y_t \) is beginning period disposable income, \( r \) denotes the real interest rate, and \( R = 1 + r \). For consistency in discussing Flavin's paper, her notion of permanent income is used.

Consider now the difference in consumption that Flavin (1981) shows [Equation (8) p. 978]:

\[ c_{t+1} = c_t + r \sum_{i=0}^{\infty} \left( \frac{1}{R} \right)^{i+1} (E_{t+i} - E_t) y_{t+i+1} - Ru_t + u_{t+1}. \]

This equation shows that, given some ARMA specification for disposable personal income, \( y_t \), and given that the PIH is correct, consumption and permanent income will only evolve as a random walk given that \( u_t = 0 \) (p. 978).

Flavin's test of the PIH centers around the term \( r \sum_{i=0}^{\infty} \left( \frac{1}{R} \right)^{i+1} (E_{t+i} - E_t) y_{t+i} \), which represents the innovation in permanent income. Once an ARMA process is specified for \( y_t \), this term may be expressed as
where $\psi$, denotes some function of the ARMA parameters (see p. 987) and $\epsilon$, denotes the disturbance term of the income series. Substituting allows the change in consumption to be expressed as

$$c_t = c_{t-1} + R\Phi \epsilon_t - R u_{t-1} + u_t,$$

which shows that the change in consumption is a function of the innovation in permanent income and transitory consumption.

The basic empirical implication of Flavin’s model is that even though the marginal propensity to consume out of current income is zero, the agent will respond to innovations in current income because these innovations provide new information about future income and thus permanent income. This implication in turn suggests that if changes in current and lagged income are included in Equation (2.9), such variables should be statistically insignificant. Coefficients for these changes in income are measures of the excess sensitivity of consumption to current income; that is, sensitivity in excess of the response attributed to new information about income (p. 990).

In testing the model, Flavin assumes that income, $y_t$, is a trend stationary series and stationarity is achieved by detrending (p. 989). Using these detrended data, along with the change in consumption, a joint system can be specified as

$$y_t = \nu_1 + \rho_1 y_{t-1} + \rho_2 y_{t-2} + \ldots + \rho_p y_{t-p} + \epsilon_{1t},$$

$$\Delta c_t = \nu_2 + k\Phi(y_t - \nu_1 - \rho_1 y_{t-1} - \rho_2 y_{t-2} - \ldots - \rho_p y_{t-p}) + \beta_0 \Delta y_t + \beta_1 \Delta y_{t-1} + \ldots + \beta_{p-1} \Delta y_{t-(p-1)} + \epsilon_{2t},$$

where:

$$\Phi = \left[ \begin{array}{c} 1 \\ 1 - \frac{\rho_1}{1 + r} - \ldots - \frac{\rho_p}{(1 + r)^p} \end{array} \right]$$

$$\epsilon_t = y_t - \nu_1 - \rho_1 y_{t-1} - \ldots - \rho_p y_{t-p}.$$
If the PIH is correct, the change in $c_t$ should only respond to the innovations in the $y_t$ process (surprises), implying that the coefficients $\beta_0, \beta_1, \ldots, \beta_p$ should all be zero. Thus, if the PIH is correct consumption should only react to surprises in the income series. These $\beta_i$'s are the so-called excess sensitivity parameters that are seen in much of the subsequent work.

Flavin's estimation utilized quarterly data and eight lags in modeling detrended income ($p = 8$). Two models were estimated, the unconstrained model, noted above, and the constrained model where $\beta_1 = \beta_2 = \ldots = \beta_p = 0$. Again, under the PIH, the restriction $\beta_1 = \beta_2 = \ldots = \beta_p = 0$ should not be rejected. A likelihood ratio test of this hypothesis provides decisive rejection of the constrained model (p. 999), suggesting that the change in consumption does react to changes in income.

One weakness with Flavin's work, which has been widely identified, is the use of detrended consumption and income series. In particular, Mankiw and Shapiro (1985) show that Flavin's procedure for testing the PIH is biased toward rejection if income has an approximate unit root. An interesting extension to Flavin's work would be the incorporation of difference stationary notions on income and consumption, an extension considered in the following chapter.

These results suggest that the PIH, as currently formulated, is incorrect. Above, it was seen that one implication of the PIH, as formulated by Hall (1978) and Flavin (1981), is that the change in consumption should be adequately modeled with a random walk. In testing this supposition, however, it was found that other lagged variables (disposable income) do indeed contain information. That is, consumption is more sensitive (excessively sensitive) to predictable changes in income than is warranted by the PIH. Explanations as to why consumption may display this property are discussed in the following section.

Another paper in the excessive sensitivity tradition, is Campbell (1987). What distinguishes this paper from Flavin (1981) is that more information than simply consumption and income is utilized in estimation. Building on the work of Flavin, Campbell posits that the agent consumes some proportion of permanent income, $c_t = \gamma y_{pt}$, where $0 < \gamma < 1$. In this framework, if the PIH is correct, dissaving anticipates rising income and saving anticipates
falling income (p. 1250). A test of the PIH follows by studying whether these theoretical hypotheses are consistent with empirical facts.

In deriving a structural form of the PIH, Campbell defines savings, $s_t$, as

$$s_t = y_t - c_t,$$

where $y_t = y_{lt} + y_{kt}$, $c_t$ denotes nondurable consumption, $y_{lt}$ is disposable labor income, and $y_{kt}$ is disposable capital income. As presented in Flavin (1981), consumption can be expressed as

$$c_t = \gamma\left(y_{kt} + r/(1+r)\sum_{i=0}^{\infty}[1/(1+r)]^i E_t y_{t+i}\right),$$

which can be rewritten into the following statement in terms of savings, $s_t$ (p. 1253):

$$s_t = -\sum_{i=0}^{\infty}[1/(1+r)]E_t \Delta y_{t+i}, \quad \text{or} \quad s_t = \Delta y_{lt} - (1+r)s_{t-1} = -r\varepsilon_t,$$

where $\varepsilon_t = [1/(1+r)\sum[1/(1+r)]^i(\varepsilon_t y_{lt+i} - E_{t-1} y_{lt+i})].$

Equation (2.10) suggests a new interpretation of excess sensitivity because Campbell describes consumption as being “... excessively sensitive if it moves too closely with income - that is, if savings the difference between consumption and income moves less than the unrestricted forecast of the present value of labor income declines.” (p. 1254)

To test the PIH, Campbell specifies a process for $\Delta y_{lt}$ and estimates this jointly with $s_t$, subject to a number of restrictions. However, estimation is contingent upon the stationarity of $s_t$. Campbell takes $s_t$ as being a cointegrating relationship among the variables $y_{lt}$, $y_{kt}$, and $c_t$, suggesting that $s_t$ is stationary. Imposing the restrictions implied by the PIH (see pp. 1257-1258), it is found that $s_t - \Delta y_{lt} - (1+r)s_{t-1}$, which should be unpredictable given lagged $\Delta y_{lt}$ and $s_t$ (p. 1258). Column 3 of Table IV of Campbell’s paper shows that the null hypothesis that coefficients for lagged $\{\Delta y_{lt}\}$ and $\{s_t\}$ are jointly zero, for lags of one and five, can be rejected implying that the PIH is also rejected.

The theoretics of the PIH, along with empirical stylized facts dealing with consumption and income, suggest another type of test of the PIH. From its beginning, the PIH was offered as an explanation for the observed smoothness of consumption relative to income. By positing that the agent consumes relative to permanent income rather than to current
income, changes in consumption will be less volatile than changes in income, which is the case (Christiano 1987, p. 2). This result suggests that the observed variance in permanent income should be less than the variance of income. Testing this supposition, through parametric and nonparametric techniques, forms the basis of the smoothness literature.

Central to this body of work is Deaton (1987), whose paper initiated the Deaton Paradox and spawned a number of papers seeking to provide an explanation for it. Deaton observed that if income is modeled as a difference stationary series, the innovations in the measured income series change permanent income by a larger-than-proportionate amount. That is, permanent income is more volatile than measured income under the assumption of difference stationary income. Assuming that the PIH holds, this implies that consumption is also more volatile than measured income. This is, of course, contrary to the stylized facts and the basic premise of the PIH itself, which was promoted as an explanation of the smoothness of consumption relative to disposable income.

We can display this property with a simple example. Consider the case where income can be modeled as an ARIMA(1, 1, 0) process, that is; \( \Delta y_t = \rho \Delta y_{t-1} + \epsilon_t \). Following Campbell and Deaton (1989, pp. 358-359) the change in consumption can be expressed as

\[
\Delta c_{t+1} = r \sum_{i=1}^{\infty} (1 + r)^{-i} (E_{t+i} - E_t) y_{t+i},
\]

where \( y_t \) denotes labor income and \( r \) is the real interest rate.

Assuming rational expectations and given that \( \text{Var}(\epsilon) = \sigma^2 \), it can be shown that

\[
\text{Var}(\Delta c_t) = \text{Var}(\Delta y_{t+1}) = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} \rho & 1 - \rho \\ 1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} \rho & 1 - \rho \\ 1 & 0 \end{bmatrix}^{-1} \epsilon_t \epsilon_t' \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} \rho & 1 - \rho \\ 1 & 0 \end{bmatrix}^{-1} \sigma^2 = \text{Var}(\Delta y_t),
\]

where \( \rho = (1 + r) \)

and \( A = \begin{bmatrix} 1 + \rho & -\rho \\ 1 & 0 \end{bmatrix} \).

Equality between the variances of permanent and labor incomes holds only if \( \rho \) is 0; that is, when income is a random walk. To illustrate this concept, consider an example from Campbell and Deaton (1989), who found that quarterly labor income can be modeled as
Ay_t = 8.2 + 0.442Ay_{t-1} + \varepsilon_t, \quad \text{Var}(\varepsilon_t) = 25.2

Assuming an annual real interest rate of 10 percent (i.e., a quarterly rate of 2.41 percent), the innovation in the change in consumption can be expressed as:

\Delta c_t = 1.759\varepsilon_t, \quad \text{or} \quad \text{Var}(\Delta c_t) = 3.09 \cdot \text{Var}(\varepsilon_t) = 3.09 \cdot 25.2,

suggesting that the variance associated with the innovation in the change in consumption is larger than that associated with income. However, the variance of \Delta c_t as computed using the data of Campbell and Deaton is only 27.3. By using a difference stationary notion of income, the variance of \Delta c_t is overpredicted. It should be noted, however, that if the coefficient of \Delta y_{t-1} is less than one, the predicted variance of consumption would be less than that of income. Empirical evidence, however, shows that this coefficient is positive.

Campbell and Deaton (1989) approach this paradox by using superior information. Their paper begins by showing the existence of excess smoothness for the case when income is assumed to be a difference stationary series. However, the authors point out that it may be too simplistic to assume that agents use only labor income in formulating their notions of permanent income. It would seem more likely that agents formulate expectations relative to a richer information set. If agents do have extra information, the effect is to smooth permanent income relative to the permanent income measure based on the univariate labor income series (Campbell and Deaton 1989, p. 363). Thus, smoothness may arise naturally in the case where the information set used by the agent is larger than that assumed by the econometrician. Consider, for example, the situation where an agent formulates his notion of permanent income using an information set that is larger than that used by an econometrician who is attempting to predict the agent’s level of permanent income. In this case, the variance of the innovation associated with the agent’s view on permanent income will be smaller than that associated with the researcher’s prediction. Consumption may appear smoother because the wrong notion of permanent income is being used by the econometrician. Thus, it may not be the PIH which is faulty in regards to excess sensitivity; rather, its existence may arise from the inappropriate manner in which permanent income has been modeled.
To examine the possible effect of a wider information set, Campbell and Deaton formulated a bivariate model that allows for the effects of superior information to be examined. Because it is impossible to model all relevant information, consumer behavior is used to reveal consumer expectations; that is, if the PIH is correct, consumers will reveal their estimates of permanent income through their consumption and savings decisions (p. 363).

The approach used builds on the work of Campbell (1987). First, a savings equation is specified such that

\[(2.11) \quad (s_i / y_i) = - \sum_{i=0}^{\infty} \rho^i \text{E}[\Delta y_{t+i} | I_t] - k,\]

where \(s_i\) denotes savings, \(y_i\) denotes income, \(I_t\) denotes the consumer's (unknown) information set, and \(k\) is some constant.

Consider now if expectations of Equation (2.11) are taken with respect to the information set \(H_t\), which is comprised of \(\{s_t / y_t, y_t\} \cup I_t\). In this case,

\[(2.12) \quad (s_i / y_i) = - \sum_{i=0}^{\infty} \rho^i \text{E}[\Delta y_{t+i} | H_t] - k\]

or, as Campbell and Deaton show,

\[(2.13) \quad (s_i / y_i) - \Delta y_{t-i} - \frac{s_{t-i}}{\Delta y_{t-i}} = - \sum_{i=0}^{\infty} \rho^i \{\text{E}[\Delta \ln y_{t+i} | I_t] - \text{E}[\Delta \ln y_{t+i} | H_t]\}\]

Equations (2.12) and (2.13) present a simple way to test the PIH, even though the information set used by the agent is not known. A VAR system containing \(\Delta \ln y_{t-1}\) and \(s_t / y_t\) can then be estimated and the resultant forecasts of future income growth can be weighted together to form the right-hand side of equation (2.12), which can be compared with the left hand side (p. 363).

In specifying this vector autoregressive (VAR) system, the model imposes two restrictions (see Campbell and Deaton 1989, pp. 364-365 for more detail). The first restriction precludes excess sensitivity while the second restriction, if satisfied, rejects the notion of excess smoothness. It is shown, however, that the condition for the absence of excess sensitivity and the condition for smoothness are identical, requiring that only one restriction be
tested. Rejection of this restriction would offer evidence against the PIH, even though superior information is used. Campbell and Deaton find that this restriction is rejected in most specifications (see Table III, p. 366), implying that consumption is smoother than it ought to be under the PIH, even when superior information is taken into account.

The final paper considered in this section is by Flavin (1993) and combines tests of orthogonality, excess sensitivity, and excess smoothness in one model. This paper generalizes the work of Campbell and Deaton (1989) with the specification of a specific alternative hypothesis following the rejection of the PIH. Flavin's alternative hypothesis posits that consumption is excessively sensitive to current income; in particular,

\[ c_t = \beta y_t^T + y_t^p, \]

where \( \beta \), the excess sensitivity parameter, is between zero and one, \( y_t^T \) denotes transitory income, and \( y_t^p \) is permanent income (p. 654). Using this notion of consumption, Flavin expresses the innovation in consumption as

\[ \varepsilon_{ct} = \beta \varepsilon_{yt} + (1 + \beta) \varepsilon_{yp}, \]

where \( \Omega_{t-1} \) is the information set used by the econometrician, \( \varepsilon_{ct} = \Delta c_t - E(\Delta c_t | \Omega_{t-1}) \), \( \varepsilon_{yt} \) is the innovation in \( \Delta y_t \) relative to the information set \( \Omega_{t-1} \), and \( \varepsilon_{yp} \) denotes the innovation in permanent income contingent upon the agent's information set \( I_{t-1} [\Omega_{t-1} \subset I_{t-1}] \). Excess smoothness will occur if

\[ \text{var}(\varepsilon_{ct}) = \beta^2 \text{var}(\varepsilon_{yt}) + 2\beta(1 - \beta) \text{cov}(\varepsilon_{yt}, \varepsilon_{yp}) + (1 - \beta)^2 \text{var}(\varepsilon_{yp}) < \text{var}(\varepsilon_{ct}) \]

It is noted that smoothness depends on the covariance between \( \varepsilon_{yt} \) and \( \varepsilon_{yp} \). In the general case where the agent lacks perfect one-period-ahead forecastability but nevertheless forecasts on the basis of a strictly larger information set than that used by the econometrician \( (\Omega_{t-1} \subset I_{t-1}) \), \( \varepsilon_{yt} \) and \( \varepsilon_{yp} \) will be correlated, albeit imperfectly. In this case the imperfect correlation of the two deviates works to smooth consumption relative to permanent income, although the strength of the smoothing is inversely related to the correlation between the two deviates. Equation (2.14) shows that excess smoothness of consumption is not a paradox, but
rather is what would be expected if consumption is generated by the excess sensitivity hypothesis (Flavin 1993, p. 657).

Following Campbell (1987) and Campbell and Deaton (1989), a bivariate autoregression of income and savings is specified. Flavin shows that the PIH implies certain restrictions on such a system (pp. 658-659). In particular, two restrictions are derived that relate to smoothness and orthogonality issues. Flavin finds that although the orthogonality condition implies the smoothness condition, the converse is not true.

An important finding in Flavin's paper is that if the PIH fails in any way, the savings series will not fulfill its crucial role of fully capturing all information available to agents, with the consequence that the variance of revisions in permanent income $\text{var}(e_{yp})$ will not be identified (p. 659). This situation shows the importance of specifying an alternative to the PIH, especially given its empirical failure. It is shown that the $\text{var}(e_{yp})$ is identified when this alternative model is used (p. 660), which permits the examination of the smoothness question, using Equation (2.14), when the PIH is rejected. In attempting to test for smoothness, Campbell and Deaton (1989) fail to use a consistent estimate of the true $\text{var}(e_{yp})$. Again, this failure arises because, as Flavin shows, $\text{var}(e_{yp})$ is not identified for any arbitrary departure from the null hypothesis. Generalizing Campbell and Deaton's model so that $\text{var}(e_{yp})$ is identified under the alternative hypothesis of excess sensitivity allows for a much stronger statement to be made with regard to smoothness.

The excess sensitivity model also imposes restrictions on the parameters of the bivariate system. Using this model, orthogonality and smoothness can also be tested. For estimation, Flavin began by examining the three closely related, yet distinct, hypotheses to be tested in the context of a bivariate autoregression of income and savings: orthogonality, excess sensitivity, and smoothness. Test results, reported in Table I of Flavin's paper (p. 662), show that in the case of nondurable consumption and services, the orthogonality and smoothness restrictions can be rejected (using the Wald test and one to five lag specification). This result shows that, in accord with the other papers discussed, evidence contrary to the PIH is found. Restrictions implied by the excess sensitivity model also cannot be rejected.
Estimates of the excess sensitivity parameter, $\beta$, are significant and range from 0.37 to 0.50, depending on the specification being used.

Using the excess sensitivity alternative Flavin, then proceeds to determine the measure of $\text{var}(\varepsilon_{yH})$ (see p. 663 for details). She finds that the estimated standard deviation of $\varepsilon_{yH}$, stated as an annualized growth rate, is 4 percent with a standard error of approximately 1 percent. This value is compared to the standard deviation of the annualized growth rate of consumption of approximately 2.2 percent, with a standard error of 0.15 (p. 663). Thus, Flavin reaches the same conclusion as Campbell and Deaton (1989), namely, that consumption is too smooth under the PIH, suggesting that the PIH as currently formulated is incorrect.

The PIH, as formulated in Hall (1978) and Flavin (1981), makes strong predictions of how consumption should evolve. One prediction, which Hall showed, was that the change in consumption ought to be orthogonal to all information in the preceding periods. A second prediction is that consumption should not respond to predictable changes in income. A third is equality of variances among consumption and permanent income, with both being less than the variance of income.

A number of the studies reviewed herein, however, find that these conditions are not upheld. Rather, the change in current consumption depends on other variables, including income, while the variability of consumption is found to be greater than that of income under certain specifications. These empirical realities suggest that the PIH as it is currently formulated is incorrect. In the following section, we examine capital market imperfections as one possible explanation for this apparent failure. A model that incorporates market failures, developed in Chapter 4, is sufficiently flexible so as to display the non-orthogonality, excess sensitivity, and excess smoothness observed in the literature.
Imperfections

With the apparent rejection of the PIH, as noted above, an explanation must be offered. A number of researchers have focused attention on the assumptions upon which the PIH is constructed. For example, Caballero (1990) has questioned the virtues of the quadratic utility function. Though used for tractability, important aspects of consumption, such as the effects of precautionary savings (which relies on the third derivative), are ignored. The assumption of constant real interest rates is addressed in papers by Wickens and Molona (1983) and Hall (1988) and the effects of stochastic interest rates are considered. However, the assumption of most relevance in this section is that of perfect capital markets.

The excess sensitivity results, central to Flavin (1981) show that aggregate consumption is responsive to anticipated changes in income. Under the PIH, agents by consuming according to their notions of permanent income and rational expectations, should not be affected by anticipated changes in income. With perfect markets an unexpected lean year in terms of earned income will not interfere with consumption because agents can borrow freely. The excess sensitivity finding may be attributed to the presence of a significant proportion of liquidity constrained consumers (Deaton 1987, p. 130) who, being unable to borrow sufficient amounts, may be forced to consume at a proportion of their current income levels.

This situation may arise in the following manner. Consider the case where there is some persistence in income, a positive income shock today signals higher future income and an upward revision in permanent income and consumption. If the agent is liquidity constrained (unable to borrow desired amounts) however, there exists an inability to borrow against this higher income. As a consequence, when the higher income is realized, the next period's consumption rises closer to the optimal level (Acemoglu and Scott 1994, p. 9). Econometrically, the change in income will have a significant impact on the current change in consumption, meaning that excess sensitivity will be observed.
To many researchers, liquidity constraints are a symptom of a more general problem, namely, imperfections in capital markets. Under the umbrella of perfect markets that the PIH presupposes are notions of perfect information, absence of transactions costs, equality of the lending and borrowing rates, and the absence of any form of discrimination. Agents operating in this environment, are able to finance consumption regardless of amount, given that the transversality condition is not violated. If any of these requirements of perfect markets fail to hold, agents may find themselves unable to achieve desired consumption levels.

Whereas the assumption of perfect markets may expedite a tractable solution, there is a consensus that the assumption may be too strong. Again, the existence of excess sensitivity to income is portended by many as suggesting liquidity constraints. Testing and modeling this broad notion of imperfect markets is the subject of this section. The literature in this area is large and growing, so only a sampling is considered. Following this review, an alternative to modeling market imperfections will be offered.

Before proceeding, one comment is offered about modeling imperfect markets. As will be seen below, researchers have used two approaches to this problem. The first is simply to posit that an agent, or a group of agents, is liquidity constrained. No explanation or insight is offered as to what may have caused this condition to arise or how long it may continue. This method may be seen as abstracting from the particular market situation that may place a limit on the agent’s actions. The second approach explicitly specifies the market situation that may limit consumption. Here, an exact imperfection is modeled and permitted to directly affect the agent’s decision process. Though this second method may be more precise, compared to the first method, it has the serious drawback that a decision must be made ex ante as to the imperfection to which the agent is reacting. For tractability, such an exercise usually ignores other possible imperfections, suggesting that although the agent is astute enough to recognize some specific imperfection, he is ignorant to the possibility of other markets imperfections. Most of the papers reviewed next use the former approach.

The first paper being considered is by Flavin (1985), who examines the presence of liquidity constraints as contributing to the excessive sensitivity of labor income noted in the
empirical estimation of the PIH. As discussed, an agent who is unable to borrow sufficient amounts will have a greater propensity to consume from current income. This behavior, however, may be fully consistent with a rational forward-looking agent (Flavin 1985, p. 118). Flavin posits the following model:

\[
\Delta c_t = \beta_p \Delta y_p^t + \beta_t (\Delta y_t^t - \Delta y_t) + \gamma \Delta Z_t,
\]

where \( y_p^t \) denotes permanent income, \( y_t \) is labor income, \( \beta_p \) and \( \beta_t \) are the respective marginal propensities to consume from permanent and transitory incomes, and \( Z_t \) is some proxy for the severity and prevalence of liquidity constraints (Speight 1990, p. 173). The PIH is a special case that arises when \( \beta_t = \gamma = 0 \). In estimation, \( Z_t \) is represented by the unemployment rate and empirical results are provided for the case when \( \gamma \) is included and also excluded from the model. Exclusion of \( \Delta Z_t \) shows that excessive sensitivity parameters cannot be rejected (p. 130). The inclusion of \( \Delta Z_t \) provides less clear results, although Flavin finds that the null hypothesis that \( \beta_t = \gamma = 0 \) can be rejected at the 0.5 percent level (p. 133). Thus, excess sensitivity is found along with a significant unemployment effect. The conclusion of excess sensitivity is not surprising given the results of the previous section. One possible explanation for the significance of the unemployment rate is provided below.

Jaffee and Russell (1976) studied the effect that imperfect information may have on consumption. By dividing consumers into honest and dishonest borrowers, with regard to the repayment of loans and assuming that lenders have imperfect information as to the group to which some particular agent belongs, various types of behavior can be observed. The authors formulate a model where a stable equilibrium is reached in which all individuals are rationed in the amount they can borrow (p. 664). Contrary to this rationing case, a second model shows that a nonrationing equilibrium may also arise where a divergence between the lending and borrowing rates will be observed. Lastly, Jaffee and Russell present a model where a monopolist operates in the lending market. Under this noncompetitive framework, rationing will not be observed, rather loan rates will increase (p. 663).
Although very simplistic in terms of modeling, the Jaffee and Russell paper reveals two important points. First is the effect that imperfect information may have on consumption. Behavior as assumed in the PIH, such as equality in lending and borrowing rates and the absence of credit rationing, can only be guaranteed when perfect information exists. Second is the variety of results that may be observed from the existence of imperfections. Currently, we see that the effect of imperfect information is a divergence in interest rates or credit rationing (that is, rationing which is not achieved through a pricing mechanism). When imperfect information is coupled with imperfect competition in the lending market, a divergence between lending and borrowing rates will be observed as lenders attempt to maximize profits. These results suggest that when imperfections are explicitly discussed, the resultant effect on consumption may not be unique. For example, divergence in interest rates may be observed, given perfect competition in the lending market and imperfect information or when there is imperfect competition in the lending market and imperfect information. Alternatively, if the researcher decides a priori that imperfect information is the relevant imperfection to be modeled, potentially important information is being ignored when the form of competition in the lending market is deemed irrelevant. Possibly ambiguous (theoretical) results would disappear once the imperfect nature of the lending markets are incorporated into the analysis. We shall say more about this explicit approach of modeling imperfections later.

Another paper that utilizes this explicit approach is by Pissarides (1978), who studied the role of liquidity and the assumption that there are transaction costs in the asset market. Effects from such an assumption are that the borrowing rate exceeds the lending rate and that assets can only be realized before maturity at a cost (p. 280). To best incorporate this notion of transaction costs, consumption is determined simultaneously with the desired asset portfolio. Joint optimization will ensure that the level of consumption undertaken results in a minimum of transactions costs.

The model solution reveals that consumption remains a homogeneous function of lifetime wealth, as predicted by the PIH, but where the discount factor used in determining wealth is a function of asset maturities and rates of return; transactions costs, and the future
income profile (Speight 1990, p. 160). Pissarides (1978) points out that one result of this is that the information contained in the current flow of income and the composition of the asset portfolios should be a more useful predictor of consumption than the market measure of total wealth, although consumption will still be a function of wealth (p. 294). It is therefore possible to use transactions costs as one possible explanation of excess sensitivity previously noted.

Similar to Jaffee and Russell (1976), King (1986) presents a model where asymmetrical information exists between lenders and borrowers and studies effect of such a situation on aggregate consumption. This asymmetrical information arises from the fact that lenders may not know the true probability of default. Assuming quadratic utility, King shows that imperfect information will manifest itself by a divergence between the lending and borrowing rates. Equilibrium is characterized by a piecewise linear budget constraint subject to which agents choose their optimal consumption plan. King notes that no agent will be rationed out of the credit market; rather, the result of incomplete information is a wedge, which is endogenous to the model, between the lending and borrowing rates (p. 69). Aggregating these results over all agents in the economy shows that current consumption is a function of lagged consumption and the expected change in wealth. Further, coefficients of these variables will be time-dependent and are functions of the wedge between the lending and borrowing rates (p. 72).

Estimation proceeds by estimating a two-equation system composed of a consumption function and the wedge between the lending and borrowing rates. In the case of the United States, this wedge is defined as the difference among the twelve-month Treasury bill and the twelve-month commercial bank loan rate. King posits that this wedge is a function of the lending rate, the unemployment rate, and the change in the unemployment rate. Figure 3 of King’s paper (p. 76) reveals that the unemployment rate and the interest rate wedge possess similar time paths, at least for the United States. Both equations are estimated using instrumental variables, which include lagged values of the money stock, interest rates, consumption, unemployment, and the real value of the stock market index (p. 74). The results
shown in Table 2 of King’s paper (p. 75), are as expected. The higher the lending rate, the fewer the number of agents who will borrow and the smaller the wedge. Unemployment enlarges the wedge; however, if unemployment is rising, future prospects are poor, borrowing is reduced, and the wedge decreases (p. 75).

A very different approach is used by Campbell and Mankiw (1990) who studied the extent to which U.S. consumers are liquidity constrained. Disposable income, \( Y_t \), is assumed to be divided among two groups: group one is assumed to receive a fixed share \( \lambda \) of total disposable income; i.e., \( Y_{1t} = \lambda Y_t \), while group two receives \( Y_{2t} = (1 - \lambda) Y_t \). Group one is additionally assumed to be liquidity constrained and to consume all of their disposable income; i.e., \( C_{1t} = Y_{1t} \). Speight (1990) shows that consumers who expect to be liquidity constrained indefinitely will consume according to their current levels of income. Group two, which is not constrained, will consume according to the PIH, that is, \( C_{2t} = (1 - \lambda) Y_t^e \). Following Hall (1978), \( \Delta C_{2t} = (1 - \lambda) \epsilon_t + u \), where \( u \) is a constant, \( \epsilon \) is the innovation in permanent income between time \( t-1 \) and \( t \), and under the assumption of rational expectations is a white noise term (Campbell and Mankiw 1990, p. 266). Aggregate consumption can thus be expressed as \( \Delta C_t = \Delta C_{1t} + \Delta C_{2t} = u + \lambda \Delta Y_t + (1 - \lambda) \epsilon_t \).

It is noted that the PIH is nested in this model and arises when \( \lambda = 0 \). Following estimation using instrumental variables, it is found that \( \lambda \) is significantly different from zero (thus negating the PIH) and that its point estimate of 0.5 suggests a substantial departure from the PIH (p. 277).

An analogous result is found by Cushing (1992) who finds that 30 percent to 40 percent of U.S. consumption is accounted for by liquidity constrained consumers (p. 135). As in Campbell and Mankiw (1990), Cushing considers two types of consumers: one type is unconstrained and consumes according to the PIH (i.e., \( c_t = \lambda_0 + \lambda_1 c_{t-1} \)) while the other is liquidity constrained and consumes a fraction of its disposable income (\( c_t = \pi y_t \)). Aggregating these two groups, Cushing shows that total (aggregate) consumption can be expressed as \( c_t = (1 - u)\lambda_0 + \lambda_1 c_{t-1} + u\pi (y_t - \lambda_1 y_{t-1}) + (1 - u)\epsilon_t \).
where $\upsilon$ denotes the fraction of total consumption accounted for by liquidity-constrained individuals (p.137). Positing a process for $y_t$ and imposing cross equations restrictions, Cushing finds strong divergence from the PIH (p. 138).

Last, we consider the results of Jappelli (1994), who used the 1983 Survey of Consumer Finances to ascertain the proportion of consumers that could be categorized as liquidity constrained. Jappelli found that the proportion of rejected applicants and discouraged borrowers was in the 20 percent range (p. 230), a result similar to the panel study undertaken by Hall and Miskin (1982).

Given the existence of market imperfections, the appropriate manner of incorporating these facts into consumption modeling must be addressed. Each of the papers just reviewed approaches the question from a particular perspective. King (1986) approaches the problem by hypothesizing ex ante that imperfect information exists about default probabilities and that such an imperfection will manifest itself as the deviation between the lending and borrowing rates that the consumer is able to proxy quite well with the unemployment rate. Jaffee and Russell (1976) obtain a deviation between the lending and borrowing rates, or of credit rationing, by assuming imperfect information on the part of lenders. Pissarides (1978) assumes that assets are prone to transactions costs and shows that consumption is dependent on more than simply some notion of wealth. These three papers share the common trait that a specific market imperfection must be explicitly specified as being relevant to the agent’s determination of consumption. Such an approach requires a priori that the market failure confronting the agent is known. This approach to modeling imperfections is very explicit compared to the following methodology, which just examines the effect that some aggregate notion of market imperfections may have on the agent.

Cushing (1992) and Campbell and Mankiw (1990) present a more general approach compared to the three studies just discussed. By simply focusing on the general notion of liquidity constraint, Cushing and Campbell and Mankiw abstract from the particulars that may cause such a situation to arise. Compared to the explicit approach this more general approach may be fortuitous because it may be difficult to explicitly single out ex ante the constraint an
individual may face. One caveat to this approach is provided by Speight (1990), who concludes that the available empirical evidence is unable to definitely assess the qualitative and quantitative significance of constrained liquidity. This caveat arises because no structural model of liquidity constraints can be derived for the general case (p. 177). Removing explicit causes thus introduces some ambiguity into the analysis.

The more general approach models imperfections in a very different manner compared to the first approach. Rather than attempting to incorporate all the imperfections that may affect the agent, this general approach simply analyzes the effects that will occur from the aggregate existence of such phenomena. That is, only the effect of all imperfections will be studied. The tangible effect of this situation is that the agent is liquidity constrained. The term liquidity constrained is characterized by Speight (1990) as an umbrella term that covers a wide variety of possible characterizations. Central to this definition is the inability of the agent to borrow desired amounts, for whatever reason. By attempting to model this notion of liquidity constraints the researcher achieves the advantage of not having to precisely specify what caused such a situation. In essence, the cause is ignored. Intuitively, this approach seems more amicable to the notion of agents who are not omniscient. By simply concentrating on the effects of imperfections, we do not force an agent to have such a level of information that he is able to identify all possible market failures that may confront him. Agents in this framework only become cognizant of imperfections by viewing the effects that such situations may have on their ability to consume. This modeling of effects, rather than causes, acknowledges that in a complex economy identifying of all possible phenomena that may affect the individual is impossible.

However, there are those who criticize this approach. Hayashi (1987) argues that unless the exact nature of the market imperfection is identified, the economic implications cannot be determined (p. 111). Although this objection may be legitimate when agents can uniquely identify the imperfections that confronts them, the objection seems irrelevant when the multifarious imperfections of a complex economy are considered. In this situation, it may
only be possible to analyze the effects that such a combination of imperfections may present to the consumer.

Hayashi also questions the ability to derive economic implications from this liquidity constraint framework. However, important policy implications are seen to arise even when this more simplistic approach is used. Consider Hall's (1978) random walk model, which hypothesizes that current consumption is orthogonal to all information at time \( t-1 \), except for lagged consumption. In this world, government intervention has no bearing on consumption. This result compares to the model specified by Flavin (1981, 1985), where consumption is excessively sensitive to changes in income, a property that may be explained in terms of liquidity constraints. Sensitivity to income, however, offers the government a means of altering consumption levels, for example, by manipulating tax rates that affect current consumption through changes in disposable income. As Hayashi (1985) concludes, if households are subject to borrowing constraints, then short-run stabilization policies will have some influence on aggregate demand (p. 185).

The following section presents criticisms of the current manner in which market imperfections have been incorporated into the modeling of consumption. These criticisms relate to the degree of information researchers have assumed upon the agent and to the type of imperfections that are assumed to exist. An alternative method of incorporating market imperfections that avoids the problems noted with contemporary models is also discussed.

Criticisms

Both the approaches described in the previous section require assumptions about the consumer. The explicit approach used by King (1986), Jaffee and Russell (1976), and Pissarides (1978) necessitates the assumption that agents are cognizant of the particular market situation confronting them. In King's paper, endogenous deviations between lending
and borrowing rates are deemed important to the agent. Imperfect information on default probabilities establishes a theoretical basis for such deviations. However, Jaffee and Russell show that imperfect information may also result in credit rationing, a situation that may produce different consumption patterns compared to those obtained given a divergence in interest rates. Thus, very different results may be found, even when the same imperfection is utilized in both cases.

Another problem with the explicit approach deals with the legitimacy of ignoring other possibly important imperfections. As discussed previously, certain important information may be lost when a researcher accepts one imperfection as being relevant to the exclusion of others. Consider the paper by Pissarides (1978), where the effects of transaction costs on assets are analyzed. It may very well be that transaction costs are incurred because of imperfect information, or even laziness, on the part of the agent. Simply utilizing transaction costs may ignore deeper possible causes and thus possible significant effects on consumption. For example, consider the case where some type of transaction costs are associated with imperfect information (e.g., higher fees generally charged in the inner city). Modeling transaction costs as the relevant failure while ignoring the existence of imperfect information results in a loss of pertinent information. Ignoring relevant imperfections may thus leave consumption misspecified, at least with regard to the true economic environment.

It would seem very unlikely; however, that all market failures could be incorporated into the determination of consumption. Beyond identifying each failure, exact specification would be necessary to model how each imperfection would affect the consumer. This task is all the more daunting given the interaction among, and the multidimensionality of, potential market failures. The relevant question for which no answer is given, is how much good or bad is done to consumption modeling by simply positing one market failure to the exclusion of all others.

Operationally, this explicit approach is seen to make an obvious informational assumption about the agent. In such a framework, the agent is assumed to possess sufficient knowledge of the economic system to allow for the relevant imperfection to be identified,
modeled, and incorporated into the determination of consumption. While a description of how
such information is obtained is not offered, application of this method proceeds by having the
researcher select an imperfection that is deemed significant or interesting. Little attention is
given as to how an agent operating in a complex economy would choose that particular
manifestation of market failures as being relevant, to the exclusion of all other possibilities. In
actuality, attempts at selection may require enormous assumptions, given the shades of gray
that distinguish possible imperfections. For example, credit rationing may be caused by
discrimination, transaction costs, or ignorance. Explaining such phenomena is difficult, given
the multitude of contributing factors and because there may be a great deal of interaction
among these factors. It may be that discrimination and/or transactions costs are incurred
because the agent is unwilling or even unable to perform necessary market research.

The liquidity constraint approach has been pursued by many because it avoids the
complications that arise when an explicit imperfection is modeled. However, this approach,
used for example by Cushing (1992) and Campbell and Mankiw (1990), is not a panacea. We
offer two objections to the approach. The first objection relates to the arbitrary manner in
which the population is divided into liquidity constrained and unconstrained. No explanation is
offered by these authors as to how agents are assigned to their respective grouping or whether
movement among groups is permitted. Whereas such a division appears natural in a panel
study, its use in an aggregate study seems dubious. Consider, for example, Campbell and
Mankiw (1990). In their paper, one segment of the population is assumed to receive some
fixed share, $\lambda$, of total disposable income. Rationale for how a particular agent is assigned to
this segment or to another is not offered.

Our second objection deals with how well this liquidity constrained approach
incorporates the general notion of market imperfections into the modeling of consumption.
This paper takes the view that the notion of being liquidity constrained is more a symptom
than a cause when discussing market failures. An agent may be considered liquidity
constrained when he is unable to borrow desired amounts, possible reasons for which are
credit rationing, discrimination, or the lack of information.
Objection is found not in the use of the notion of using liquidity constraints to proxy market imperfections, but rather in that such use necessarily limits the extent to which imperfections are permitted to affect the agent. This objection arises because there is no guarantee that such a representation will reflect all the imperfections that may confront agents. Phenomena such as transaction costs, whether arising from losses on liquidated assets or from the breaking of contracts, have the effect of limiting consumption without diminishing the ability of the agent to borrow. By ignoring these possibilities, the liquidity constrained approach is seen to reflect only a portion of the total set of market failures that may influence the agent’s consumption decision.

Applications by Cushing (1992) and Campbell and Mankiw (1990) seem to compound the problem because the unconstrained segment of the population is assumed to consume according to the PIH, which depends on the implicit assumption that this group is completely unaffected by any market failure. This assumption, however, appears to be very strong. Market realities, such as the divergence in interest rates, transaction costs, and imperfect and costly information, contrive to make the predictions of the PIH untenable. These conditions exist for all agents in the economy, albeit by varying degrees. Viewed relative to this point, the assumption that some population segment [estimated to be a majority of the population in studies by Cushing (1992) and Campbell and Mankiw (1990)] encounters no imperfections in determining a consumption path is very unlikely. It would seem more appropriate to allow market failures to affect all agents, although possibly by varying degrees, as consumption is being determined.

Incorporating imperfections using the approach exemplified by King (1986) or the method characterized by Cushing (1992) is thus seen to encounter problems. We now introduce an alternative approach that addresses the objections just noted. This approach, which forms the basis for the remainder of this paper, attempts to model imperfections in a manner that is tangible to the economic agent. In particular, this alternative approach relies on the supposition that it is not the imperfections per se to which the agent is reacting, rather, the agent will react to the costs that these market failures may impose upon that agent.
Market imperfections may manifest themselves to the agent in two ways. First, like any other maximization problem, imperfections may act as constraints which may result in lower or unchanged consumption levels. In determining consumption, the existence of market failures will generally lead to a reduction in the consumption levels that would have occurred in the absence of such imperfections, at least when viewed relative to the PIH. The second effect is the elimination of instantaneous adjustment because once imperfections are encountered, rigidities will generally occur. Altering consumption levels in this situation may thus expose the agent to pain.

One quite plausible method of modeling such manifestations is through the use of costs. In particular, we seek to interpret the effects of market imperfections by measuring the costs that such phenomena may generate. As mentioned, two types of costs may become evident as the agents attempt to formulate a consumption path. One cost may arise from the inability to obtain desired (unconstrained) levels of consumption, while the second is associated with the pain incurred from altering consumption levels, ostensibly to obtain desired consumption levels. By modeling imperfections in this very tangible manner, we are able to avoid specifying particular manifestations of market failure, thereby avoiding the invocation of an omniscient agent who completely understands the economic system. As with the liquidity constrained approach, only an aggregate notion of imperfections needs to be modeled.

An additional advantage of this approach, especially when compared to the liquidity constraint models of Cushing (1992) and Campbell and Mankiw (1990), is that it allows market failures to affect all agents while permitting a greater variety of imperfections to influence the consumption decision. By modeling the impact that costs may have on the agent's consumption decision, no arbitrary division in the population must be made. Imperfections are thus permitted to affect all individuals in the economy. A second advantage is that we are able to use a broader notion of market failure than those that result in liquidity-constrained behavior. Additional market imperfections that agents may encounter when formulating consumption plans include transaction costs, losses on liquidated assets, and
possible search costs. This concept is important because no explicit assumption must be made as to what affects the agent. Because all imperfections are permitted to influence the consumption decision, the agent is placed in a more realistic environment.

Chapter 4 provides more detail on incorporating this notion of costs into the determination of consumption. Explanations are also provided as to what may cause costs associated with deviations from desired levels or from altering consumption from previous levels.

Before considering costly adjustment in greater detail, we use Chapter 3 to present the data used in Chapters 4 and 5. Modeling consumption as undertaken by the numerous studies reviewed herein, suggests particular variables. Nondurable consumption, disposable labor income, and capital income are considered the minimum group of variables necessary in studying consumption determination under the PIH. In the following chapter, we broaden this set of variables by presenting a reformulation of permanent income. Incorporating more information, with theoretical justification, offers a more interesting view of consumption determination.

After expanding the variable set, we can study whether these data are consistent with the PIH as developed in this chapter. Tests of orthogonality, excess sensitivity, and excess smoothness are all performed to ascertain whether our data are consistent with the PIH. By rejecting the PIH with our data, whose composition is implied by an alternative view of permanent income, we can consider formulating a consumption model that may explain this result. One model that may provide a possible explanation, a costly adjustment model, forms the basis of Chapter 4.
CHAPTER 3: DATA

Introduction

This chapter explicitly discusses the variables used in modeling consumption. Previous studies have tended to use nondurable goods and services, disposable labor income, and capital income in modeling consumption. We present a reformulation of permanent income that allows for the use of a broader information set, with theoretical justification, in determining consumption. Incorporation of additional information makes the analyses of the following chapters more interesting.

Beyond introducing the variables used in later chapters, definitions and analyses of these data are also considered in this chapter. Questions of whether the data are stationary, how many lags are necessary, and whether a group of variables are cointegrated are important for all the work to follow and are explicitly considered. Additionally, we examine whether these data are consistent with the observed non-orthogonality, excess sensitivity, and excess smoothness properties that were found to be displayed in much of the literature. Tests similar to those performed by Hall (1978), Flavin (1981), and Campbell and Deaton (1989), using the data of this chapter, will reveal whether our data are consistent with such reality. Confirmation of these properties offers a rationale for reformulating the permanent income hypothesis (PIH), which is the subject of Chapter 4.

The remainder of this chapter is organized as follows. The following section, entitled "Reformulation of Permanent Income", introduces a reformulation of permanent income which, with the assumption of Ricardian equivalence, shows that this quantity is a function of gross labor income, capital income net government bonds, and government expenditures. It is
this notion of permanent income that is used in Chapters 4 and 5. The third section, entitled "Variables", provides definitions of all the variables used in the following chapters. Descriptions of data collection and how these variables are formed are discussed, as are sources and conventions used in gathering these data.

The fourth section, entitled "Analysis of the Data", studies the statistical properties of the data. First, plots of each data series are provided, allowing us to examine how these series evolve over the sample. The obvious trend that these plots reveal raises the question of whether these variables are stationary; that is, whether a stochastic process possesses a time-invariant mean, variance, and autocovariances. With the obvious trend in these variables, stationarity appears doubtful. The possibility of nonstationarity is important because estimation using such variables is subject to well-known problems. Dickey-Fuller and augmented Dickey-Fuller tests are used in examining this possibility. These tests suggest that each series possesses a unit root, implying that the relevant variables are difference stationary.

Difference stationarity implies that each variable must be first differenced for proper modeling, something that is done for each variable. Because all the modeling in this paper is multivariate, however, the question of whether the relevant variables are cointegrated must also be addressed. Cointegration refers to the existence of some combination of difference stationary variables that is itself stationary. Testing for this possibility is important because a vector autoregressive system (VAR) using only first differences may be misspecified if the relevant variables are cointegrated. The maximum likelihood method of estimation developed by Johansen and Juselius (1990) is used to determine whether this property exists.

The last section, "Tests of PIH", presents simple tests of the PIH using the data described in this chapter. Tests of orthogonality, excess sensitivity, and excess smoothness are performed to examine whether the data are consistent with previous results. Rejection of the PIH with our data offers impetus for the material in Chapter 4, where an alternative method of modeling consumption is discussed.
Reformulation of Permanent Income

In the previous chapter, we explicitly assumed that the agent uses disposable labor income, defined as labor income net of taxes, in modeling permanent income. By modeling income in this manner we have implicitly assumed that taxation affects consumption only by affecting the level of disposable income. This section attempts to incorporate taxation explicitly and to analyze what effect this may have on the determination of consumption. An important result of this reformulation is to expand the variable set, with theoretical justification, that may be used in modeling consumption. By utilizing more information, model results may provide more insight. We start by representing disposable labor income as gross labor income net of taxes. With this redefinition of disposable labor income, wealth is assumed to evolve as

\[ A_{t+1} = R[A_t + y_t - \tau_t - c_t] \]

where \( \tau_t \) denotes real taxes, \( y_t \) is now defined as real gross labor income, and \( R \) is one plus the real interest rate, \( r \). With this alternative notion of wealth, consumption as discussed in the previous chapter may be rewritten as

\[
(3.1) \quad c_t = \left( \frac{-\alpha}{R - 1} \right) + \left( 1 - \frac{1}{BR^2} \right) \left[ \sum_{t=0}^{\infty} \left( \frac{1}{R} \right)^t E_t y_{t+1} - \sum_{t=0}^{\infty} \left( \frac{1}{R} \right)^t E_t \tau_{t+1} + A_t \right].
\]

An assumption is also made that will allow consumption in Equation (3.1) to be expressed as a function of expected current and future labor income, current wealth, and expected current and future government expenditure. Specifically, the assumption of Ricardian equivalence is made.

Ricardian equivalence proposes that deficit financing is no different from current taxation because individuals fully take into account the future taxes they will have to pay (Blanchard and Fischer 1989, p. 129). In the present case, such a theorem allows for
\[ \sum \left( \frac{1}{R} \right)^i E_t \tau_{t+i} \] to be expressed in terms of expected government expenditures; that is,

\[ \sum \left( \frac{1}{R} \right)^i E_t \tau_{t+i} = \sum \left( \frac{1}{R} \right)^i E_t g_{t+i} + \beta_t, \]

where \( g_t \) represents real government expenditures and \( \beta_t \) represents the net issue of government bonds in the current period (see Sargent 1989, p. 382). In this paper, \( \beta_t \) is assumed to equal the federal government's deficit/surplus. Substituting this expression for expected taxes into Equation (3.1) results in:

\[ c_t = \left( \frac{-\alpha}{R - 1} \right) + \left( 1 - \frac{1}{BR^2} \right) \left[ \sum_{i=0}^{\infty} \left( \frac{1}{R} \right)^i E_t y_{t+i} - \sum_{i=0}^{\infty} \left( \frac{1}{R} \right)^i E_t g_{t+i} + y_k \right], \]

or, with \( RB = 1 \),

\[ (3.2) \quad c_t = \left( \frac{R - 1}{R} \right) \left[ \sum_{i=0}^{\infty} \left( \frac{1}{R} \right)^i E_t y_{t+i} - \sum_{i=0}^{\infty} \left( \frac{1}{R} \right)^i E_t g_{t+i} + y_k \right], \]

where \( y_k = (A_t - \beta_t); \) i.e., wealth minus bond holdings.

The assumption of Ricardian equivalence is not inconvertible. Empirical evidence for and against this hypothesis is as large as it is diverse. Rather than entering into this foray, the reader is referred to a review article by Seater (1993), which provides some insight into this area of conflicting results. For our purposes, we simply assume the existence of Ricardian equivalence. Because we are operating in a paradigm in which a representative agent is infinitely lived, however, the assumption of Ricardian equivalence does not appear excessively strong.

Incorporating the assumption of Ricardian equivalence allows for consumption to be modeled as a function of \( (y_k, y_{kt}, g_k) \). By modeling taxes, we are able to determine explicitly what effect government expenditure may have on the evolution of consumption. This method, it is believed, will allow for a more interesting view on consumption modeling.

Before concluding this section, we must make a final point. In deriving \( A_t \), we follow the methodology shown in West (1988, p. 22), where \( \frac{R - 1}{R} A_t \) is taken to be the difference
between labor income and disposable income. As shown later, we modify this notion by
defining \( \frac{R-1}{R} A_t \) to be equal to the difference between labor income (before taxes) and
national income.

Variables

To model consumption, the following variables are used: nondurable consumption, \( c_t \),
which is comprised of nondurable goods and services; gross labor income, \( y_{lu} \); capital income
net government bonds, \( y_{ki} \); government expenditure, \( g_t \); and disposable income, \( y_{di} \). Disposable
income is included to perform tests of previous consumption theories, which is the subject of
the fourth section in this chapter. We define consumption as nondurable goods and services to
avoid the problems of imputing a service flow to the stock of durables (Hall 1978, p. 979).
This definition of consumption is in accord with much of the literature [e.g., Hall (1978),
Flavin (1993), and Campbell and Mankiw (1990)]. The utilization of gross labor income,
capital income net government bonds, and government expenditure follows directly from the
reformulation of permanent income presented in the previous section.

In accord with the literature, each variable is deflated by the personal consumption
expenditure deflator using 1987 as the base year, and each resultant quantity is divided by a
population measure, thus allowing us to express each variable in real per capita terms.
Quarterly U.S. data are used throughout this paper. Data samples start with Quarter 1 1957
(57Q1) and continue through Quarter 4 1993 (93Q4).

All the data in this paper were obtained from U.S. Commerce Department
publications. Data for 1960 through 1991 were obtained from Business Statistics 1961-1991
(U.S. Department of Commerce 1989) and data from 1957 to 1960 are from the 1986 edition
of the same publication. Data after 1991 were taken from various issues of the Survey of
Current Business (U.S. Department of Commerce, various issues). The latest revised values are used in all instances. When necessary, the price deflators were converted so as to have 1987 as the base year (1987 = 100). Page references in the following text refer to Business Statistics 1961-1988.

We now present definitions of the variables noted above:

\( c_i \) is real per capita consumption of nondurable goods and services.

\( y_i \) is a measure of real per-capita gross labor income. This variable includes wage and salary income (defined below), other labor income, and transfer payments. Taxes are not subtracted out.

\( y_{kt} \) represents a form of real per capita capital income. Capital income is attained by subtracting the component of national income generated by labor (wage and salary and other labor income) from total national income. From this amount currently issued government bonds, which are proxied by the federal government deficit/surplus, are subtracted. This net result is \( y_{kt} \), which represents capital income minus government bonds. As shown in the previous section, \( y_{kt} \) is taken to equal \( \frac{R-1}{R} A_1 \).

\( g_i \) is real per capita government expenditures. The components of \( g_i \) are noted later in the text.

\( y_{dt} \) denotes real per capita disposable income, defined later in the text.

Definitions

We now offer definitions of all the components that comprise the series just noted. All of these definitions were obtained from Business Statistics 1961-1988. Page numbers denoted in the following text refer to this publication.

(1) Personal consumption expenditures, total (billion $). Goods and services purchased by individuals; operating expenses of nonprofit institutions; and the value of food, fuel,
clothing, rental of dwellings, and financial services received in kind by individuals. Net purchases of used goods are also included. Purchases of residential structures by individuals and by nonprofit institutions serving individuals are classified as gross private domestic investment because home ownership is treated as a business in the national income and product accounts (p. 303).

(2) National income. Income that originates from the production of the goods and services attributable to the labor and property supplied by U.S. residents. Incomes are recorded as they accrue to residents and are measured before taxes are deducted. Income consist of compensation of employees, proprietor’s income, net interest, corporate profits, and the rental income of persons (p. 303).

(3) Compensation of employees. Relates to income accruing to employees as remuneration for their work. This compensation is the sum of wages and salaries and supplements to wages and salaries (p. 303).

(4) Wages and salaries. Consists of the monetary remuneration of employees, including compensation to corporate officers, commissions, tips, bonuses, and receipts in kind that represent income to the recipients (p. 303).

(5) Supplement to wages and salaries. Employer contributions for social insurance and other labor income comprise this category. Employer contributions for social insurance includes employer payments under the following programs: Federal old age, survivor, disability, and hospital insurance; state and federal unemployment insurance; railroad retirement and unemployment insurance; government employee unemployment insurance and retirement; military medical insurance; and publicly administered workers compensation. Other labor income consists primarily of employer contributions to private pension and welfare funds (p. 303).

(6) Government Expenditure. Covers 5 categories: (a) purchases of goods and services consist of the compensation of government employees, purchases from business and from abroad, and gross investment of government enterprises. (b) Transfer payments consist of transfer payments to U.S. residents and foreigners. Transfer payments to
persons are income payments to persons for which they do not render current services and include payments for programs such as social security; unemployment insurance; medical; and federal civilian, military, and veterans pensions. Transfer payments to foreigners consist of U.S. government nonmilitary grants to foreign governments in cash and in kind, military retirement benefits to former US residents living abroad. (c) Grants-in-aid to state and local governments consists of Federal payments to state and local governments, other than for net interest payments. Major grants-in-aid are for medical, aid to families with dependent children, highways, and education. (d) Net interest paid is interest paid by the government less interest received. (e) Subsidies less the current surplus of government enterprises consist of subsidies such as payments to farmers and the current surplus of government enterprises, calculated by subtracting current operating expenses from current operating income. Federal surplus of deficit is federal receipts less federal expenditures (p. 304).

(7) Price deflator. Implicit price deflator, personal consumption expenditures, and nondurable goods index numbers, 1987 = 100 (p. 289).

(8) Population measure. Total non-institutional population of persons 16 years of age and over (in thousands) (p 43 and p 246).

(9) Transfer payments. Annual estimates of business transfer payments to persons are based on Internal Revenue Service tabulations of business tax returns. Other components, such as liability payments for personal injury, are based on information from other government and trade sources. Annual estimates of Federal government transfer payments are derived largely from the Budget of the United States, Treasury Department data, and agency data for individual programs. State and local government transfer payments are derived from Census Bureau surveys and from reports by Federal agencies that fund certain state and local government programs. Monthly estimates of business transfer payments are based on judgmental trends. Monthly estimates of federal government transfer payments are based largely on monthly Treasury Department data (p. 146).
Disposable Personal Income. Disposable personal income is defined as personal income less personal tax and nontax payments. Personal Income is defined as income received by persons from all sources; that is, from participation in production, from transfer payments from government and business, and from government interest, which is treated like a transfer payment. Personal taxes include income, estate and gift, and personal property taxes (p. 304).

Analysis of the Data

A useful first step in analyzing the data is to plot of the variables $c_t$, $y_t$, $y_k$, $g_t$, and $y_{dt}$. Beyond providing important insights into how these variables evolve over time, such plots also offer suggestions on modeling. We start by considering nondurable consumption, plotted in Figure 3.1.

The consumption series, as seen in Figure 3.1, reveals a distinct upward trend over the sample. Large downward movements occurred in 1970, 1981, and 1990. It is also noted that a simple linear trend does not appear appropriate because the growth rates during the periods 1957-1963, 1965-1971, and 1978-1993 all appear different.

Labor income, as represented in Figure 3.2, displays a smoother plot compared to Figure 3.1. Two large real decreases are evident in 1973 and 1979. It may also be argued that a shift in trend occurred following the first oil crisis. The growth rate after 1979 appears to be lower.

Capital income, presented in Figure 3.3, displays a graph for which a definitive statement on trend is difficult to determine. It is arguable that there exists some upward trend until 1969. The two large real decreases occurring within the years 1973-1974 and 1980-1982, appear related to the recessions following the two oil crises. A large decline is also noted following 1990.
Government expenditure, displayed in Figure 3.4, offers a plot where any fixed notion of trend implies a strong assumption. Distinct changes are noted in the intervals 1957-1967, 1967-1980, and 1985 through the present. The large real increase in government expenditures associated with the Reagan administration is quite evident.

Last, we consider the graph of real disposable income in Figure 3.5, which has a shape similar to that of labor income (Figure 3.2). Note, however, that this plot is less smooth compared to the former. A significant downturn is seen following 1973.
Figure 3.3. Capital income

Figure 3.4. Government expenditure

Figure 3.5. Disposable income
Based on the plots of these series, a decision must be made as to the proper method of modeling these variables. Consumption, labor and disposable income, and government expenditure are all clearly nonstationary. The nonstationarity of capital income is more questionable. Before estimation can proceed, the possibility of nonstationarity must be taken into account. Two possibilities in modeling these data exist, namely, incorporating a trend or differencing the data as a means of accounting for the observed nonstationarity.

One apparent commonality among the relevant data series is our inability to assume a constant growth rate. Albeit, by varying degrees, each plot reveals growth rates that alter over the sample. Thus, it would appear that a deterministic trend model [as used in Flavin (1981) to model consumption and disposable income] is inappropriate because, as Christiano (1987) argues, deterministic trend models embody the unlikely assumption that even after a significant event, such as an oil crisis, the economy will ultimately regain previous growth rates. Recent economic history has shown that for many macroeconomic series this is simply not true. Consider, for example, the plot of gross labor income presented in Figure 3.2, which shows a distinctly lower growth rate following the first oil crisis. A more likely situation is a stochastic trend, where two approaches of incorporating this idea are possible. The first approach uses first differences, which Nelson and Plossar (1982) conclude is the proper representation for most macroeconomic series. The second approach explicitly models the stochastic trend using, for example, a structural time-series method as found in Harvey (1985), Watson (1986), and Clark (1989). This paper adopts the former approach.

Formal tests must be provided before we can conclude that the variables \( \{c_t, y_{ht}, y_{kt}, g_t, \text{ and } y_{dt}\} \) are indeed difference stationary. Dickey-Fuller and augmented Dickey-Fuller tests are performed for each series to ascertain this conjecture. Before proceeding, however, we must determine if these differenced series possess drift, because the existence of drift will affect critical values. Testing for this possibility involves determining the significance (or lack thereof) of the mean of the differenced series, because a significant mean implies drift (Harvey
Drift terms were found to be significant for \( \{c_t, y_{kt}, g_t, y_{dt}\} \), while it appears to be absent for capital income, \( y_{kt} \).

Determination of the appropriate number of lags to be used in the augmented Dickey-Fuller test follows the methodology set forth in Enders (1995). Starting with a model with 12 lags of the differenced variable, the model was pared down by means of an F-test until no further lags could be rejected. Following estimation, the chosen lag length should produce a residual series that lacks any serial correlation. This procedure was repeated separately for each series. Estimation reveals that a lag of 8 is appropriate for \( \Delta c_t, \Delta y_{kt}, \Delta y_{kt}, \) and \( \Delta y_{dt} \). P-values for going from 12 lags to 8 lags were found to be 0.11, 0.08, 0.20, and 0.41, respectively. Even though an insignificant test statistic was found in the case of \( \Delta y_{kt} \), we use a lag of 8 because it accounts for a large correlation that appears at this period. In the case of \( \Delta y_{dt} \), a lag of 8 was used to ensure that the residual series were white noise. A lag of 3 proved to be sufficient in the case of \( \Delta g_t \). Before proceeding, we must check the residuals of these estimated equations to ensure that nothing changed structurally over the sample and to also determine whether the residuals are serially uncorrelated. For this purpose residual plots for equations \( \Delta c_t, \Delta y_{kt}, \Delta y_{kt}, \Delta g_t, \) and \( \Delta y_{dt} \) are presented in Figures 3.6 through 3.10.

![Figure 3.6. \( \Delta c_t \) residuals](image-url)
Figure 3.7. $\Delta y_h$ residuals

Figure 3.8. $\Delta y_{kt}$ residuals

Figure 3.9. $\Delta y_{k}$ residuals
These residual plots suggest the absence of significant structural change over the
sample, though all the plots except \( \Delta c_t \) and \( \Delta g_t \) reveal large changes during the 1970s. Next,
we test for significance of the autocorrelations among each residual series. Under a properly
specified model these correlations would be expected to be zero. To test the hypothesis that
correlations are jointly zero, a Ljung-Box test for various lags was performed for each series.
Description of this test is deferred until the last section of this chapter. A significant test
statistic would imply rejection of autocorrelations being jointly zero up to some specified lag
and thus would be indicative of a poorly specified model. These \( Q \) statistics are provided in
Table 3.1, where p-values are noted in parentheses.

These results indicate that the null hypothesis of the lack of serial correlation cannot be
rejected for each residual series, implying that the specified lag lengths are appropriate. Using
these lag lengths, tests of unit roots can now be performed.

Denoting \( m \) as the appropriate number of lags, the augmented Dickey-Fuller test
(Harvey 1992, pp. 130-132) amounts to estimating the following equation
\[
\Delta x_t = \alpha_0 + \alpha_1 x_{t-1} + \sum_{i=0}^{m} \beta_i \Delta x_{t-i-1}
\]
and comparing the t-statistic on \( \alpha_1 \) to the relevant Dickey-Fuller table of critical values. The
standard Dickey-Fuller test is the special case where \( m = 0 \). Results from these unit roots

Figure 3.10. \( \Delta y_{dt} \) residuals
tests, along with the appropriate number of lags (m), are presented for each variable. Table 3.2 presents results associated with augmented Dickey-Fuller tests, while Table 3.3 presents results using the standard Dickey-Fuller. Critical values for \( y_t \), \( c_t \), and \( g_t \) refer to the \( \tau \) table; i.e., the table that incorporates a drift term (Harvey 1990, p. 368). The critical value for \( y_{kt} \) uses the table with no drift or time trend terms.

<table>
<thead>
<tr>
<th>Lag</th>
<th>( \Delta c_t )</th>
<th>( \Delta y_{kt} )</th>
<th>( \Delta y_{kt} )</th>
<th>( \Delta g_t )</th>
<th>( \Delta y_{kt} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>0.18</td>
<td>0.26</td>
<td>0.37</td>
<td>0.35</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td>(1.0)</td>
<td>(0.99)</td>
<td>(0.99)</td>
<td>(0.99)</td>
<td>(0.99)</td>
</tr>
<tr>
<td>8</td>
<td>0.51</td>
<td>0.54</td>
<td>0.88</td>
<td>2.63</td>
<td>0.56</td>
</tr>
<tr>
<td></td>
<td>(1.0)</td>
<td>(1.0)</td>
<td>(1.0)</td>
<td>(0.96)</td>
<td>(1.0)</td>
</tr>
<tr>
<td>12</td>
<td>4.13</td>
<td>3.75</td>
<td>7.48</td>
<td>5.49</td>
<td>5.06</td>
</tr>
<tr>
<td></td>
<td>(0.98)</td>
<td>(0.99)</td>
<td>(0.83)</td>
<td>(0.94)</td>
<td>(0.96)</td>
</tr>
<tr>
<td>16</td>
<td>4.97</td>
<td>7.61</td>
<td>10.18</td>
<td>7.69</td>
<td>9.42</td>
</tr>
<tr>
<td></td>
<td>(0.99)</td>
<td>(0.96)</td>
<td>(0.86)</td>
<td>(0.96)</td>
<td>(0.90)</td>
</tr>
</tbody>
</table>

Note: p-values appear in parentheses.

<table>
<thead>
<tr>
<th>Series</th>
<th>Sign. Lags</th>
<th>Test Statistic</th>
<th>95% Critical Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( c_t )</td>
<td>8</td>
<td>-2.1</td>
<td>-2.89</td>
</tr>
<tr>
<td>( y_{lt} )</td>
<td>8</td>
<td>-1.69</td>
<td>-2.89</td>
</tr>
<tr>
<td>( y_{kt} )</td>
<td>8</td>
<td>-0.29</td>
<td>-1.95</td>
</tr>
<tr>
<td>( g_t )</td>
<td>3</td>
<td>-2.2</td>
<td>-2.89</td>
</tr>
<tr>
<td>( y_{at} )</td>
<td>8</td>
<td>-1.38</td>
<td>-2.89</td>
</tr>
</tbody>
</table>
Table 3.3. Dickey-Fuller tests: data

<table>
<thead>
<tr>
<th>Series</th>
<th>Test Statistic</th>
<th>95% Critical Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>c_{it}</td>
<td>-2.06</td>
<td>-2.89</td>
</tr>
<tr>
<td>y_{it}</td>
<td>-1.51</td>
<td>-2.89</td>
</tr>
<tr>
<td>y_{kt}</td>
<td>-0.07</td>
<td>-1.95</td>
</tr>
<tr>
<td>g_{it}</td>
<td>-2.06</td>
<td>-2.89</td>
</tr>
<tr>
<td>y_{dt}</td>
<td>-0.91</td>
<td>-2.89</td>
</tr>
</tbody>
</table>

As shown, the null hypothesis of a unit root cannot be rejected at the 95 percent level for any of the five series, using either the standard or the augmented Dickey-Fuller test. These five series are thus assumed to be difference stationary.

Given that each variable is difference stationary, we present plots of the autocorrelation functions (ACF) and the partial autocorrelation functions (PACF) for each variable, $\Delta c_{it}$, $\Delta y_{it}$, $\Delta y_{kt}$, $\Delta g_{it}$, and $\Delta y_{dt}$, in Figures 3.11 through 3.15. These figures will aid in the specification of time-series models for each series.

What is noteworthy for the change in consumption shown in Figure 3.11 are the three large correlations at lags 1 through 3. As will be shown, one theory of consumption posits that the change in consumption should be white noise. A test of this prediction is performed in the succeeding section.
Figure 3.11. ACF $\Delta c_t$

Figure 3.12. ACF $\Delta y_{ht}$
Figure 3.13. ACF $\Delta y_{kt}$

Figure 3.14. ACF $\Delta g_t$
Cointegration Results

Over the past few years, the notion of cointegration has assumed a position of increasing importance in economic modeling and theory. A group of difference stationary variables are said to be cointegrated if some combination of these variables is stationary. This particular combination is the cointegrating vector.

An important result in cointegration analysis is the Granger Representation Theorem (Engle and Granger 1987a), that states that if a group of variables is cointegrated, there exists a valid error correction representation (Cutherbertson, Hall, and Hendry 1992, p. 133). The error correction model may be expressed as

\[(1 - A)(1 - L)\Delta X_t = -\beta \alpha' X_{t-1} + \epsilon_t, \quad \text{or} \]
\[\Delta X_t = A \Delta X_{t-1} - \beta \alpha' X_{t-1} + \epsilon_t, \quad \text{(3.3)} \]

assuming that \(X_t\) is difference stationary and that the change in \(X_t\) can be adequately represented as a VAR(1). The matrix \(\alpha\) denotes the cointegrating vector(s), and \(\beta\) measures the impact with which the cointegrating vector(s) enters into the model. One important
implication of Equation (3.3) is that if cointegration is ignored and $\Delta X_t$ is modeled as a simple VAR(1), the resultant system will be misspecified if the variables $X_t$ are truly cointegrated (Campbell and Perron 1991, p. 130).

Cointegration is a very relevant concept for the work in this paper. Models I and II, presented in Chapter 4, both utilize a VAR system composed of $\{\Delta c_t, \Delta y_t, \Delta y_{kt}, \Delta g_t\}$. If cointegration is present among $\{c_t, y_t, y_{kt}, g_t\}$ a VAR system which ignores the possible cointegration effect will not be properly specified. A VAR system composed solely of first differences is warranted only when cointegration does not exist. A test of whether such a situation is present or not assumes special importance in the work to follow and is carried out below.

We test for cointegration using the maximum likelihood method of estimation found in Johnansen and Juselius (1990). (For the methodology of this procedure, we refer the to that paper.) One element that needs to be specified in implementing this test is the number of lags in the VAR system. To obtain this number, we start with a VAR(8) system and proceed to reduce the number of lags using a likelihood ratio test. Lags are reduced until a significant test statistic is found. The results of this test are presented in Table 3.4.

<table>
<thead>
<tr>
<th>Lag Specification</th>
<th>Test Statistic</th>
<th>Degrees of Freedom</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>VAR(8) → VAR(4)</td>
<td>44.9</td>
<td>64</td>
<td>0.970</td>
</tr>
<tr>
<td>VAR(4) → VAR(3)</td>
<td>12.1</td>
<td>16</td>
<td>0.740</td>
</tr>
<tr>
<td>VAR(3) → VAR(2)</td>
<td>26.6</td>
<td>16</td>
<td>0.046</td>
</tr>
</tbody>
</table>

Because we can reject the reduction from the VAR(3) system to the VAR(2), at the 5 percent level, the appropriate lag length is assumed to be 3. Thus, a VAR(3) is believed to provide an adequate representation of the variables $\{\Delta c_t, \Delta y_t, \Delta y_{kt}, \Delta g_t\}$ and will be used throughout this paper. Having obtained the proper lag specification, we can proceed in testing
for cointegration. Two tests are utilized: the eigenvalue test and the trace test. Test results and the associated null and alternative hypotheses are noted in Tables 3.5 and 3.6.

From the eigenvalue test, the null hypothesis of one or no cointegrating vector is rejected at the 95 percent significance level. However, the null hypothesis of 2 or less cointegrating vectors could not be rejected. The trace test rejects the null hypothesis of no cointegrating vectors at the 95 percent significance level. As with the eigenvalue test, the null hypothesis of two or less such vectors could not be rejected. These results imply that two cointegrating vectors exist among these four variables.

Table 3.5  Eigenvalue test results

<table>
<thead>
<tr>
<th>Null Hypothesis</th>
<th>Alternative Hypothesis</th>
<th>Statistic</th>
<th>95% Critical Value</th>
<th>90% Critical Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>r = 0</td>
<td>r = 1</td>
<td>51.99</td>
<td>28.138</td>
<td>25.559</td>
</tr>
<tr>
<td>r&lt;= 1</td>
<td>r = 2</td>
<td>31.82</td>
<td>22.002</td>
<td>19.766</td>
</tr>
<tr>
<td>r&lt;= 2</td>
<td>r = 3</td>
<td>6.72</td>
<td>15.672</td>
<td>13.752</td>
</tr>
<tr>
<td>r&lt;= 3</td>
<td>r = 4</td>
<td>3.8</td>
<td>9.243</td>
<td>7.525</td>
</tr>
</tbody>
</table>

Table 3.6  Trace test results

<table>
<thead>
<tr>
<th>Null Hypothesis</th>
<th>Alternative Hypothesis</th>
<th>Statistic</th>
<th>95% Critical Value</th>
<th>90% Critical Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>r = 0</td>
<td>r&gt;= 1</td>
<td>74.33</td>
<td>53.116</td>
<td>49.648</td>
</tr>
<tr>
<td>r&lt;= 1</td>
<td>r&gt;= 2</td>
<td>42.34</td>
<td>34.910</td>
<td>32.003</td>
</tr>
<tr>
<td>r&lt;= 2</td>
<td>r&gt;= 3</td>
<td>13.08</td>
<td>10.520</td>
<td>17.852</td>
</tr>
<tr>
<td>r&lt;= 3</td>
<td>r = 4</td>
<td>4.55</td>
<td>3.800</td>
<td>7.525</td>
</tr>
</tbody>
</table>
These two estimated cointegrating vectors can be expressed as:

vector (1): \(-0.0797c_t = -0.042y_t + 0.06207y_{kt} + 0.1544g_t - 0.5953\)

vector (2): \(0.1296c_t = 0.207y_t + 0.1812y_{kt} - 0.2431g_t - 0.794\)

Cointegration implies that the error associated with these vectors is stationary. To determine if vectors (1) and (2) meet this condition, such errors are determined and ACF plots and Dickey-Fuller statistics produced. ACF plots for vectors (1) and (2) are shown in Figures 3.16 and 3.17, respectively. Results of unit root tests are presented in Table 3.6.

The unit root test results and ACF plot for vector (1) suggest that such a combination may not be stationary. Note that the null hypothesis of a unit root could not be rejected at the 5 percent level, regardless of whether the Dickey-Fuller or the augmented Dickey-Fuller test was used. The ACF plot also reveals autocorrelations that decline slowly, indicative of a nonstationary series.

Vector (2), on the other hand, seems to display the necessary stationarity. Dickey-Fuller test results show that the null hypothesis can be comfortably rejected, whereas the ACF plot, Figure 3.17, displays the required rapid decreases in autocorrelations. Owing to the nonstationary errors associated with vector (1), the data series \(\{c_t, y_t, y_{kt}, g_t\}\) are assumed to possess only one cointegrating vector, namely, vector (2). This assumption is made even though the maximum likelihood test of Johnansen and Juselius (1990) suggests two vectors.

![Figure 3.16. ACF vector (1)](image-url)
Figure 3.17. ACF vector (2)

Table 3.6 Unit Root Tests: Cointegrating Vectors

<table>
<thead>
<tr>
<th>Lag Length</th>
<th>Vector (1)</th>
<th>Vector (2)</th>
<th>Signific (95%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DF</td>
<td>-1.6</td>
<td>-3.0</td>
<td>-2.88</td>
</tr>
<tr>
<td>ADF(1)</td>
<td>-1.74</td>
<td>-3.66</td>
<td>-2.88</td>
</tr>
<tr>
<td>ADF(2)</td>
<td>-1.76</td>
<td>-4.07</td>
<td>-2.88</td>
</tr>
<tr>
<td>ADF(3)</td>
<td>-1.66</td>
<td>-4.31</td>
<td>-2.88</td>
</tr>
<tr>
<td>ADF(4)</td>
<td>-1.8</td>
<td>-4.58</td>
<td>-2.88</td>
</tr>
</tbody>
</table>

Testing the PIH

The finding in the previous chapter that consumption is non-orthogonal, excessively sensitive, and excessively smooth proved sufficient in rejecting the PIH as derived by Hall (1978) and Flavin (1981). We now test whether these properties are observed in the data presented in this chapter. Rejection of the PIH using our data would offer rationale for
providing an alternative consumption model that may provide insights into how these properties may arise. To examine this possibility, we consider tests similar to those used by Hall (1978) and Flavin (1981). A weak test of excess smoothness is also conducted.

The first model we consider is by Hall (1978). As discussed in the previous chapter, this model predicts that the change in consumption is orthogonal to all preceding information. Because Hall's theory for the evolution of consumption is so important to consumption theory, we test this hypothesis using the data presented in this chapter. If the current measure of consumption does indeed follow a random walk, there would be little gained from formulating a more complex consumption model, as will be seen in the succeeding chapter.

An indication that consumption does not follow a random walk is provided by the ACF plot for the \( \Delta c_t \) series (Figure 3.12), which displays two significant correlations at lags 1 and 3. To provide a more formal test of whether consumption evolves as a random walk, we consider the following two models for the change in consumption:

(a) \( \Delta c_t = \varepsilon_t \),

(b) \( \Delta c_t = a_1 \Delta c_{t-1} + a_2 \Delta c_{t-2} + a_3 \Delta c_{t-3} + b_1 \Delta y_{d,t-1} + b_2 \Delta y_{d,t-2} + b_3 \Delta y_{d,t-3} + \varepsilon_t \),

where \( c_t \) denotes nondurable goods and services and \( y_{d,t} \) is disposable income, as previously defined. Model (a) embodies Hall's (1978) model and implies that consumption is a simple random walk. Under such a specification, the disturbance term \( \varepsilon_t \) would be expected to be white noise because if this specification of the PIH is correct, all information relevant in formulating \( c_t \) is embodied in \( c_{t-1} \). Differences in consumption would thus be attributed to surprise which, with the assumption of rational expectations, is the white noise series \( \varepsilon_t \). This implication offers an easy method of testing Hall's theory by simply examining whether the disturbance series \( \varepsilon_t \) is indeed white noise.

Model (b) offers an alternative method of testing the random walk hypothesis. If consumption is consistent with theory, we would expect that the coefficients \( a_1, a_2, \ldots, b_4 \) would all be jointly zero because, again, once \( c_{t-1} \) is included into the model no other variables would provide relevant information. We now consider estimating these two models using our data.
In testing model (a), we concentrate on the disturbance series \( e \). To ascertain whether such a series is white noise, Ljung-Box statistics for various lags were computed. These statistics test whether all the autocorrelations up to a given lag are jointly equal to zero. A significant test statistic would indicate failure of the random walk model because the disturbances contain information as implied by nonzero correlations. These statistics were calculated as in Harvey (1990, p. 217); in particular:

\[
Q^* = T(T + 2) \sum_{t=1}^{p} (T - t)r_t^2
\]

where \( r \) denotes sample correlations, \( p \) specifies the lag up to which we are testing and \( T \) denotes the sample size. The statistic \( Q^* \) is distributed as \( \chi^2(p) \). Calculating these statistics for various lags, \( p \), along with the sum of squared residuals (SSR) and the Durbin-Watson statistic, produces the results found in Table 3.7.

### Table 3.7 Model (a): results

<table>
<thead>
<tr>
<th>Sum of Squared Residuals: 0.21603</th>
</tr>
</thead>
<tbody>
<tr>
<td>Durbin-Watson: 1.61</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Number of Lags</th>
<th>Q*</th>
<th>p-value of Q*</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>11.28</td>
<td>0.02</td>
</tr>
<tr>
<td>8</td>
<td>15.93</td>
<td>0.04</td>
</tr>
<tr>
<td>12</td>
<td>19.37</td>
<td>0.08</td>
</tr>
<tr>
<td>16</td>
<td>20.08</td>
<td>0.22</td>
</tr>
</tbody>
</table>

\( Q^* \) statistics appear marginal because at lag 8, for example, the hypothesis of white noise could be rejected at the 5\% level but not at the 1 percent level. The two large correlations at lags 1 and 3 contribute to a significant \( Q^* \) test statistic when the first four lags are utilized, while the Durbin-Watson statistic, which tests for autocorrelation, is significant at
the 1 percent level (Greene 1990, p. 449). Given the large autocorrelation in lags 1 and 3 and a significant Durban-Watson result, we conclude that the residual series $e_i$ is not white noise.

Another test can be performed by estimating model (b) and using an exclusion test noted above. Parameter estimates of this model are reported in Table 3.8, while various diagnostics are reported in Table 3.9.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_1$</td>
<td>0.089</td>
<td>0.32</td>
</tr>
<tr>
<td>$a_2$</td>
<td>0.016</td>
<td>0.86</td>
</tr>
<tr>
<td>$a_3$</td>
<td>0.071</td>
<td>0.40</td>
</tr>
<tr>
<td>$b_1$</td>
<td>0.053</td>
<td>0.02</td>
</tr>
<tr>
<td>$b_2$</td>
<td>0.027</td>
<td>0.24</td>
</tr>
<tr>
<td>$b_3$</td>
<td>0.036</td>
<td>0.12</td>
</tr>
</tbody>
</table>

Table 3.9 Diagnostics

Sum of Squared Residuals: 0.187
Durbin-Watson: 2.01

<table>
<thead>
<tr>
<th>Number of Lags</th>
<th>Q*</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>0.94</td>
<td>0.92</td>
</tr>
<tr>
<td>8</td>
<td>6.16</td>
<td>0.63</td>
</tr>
<tr>
<td>12</td>
<td>9.41</td>
<td>0.67</td>
</tr>
<tr>
<td>16</td>
<td>10.23</td>
<td>0.85</td>
</tr>
</tbody>
</table>
Estimation of model (b) produces superior results, as seen in Table 3.9. $Q^*$ statistics for various lags reveal that the null hypothesis of randomness in the residual series cannot be rejected. Parameter estimates show that only the coefficient associated with $\Delta y_{t-1}$ is significantly different from zero at the 5 percent level. To test for orthogonality, however, a joint test must be performed, for which an F-test of the form

$$ F^0 = \frac{(SSR_a - SSR_b)}{SSR_b / 138} $$

is used. From the results in Table 3.9, it was found that $F^0 = 3.59$, which is distributed with 6,138 degrees of freedom with a p-value of 0.0003. This result suggests that the null hypothesis that all the coefficients in model (b) are jointly zero can be confidently rejected.

Using the fact that the residuals of model (a) are not white noise and given our inability to exclude lagged values of $\Delta y_{t-1}$, along with $\Delta c_{t-2}$ and $\Delta c_{t-3}$ using model (b), we conclude that Hall's (1978) model fails to hold when the data of this chapter is used. Rejection of the random walk specification implies that consumption is dependent upon more than simply its lagged value. The change in consumption is thus not orthogonal to all preceding information.

Next, we consider Flavin (1981), who found that consumption is excessively sensitive to changes in income. As noted in the previous chapter, consumption should not be affected by predictable changes in income. Flavin's test, however, has been criticized because it uses detrended data. We now reformulate Flavin's model by incorporating unit roots and use the data presented in this chapter to test the PIH. Reformulation is important because Flavin (1981) has been widely cited by researchers as offering evidence against the PIH. By estimating such a model using the data of this chapter, we can show whether our data are consistent with such rejection. Further, by explicitly incorporating unit roots, which were noted to be the proper way to model disposable income, the objection of Mankiw and Shapiro (1985) can be avoided and a more powerful test of the PIH obtained. Estimation in this framework will reveal whether the notion of consumption used in this paper possesses the excess sensitivity observed by Flavin (1981).
Formulation of Flavin's (1981) model under the assumption that disposable income and consumption are unit root processes is relatively straightforward. Derivations of the modified system are presented in Appendix A, at the end of this paper. (We relegate these derivations to an appendix because such an exercise is not of central importance to this paper.) Assuming that $\Delta y_{dt}$ is a AR(3) process, these derivations show that the following system needs to be estimated:

$$
(3.4) \quad \Delta y_{dt} = \rho_1 \Delta y_{dt-1} + \rho_2 \Delta y_{dt-2} + \rho_3 \Delta y_{dt-3} + \varepsilon_t
$$

$$
\Delta c_t = (\beta_0 \rho_1 + \beta_1) \Delta y_{dt-1} + (\beta_0 \rho_2 + \beta_2) \Delta y_{dt-2} + (\beta_0 \rho_3) \Delta y_{dt-3} + \varepsilon_t
$$

where $c_t$ is nondurable consumption and $y_t$ denotes disposable income [as used by Flavin (1981)]. Now, the PIH as utilized by Hall (1978) is seen to arise when the parameters $\beta_0$, $\beta_1$, $\beta_2$, and $\beta_3$ are jointly insignificant. Under the null hypothesis $\beta_0 = \beta_1 = \beta_2 = \beta_3 = 0$ it is seen that $\Delta c_t = \varepsilon_t$; that is, Hall's random walk. If the null hypothesis is rejected, we can conclude that the change in consumption is excessively sensitive to changes in income.

Estimation proceeds by imposing the cross-equation restrictions and estimating the system. Note that system (3.4) is just identified, meaning that indirect least squares can be utilized. Parameter estimates are presented in Appendix A. Of importance here is whether the null hypothesis $\beta_0 = \beta_1 = \beta_2 = 0$ can be rejected. To test this hypothesis, we calculate the determinant of the sample variance/covariance matrix under both the null and alternative hypotheses and calculate a likelihood ratio test, which is distributed as $\chi^2(3)$ (Harvey 1990 pp. 162-166). In particular

$$
LR = 138[\ln(0.7612) - \ln(0.696)] = 12.54,
$$

with a p-value of 0.007, suggesting that we can confidently reject the null hypothesis that $\beta_0 = \beta_1 = \beta_2 = 0$ and thus conclude that excess sensitivity does indeed exist.

These results suggest that the PIH as currently formulated is incorrect. As discussed, one implication of the PIH, as formulated by Hall (1978) and Flavin (1981), is that the change in consumption should be adequately modeled with a random walk. In testing this supposition, however, we found that other lagged variables (disposable income) do indeed contain
information. That is, consumption is more sensitive ("excessively sensitive") to predictable changes in income than is warranted by the PIH. Explanations as to why consumption may display this property are discussed in Chapter 4.

Last, we consider the question of excess smoothness. Consumption was noted to be excessively smooth in the previous chapter because the variance of consumption (as predicted using the PIH and a difference stationary time-series model of income) was found to be greater than the variance of income. Assuming that the PIH is correct, the variance associated with consumption should equal the variance of permanent income, both of which should be less than the variance of income. Only a very simple view of this question will be addressed because our sole intent is to show that the PIH as it is currently formulated displays such smoothness.

We start by considering standard deviations of consumption and disposable income using the data presented herein. These results are presented in Table 3.10.

<table>
<thead>
<tr>
<th>Table 3.10 Comparing Standard Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Series</td>
</tr>
<tr>
<td>Nondurable Consumption</td>
</tr>
<tr>
<td>Disposable Income</td>
</tr>
<tr>
<td>Change in Nondurable Consumption</td>
</tr>
<tr>
<td>Change in Disposable Income</td>
</tr>
</tbody>
</table>

Whether levels or first differences are considered, the consumption series is less volatile than disposable income. Now, following Campbell and Deaton (1989), assume that labor income can be adequately modeled as a AR(1). Estimating this model produces

\[ \Delta y_t = 0.063 + 0.163 \Delta y_{t-1} + \varepsilon_t, \quad \text{Var}(\varepsilon_t) = 0.0112. \]

With this time-series representation of the change in income, we can use results from the previous chapter to examine what such a process implies about the variance of permanent
income. Using the results derived from Campbell and Deaton, and presented in Chapter 3, it can be shown that

\[ \text{Var}(\Delta c_t) = \text{Var} \text{(permanent income)} = 0.018 > \text{Var}(\Delta y_t) = \text{Var}(\varepsilon_t) = 0.0112. \]

Thus, assuming a difference stationary process for labor income, the variance of the change in consumption is greater than that associated with income. Again, this result contradicts the theory of the PIH as well as the stylized facts, which show exactly the opposite.

The results in this section reveal that the data used in this paper possess properties similar to those observed in previous studies. In the following chapter, we propose a model which incorporate market imperfections and displays behavior quite consistent with that discussed above. That is, the model to be presented generates results which show that consumption is non-orthogonal, excessively sensitive, and excessively smooth, suggesting that the model is consistent with the stylized facts previously noted.
CHAPTER 4: COSTLY ADJUSTMENT

Introduction

We now consider a possible reformulation of the PIH, given its apparent rejection both in the literature and using the data of Chapter 3. As discussed, causes for failure are numerous. Possible explanations range from the use of a quadratic utility function, to stochastic interest rates, to capital market imperfections. This paper focuses on market imperfections, which will be incorporated into the determination of consumption through the costs that such phenomena may generate. These costs, introduced in Chapter 2, appear to be a very natural way in which market imperfections may be modeled as compared with previous approaches. Unlike previous attempts, which requires either extraordinary assumptions about the agent's understanding of the economy or the division of agents into constrained and unconstrained groups, we simply posit that the agent responds to the costs associated with market failures.

To model this costly dimension of market imperfections as they relate to the determination of consumption, this chapter breaks the discussion into five sections. The following section, entitled "An Alternative Approach: Costly Adjustment," discusses one method by which market imperfections may be incorporated into the determination of consumption. We model imperfections by assuming that a dichotomy may exist among the formulation and implementation of a consumption plan. Thus, although we assume that the agent operates in an optimizing environment, implementation of consumption plans based on such maximization may be hampered by reality. If agents are unable to borrow sufficient amounts and encounter market rigidities, which may be caused by imperfect markets, their
desired levels of consumption may be unattainable. These desired levels are those predicted by the PIH.

To incorporate these market imperfections, while remaining in the PIH framework, this paper studies the costs that these imperfections may present to the agent. Costs may arise from two sources: First is with the inability to achieve desired consumption levels, while a second occurs as market rigidities preclude instantaneous change in consumption levels. Two examples of incorporating adjustment costs are also considered.

The section, entitled “Model I”, provides two methods by which adjustment costs may be incorporated into consumption modeling. Although these two methods are formed on very different bases, they are shown to provide an equivalent estimable function for the change in consumption. The first approach considered uses a quadratic loss function to model the two costs that may be incurred while formulating an optimal consumption path. This function presents costs to the agent in a very amenable manner. The expected change in permanent income is crucial in this discussion and is considered in the subsection, “Determining $E_{t-1}\Delta y_{Pt}$."

A second method, presented in the subsection “Alternative Derivation,” approaches the adjustment question in terms of the utility function where costs of adjustment affect agents directly through the utility function. Regardless of which theoretical basis is assumed, the change in consumption is shown as a function of the lagged change in consumption and the expected change in permanent income. An important question in this analysis is whether the derived model displays results consistent with consumption determination as seen in Chapters 2 and 3. In particular, we consider whether the costly adjustment model displays non-orthogonality, excess sensitivity, and excess smoothness. This material is contained in a subsection entitled “Consumption Properties.”

Estimation proceeds under the assumption of rational expectations, where an auxiliary system is specified and joint estimation performed. The assumption of rational expectations, however, necessitates the imposition of formidable cross-equation restrictions. Explicit tests are used in determining the validity of the costly adjustment framework. Following estimation the appropriateness of the model is explored by analyzing various diagnostics. One test which
will prove important in the following chapter examines whether the assumption of time-invariant parameters is warranted.

Specifying Model I requires a high level of analytical skill on the part of the agent as revealed by the cross-equation restrictions, implied by the rational expectations hypothesis. To consider an alternative specification, a section entitled "Model II," offers a simplified version of Model I. Model II differs from Model I in the level of knowledge the agent is assumed to possess. Rather than assuming that the agent "knows" the complete system as Model I presupposes, the milder assumption is made that the agent uses some simple linear rule to determine expected permanent income. The use of a simple rule rather than the more complex structure implied by the rational expectations hypothesis provides a precursor to Chapter 5, where a type of learning model is considered. Following specification of the rule, Model II is estimated and explicit tests are performed to determine the legitimacy of the costly adjustment framework. Various diagnostics are used to ascertain the validity and properties of the model. A test of whether parameters are time-invariant is also performed.

Cointegration is the subject of the subsequent section. As noted in Chapter 3, the variables \( \{c_t, y_{kt}, y_{kt}, g_t\} \) appear to possess one cointegrating vector. Models using first differences of these variables may be misspecified if this cointegrating vector is ignored. To avoid charges of misspecification, cointegration is explicitly incorporated into revised versions of Models I and II. As shown in this section, even though such an addition is statistically significant, nothing of substance changes in terms of the acceptance or rejection of the costly adjustment framework.

The final two sections of this chapter present various results associated with the costly adjustment framework. The sixth section examines policy implications given the existence of the costly adjustment. Of primary importance is the possibility of government policies affecting current consumption levels. The last section summarizes the most significant conclusions that arise from estimating Models I and II. These conclusions provide insight into how consumption may best be estimated.
An Alternative Approach: Costly Adjustment

Chapters 2 and 3 reveal problems with the manner in which consumption is currently modeled. Papers reviewed in Chapter 2 and tests using our data in Chapter 3 suggest that consumption does not evolve in accord with the PIH. One explanation for this empirical failure is the existence of capital market imperfections. Imperfections, by violating some of the assumptions upon which the PIH is based, was shown in Chapter 2 to produce consumption behavior consistent with reality, namely, that current consumption relies upon more than simply its lagged level. Consider the results of Flavin (1993), for example, which showed that consumption was non-orthogonal, excessively sensitive, and excessively smooth.

However, the present method of incorporating market failures into consumption modeling is believed to be lacking. As discussed in Chapter 2, two approaches have been used to model imperfections. The first explicitly models an imperfection that was deemed relevant, while the second assumes that some segment of the population is liquidity constrained. Both approaches, however, were shown to make strong assumptions upon the agent and the economy. For example, the former assumes that agents can uniquely identify all imperfections that may affect them, while the latter posits that the economy can be divided among constrained and unconstrained agents, with imperfections affecting only the constrained agents. We now introduce a method of incorporating market failures which reduces the assumptions imposed upon the agent. The alternative developed in the following text does not require explicit knowledge of particular market imperfections, nor does it incorporate the unlikely assumption that the majority of the population is completely unaffected by market failures. Instead, our alternative method allows market imperfections to affect all agents, albeit by varying degrees. All market failures are thus permitted to influence consumption, not just those failures that initiate liquidity constraints.
In this section, we introduce an alternative approach to incorporating market imperfections that addresses the criticisms of previous attempts noted in Chapter 2. Imperfections presently, are modeled by the costs which these phenomena may generate, costs which the rational agent will take into account when formulating a consumption plan. Thus, we assume that agents are not reacting to the imperfection per se, but rather to the costs that such failures may generate.

Costs seem a natural summary expression for any imperfection that may confront the consumer. Imperfections that are most relevant to consumption, such as credit rationing, discrimination, transaction costs, search, and deviations between the lending and borrowing rates, all have the effect of increasing the cost of consumption. More severe manifestations will generally result in higher costs. Consider, for example, an individual who encounters discrimination and is thus only able to borrow funds at a rate higher than the fair market rate. To this consumer, the visible results of such discrimination are higher interest payments that must be made. It is these costs that this individual will be reacting to when determining the appropriate consumption path. By modeling the effect and not the cause, a more appropriate manner of incorporating imperfections into a representative consumption model may be achieved. Speight (1990, p. 161) takes costly to mean that: "... assets can only be realized or made "liquid" at a value lower than their "illiquid" value; that it is not possible to borrow at the same interest rate as applies when lending; or that it is not possible to secure debt on terms comparable with those on existing borrowing." It is precisely this notion that we use.

The costly adjustment framework differs fundamentally from the second approach considered in the proceeding section. Rather than assuming that imperfections are selective with regard to whom they affect, we take a more general view and allow market imperfections to affect all agents. Note, however, that we are not presupposing that all agents are affected equally. It may be that the costs associated with consumption are higher for one segment of the population compared to another. In this manner, no arbitrary decision must be made as to whether some group of consumers are liquidity constrained. The more realistic case that market imperfections affect all consumers can then be modeled.
This approach differs from the "liquidity constrained" approach discussed earlier, whereby one segment of the population remains unaffected by market imperfections. Being unconstrained this group would consume according to the PIH. What such an approach fails to realize is that market failures other than those that contribute to the notion of liquidity constraints may exist in the economy. A more realistic approach, and the one taken in this paper, is to allow all market imperfections to affect the determination of consumption. By incorporating the costs that such failures may generate and permitting these costs to affect all agents, no arbitrary decision must be made as what constitutes market failure or whether or not some segment is affected by such phenomena.

It is, of course, an assumption that the agent responds to imperfections by incorporating the associated costs into the decision-making process. However, because imperfections tend to make the attainment of a desired level of consumption more costly, whether in real or psychological terms, it seems plausible that the rational consumer would take such costs into account. When formulating a consumption path, the agent will attempt to maximize utility while simultaneously attempting to incur as little additional cost as possible.

Under the PIH with all its accompanying assumptions, the agent will be able to reach desired levels of consumption with no excessive costs. As mentioned, the absence of excessive costs exists for several reasons. First, it is assumed that the individual can borrow any amount at a constant rate, r. Further, it is assumed that there are no rigidities or discrimination in the market. Last, equality between lending and borrowing rates is assumed, which implies that the present value of the loan is 0. Allowing the lending rate to differ from the borrowing rate leads to excessive cost because the present value of any loan is no longer 0.

The subject that forms the basis for the remainder of this paper is how best to incorporate this notion of costs into the PIH framework. We choose to remain within the PIH framework because many elements that comprise the hypothesis make intuitive sense. For example, the idea that the agent uses expected future income in determining consumption seems very plausible. Empirical shortcomings of the PIH may have less to do with the formation of some desired level of consumption than with the ability of the agent to achieve
that level. In this situation, remaining within the PIH framework is appropriate because
deriving consumption based on some notion of permanent income seems quite reasonable.
Beyond this, as shown in the following section, augmenting such costs to the PIH will nest the
standard PIH in a more general costly adjustment model, allowing us to test the standard PIH
framework relative to a more complex framework.

Costs that may confront the agent arise from two sources. First, costs can arise from
the inability of the agent to attain some desired level of consumption. This concept represents
the idea that market imperfections are in essence restrictions that prevent the achievement of
desired consumption levels (Flavin 1985, p. 118). As in any optimization problem, the
addition of (binding) restrictions precludes the attainment of the unconstrained value which, in
this case, is the level of consumption predicted by the PIH. An implicit assumption is being
made, however, in that the PIH will hold in the absence of all market imperfections and that
consumers will use this level as a proxy for desired consumption. Desired consumption may
thus be defined as that level which would have been selected in the absence of any restrictions.

By discussing deviations between actual and desired levels, we explicitly allow for a
dichotomy between the formation and implementation of a consumption plan as formed under
the PIH. If all the assumptions upon which the PIH is built are upheld, formulation and
implementation are the same, because the agent is able to consume at the level predicted by
the theory. Currently, we allow for the possibility that such an assumption may not be upheld.
In this case, even though agents may formulate consumption in accord with the PIH, they may
be unable to implement such plans. The inability to borrow sufficient funds to achieve desired
consumption levels is one example of the failure to implement some course of action.

Deviations from desired levels of consumption are deemed to be costly because the
agent is forced to consume quantities that would not have been accepted in the absence of
market failures. These deviations may generate both financial and physiological costs as
viewed by the agent. The inability of the agent to achieve desired consumption levels may
force the consumer to break habits or incur discomfort from altering consumption plans.
Financial costs may occur as desired levels of consumption are realized to be unattainable.
Altering consumption plans may present costs to the agent in the form of breaking contracts, possibly at a cost, losses sustained when precommitted consumption and investment plans must be changed, or possible search costs as the agent strives to reconstruct a consumption plan.

Existence of this cost depends crucially on some notion of desired consumption. In this paper, desired consumption is assumed to be represented by the level of consumption predicted by the PIH. Although based on an assumption, this definition seems appropriate because we are working in the PIH framework that predicts this consumption level if all assumptions are met. Because we are interested in ascertaining the importance of market imperfections, at least with regard to the PIH, this definition appears reasonable because it represents the case where market imperfections do not exist. The inability to achieve desired levels of consumption within the PIH framework, suggests the existence of market failures. By assuming that the agent reacts to the costs generated by such deviations, we are able to model the impact of market failures while leaving the causal imperfection(s) unspecified.

Because rational agents will attempt to minimize costs, the existence of deviations between desired and actual consumption will induce the agent to act toward eliminating such a gap. Such action, however, generates a second type of cost, related to altering consumption levels. Whereas the agent may freely change consumption levels with impunity under the PIH, the existence of market imperfections may limit this ability. Imperfections may translate into rigidities that make altering consumption levels costly.

Market failures such as credit rationing, imperfect information, and transaction costs may all contribute to making instantaneous change in the level of consumption impossible or extremely unpleasant. With the existence of these phenomena, the agent may be forced to adjust over a number of periods, making abrupt change impossible or subject to large financial costs. Examples of possible costs include exorbitant interest rates associated with credit card purchases or the necessity of liquidating assets at a loss as the agent attempts to finance an expansion in consumption. Based on these costs, the agent will undertake a change in consumption that incurs the least possible additional cost.
Confronted with these costs, the agent will attempt to formulate a consumption path. One cost will induce the agent to consume near the desired level, while the second will inhibit any movement to eliminate any such deviations. The final consumption path will be that which "balances" these two opposing costs so that actual consumption generates the least possible total additional cost. By viewing the magnitude and significance of these two costs, we are able to infer the existence (or lack thereof) of market imperfections while incorporating such phenomena into the determination of consumption in a very tractable manner.

The following section expresses these notions of costs in a reasonable manner. In particular, the costs are incorporated into the determination of consumption through the use of a quadratic loss function. The first term in such a function will represent the first cost just discussed, namely, that relating to deviations between actual and desired consumption. An additional term incorporates the second cost, or that associated with altering consumption levels. Minimizing this function relative to consumption will produce a consumption path that minimizes expected costs. Before proceeding to this material, however, we examine two papers that apply the notion of costly adjustment in the determination of consumption. These papers reveal how costly adjustment has been applied and what may be expected when operating within such an framework.

Examples of Costly Adjustment

We first consider a paper by Cushing (1992), who contends that costs associated with rapidly changing consumption could provide rationale for Deaton's (1987) result that consumption is excessively smooth (p. 145). To test whether costly adjustment may explain the rejection of the PIH, Cushing specifies the following utility function:

$$u(c_i) = u_o + u_1c_i - (u_2 / 2)c_i^2 - (u_3 / 2)(c_i - c_{i-1})^2 + c_i a_i,$$

where $a_i$ denotes a mean zero, serially uncorrelated preference shock. This utility function is the standard quadratic utility function to which the term $(u_3 / 2)(c_i - c_{i-1})^2$ has been
augmented. The addition of this term reflects the possibility that changes in the level of consumption may have an adverse effect on utility.

Allowing for the existence of constrained and unconstrained segments of the population, Cushing shows that the change in consumption can be expressed as

$$\Delta c_t = (1 - \mu)\lambda_o + \gamma \Delta c_{t-1} + \mu \pi (\Delta y_t + \gamma \Delta y_{t-1}) + (1 - \mu)\epsilon_t,$$

where $\gamma$ reflects the penalty for adjusting consumption and is bounded between 0 and 1. The disturbance term, $\epsilon_t$, is a moving average 1 [MA(1)] error term, because Cushing allows for preference shocks, $\pi$ denotes the fraction of total income accounted for by the liquidity-constrained portion of the population, and $\mu$ represents the fraction of consumption accounted for by liquidity-constrained individuals.

Following estimation, Cushing concludes that $\gamma$ is not significantly different from zero, implying that in this particular specification the costly adjustment model can be rejected (see pp. 146-147 for details). It is noted however, that Cushing's model differs fundamentally from the model we are espousing. In his work, Cushing models only the cost associated with altering consumption levels (the second cost noted above), and ignores the cost arising from deviations between actual and desired consumption levels.

We now consider a paper by Attfield, Demery, and Duck (1992) that is more in keeping with the notion of costs as laid forth in this section. Attfield, Demery, and Duck (1992) posit that the consumer minimizes the function

$$L = E_t \left\{ \sum_{i=0}^{\infty} p_t^i \left[ (1/2)(b - c_t)^2 + (\alpha/2)(c_{t+i} - c_{t+i-1})^2 \right] \right\},$$

subject to the constraint

$$\sum_{i=0}^{\infty} p_t^i c_{t+i} = (1 + r)A_t + \sum_{i=0}^{\infty} p_t^i y_{t+i},$$

where $c_t$ is nondurable consumption, $y_t$ is disposable labor income, $A_t$ is real beginning period wealth, $p = 1/(1 + r)$, $b$ is the "bliss point" and $\alpha$ is a cost of adjustment parameter. Equation (4.1) presents the notion of costs used in this section, namely, that there are costs associated with being away from "bliss" (or desired consumption), which is noted to be fixed over the
sample, and costs associated with adjusting to this consumption level (p. 1209). Without going into detail, the authors derive the following model for the change in consumption:

\[ \Delta c_t = (1 + r)(1 - \theta)\Delta c_{t-1} + \theta \sum_{i=0}^{\infty} (E_i - E_{t-1}) \Delta y_{t+i}, \]

where \( \theta \) is a function of the real interest rate \( r \) and the stable root associated with the solution of Equations (4.1) and (4.2). If the adjustment parameter \( \alpha \) is zero (i.e., if adjustment is costless), then \( \theta = 1 \); otherwise \( \theta < 1 \).

Equation (4.3) implies that once the lagged change in consumption is controlled for, the error term \( \theta \sum_{i=0}^{\infty} (E_i - E_{t-1}) \Delta y_{t+i} \) should be white noise, under the null hypothesis of costly adjustment. If added to Equation (4.3), lagged variables should be statistically insignificant (p. 1209).

Attfield, Demery, and Duck test Equation (4.3) by analyzing the residual series produced with estimation and by adding terms \( \Delta c_{t-2}, \ldots, \Delta c_{t-5}, \Delta y_{t-1}, \ldots, \Delta y_{t-5} \) to test for orthogonality. In the case of the United States the null hypothesis that the residual series is white noise can be rejected at the 5 percent level, but not at the 7 percent level. Orthogonality can be rejected at the 5 percent level, but not at the 6 percent level. Thus, the costly adjustment model seems marginal, at least at the 5 percent level of significance.

Incorporation of the costs utilized by this paper, which is the subject of the succeeding sections, follows a form similar to that used by Attfield, Demery, and Duck (1992). Estimation presents costs in a very intuitive and plausible manner. As will be seen, however, Cushing's (1992) model can be obtained as a special case of a more general model. The work of Attfield, Demery, and Duck is extended in the following work by allowing the level of consumption predicted by the PIH \( (y_{\mu}) \) to represent bliss, which we call desired consumption. This approach is more general, because we explicitly allow the notion of desired consumption to alter over the sample, as compared to the time-invariant representation utilizing Equation (4.1). Incorporating a changing notion of desired consumption is believed to be important because it is unduly restrictive to assume that such a quantity will be fixed over the sample. In
an evolving economy an agent's notion of desired consumption would be expected to change, reflecting an altering environment. As shown in the following section, however, allowing for such change introduces more complexity into the system. An additional extension to the model presented by Attfield, Demery, and Duck, is to impose no constraint on the term \((b - c_{t+1})\), appearing in Equation (4.1). Again, given that we use a notion of \(b\) that changes each period, the two alterations previously noted above provide a costly adjustment model that is more flexible than that used in Attfield, Demery, and Duck.

We now turn to our presentation of a costly adjustment model, which forms the basis of the following section. Although the model is similar to that utilized by Attfield, Demery, and Duck important differences exist. Estimation will show whether our more general representation of costly adjustment produces appreciably different results. The model that we discuss next permits explicit statements to be made about the two costs of adjustment and, unlike Attfield, Demery, and Duck, strong evidence for the existence of the two costs of adjustment.

Model I

As discussed in the previous section, market imperfections may be modeled in terms of the costs that these phenomena may generate. The first cost considered was that associated with the inability to consume at desired levels because of restrictions imposed by the market may prohibit the agent from consuming what was planned. The second cost arises when the agent attempts to alter consumption levels to alleviate this first cost, for rigidities may be introduced by market failures thus ruling out instantaneous change. We now consider the precise nature of these costs and how they may be incorporated into the optimal consumption problem. We begin by defining desired consumption as that level the consumer would choose in the absence of all imperfections. Assuming quadratic utility, the PIH as formulated by Hall
(1978) and Flavin (1981), specifies the optimal level of consumption as $c^*_t = y_{pt}$, where $y_{pt}$ denotes permanent income. Because we are working in a PIH framework the level of permanent income seems to be a reasonable quantity with which to proxy the notion of desired consumption.

Using this level of desired consumption as proxied by the level of permanent income ($y_{pt}$), the agent will attempt to realize this consumption goal. Confronted with an uncertain economy, it is implicitly assumed that the agent will use the notion of desired consumption as the basis of a consumption plan and undertake action toward its attainment. It is only when the agent attempts to achieve desired levels, however, that the imperfect nature of the market is seen. As mentioned, these imperfections have the effect of limiting the ability of the agent to consume desired levels, thus frustrating prior plans. Deviation between actual and desired levels of consumption may generate tangible costs that the agent must take into account when formulating an actual, or achievable, consumption path. Costs may be incurred if the agent accepts this desired level as the quantity he wishes to consume and undertakes action, whether real or physiological, toward its attainment. Inability to achieve the desired level may leave the consumer with financial liabilities or a heightened sense of uncertainty through the forcible breaking of habits. In either case, market imperfections are seen to frustrate consumption plans that were formed under the assumption that the PIH offers the best prediction of consumption in an uncertain environment.

One possible objection to using this definition of desired consumption is that the agent is assumed to use continuously the PIH as providing the correct measure, even though the actual level of consumption obtained shows that market imperfections may exist. That is, the agent never learns to incorporate imperfections into the determination of desired consumption. Though this objection is legitimate, we choose to ignore it for two reasons. First, by assuming that the agent uses the PIH to determine desired consumption, we do not need to ascertain a priori what imperfections may be present in the system. As discussed in the previous chapter, much complexity can be avoided by attempting to model the general notion of imperfections rather than modeling specific manifestations. Further, by incorporating this notion of desired
consumption, a test of the PIH can be carried out in terms of a more general model, which is discussed later. A second reason is that the agent may not be fully cognizant of all imperfections when consumption plans are made. Reality is only encountered when the agent attempts to implement his desired consumption plan. With uncertainty in the system, the level predicted by the PIH may represent a best guess in projecting a consumption path.

With the existence of deviation between actual and desired consumption and its subsequent cost comes an incentive to alleviate it. However, a second cost is incurred as the agent acts to eliminate the first. This second cost, denoted as a cost of adjustment, inhibits the agent from attempting to eliminate the difference between actual and desired consumption too abruptly. As mentioned, altering consumption levels should not be assumed to be frictionless; rigidities in the market may make large changes prohibitively expensive. Changes in consumption may thus be spread out so as to incur as little additional cost as possible.

With the existence of these costs, the agent will attempt to formulate optimal consumption. Optimization will smooth consumption sufficiently so that the costs associated with deviations between actual and desired are balanced with the costs of adjustment. The procedure that we use for this purpose is sufficiently flexible so that the special case that no such costs exist is nested within a more general model. Specifically, we use a quadratic cost function in which to incorporate the two costs. Use of this type of function implicitly assumes that the agent will choose consumption levels that minimize these costs.

Quadratic loss functions are quite common in the macroeconomic literature, examples of which include Cuthbertson and Taylor (1987, 1992) and Sargent (1989). As will be shown, adjustment costs appear in a reasonable and amicable manner. The quadratic loss function used in this section, denoted as Model I, can be expressed as

\[
(4.4) \quad \Psi = E_t \left\{ \sum_{t+1} \delta^t \left[ \alpha_0 (c_{t+i} - y_{t+i})^2 + \alpha_1 (c_{t+i} - c_{t+i-1})^2 \right] \right\}.
\]

where $\delta$ is a discount factor ($0 < \delta < 1$) that measures the rate of the agent's time preference and $\alpha_0, \alpha_1 \geq 0$. Permanent income, $y_{t+i}$, is used to represent desired consumption at period $t+i$. The first term in Equation (4.4), $(c_{t+i} - y_{t+i})$, reflects the possibility that the agent may
not be able to reach desired levels as proxied by \( y_{p+i} \). Because this squared deviation is weighted by the nonnegative cost term, \( \alpha_0 \), any deviations will increase or leave unchanged, \( \Psi \). Actual consumption greater than or less than desired levels is all treated alike, reflecting the belief that any deviation from desired consumption will force the agent to reformulate consumption plans. The second term, \( (c_{t+i} - c_{t+i-1}) \), describes the change in consumption over subsequent periods with \( \alpha_i \) providing a measure of the cost associated with this type of change. Abrupt movements in the level of consumption are thus seen to increase \( \Psi \). The cost terms \( \alpha_0 \) and \( \alpha_i \) reveal how the agent behaves in this particular model. Although these terms are not individually identified, the ratio of \( \alpha_0/\alpha_i \) is specified and provides sufficient information to comment upon the behavior of the agent in this particular model.

Using this function, \( \Psi \), and given the information available at time \( t \), the agent is assumed to determine an actual consumption path that provides an expected minimum to the function. Intuitively, given the nonnegative costs \( \alpha_0 \) and \( \alpha_i \), we seek to avoid deviations from desired consumption levels and from previous consumption levels. Minimizing such a function will provide a consumption path where these expected deviations are balanced. It should be noted, however, that even though this notion of quadratic costs is infinitely summed, the introduction of the discount factor \( D \) (0 < \( D < 1 \)) depreciates the importance of expected consumption levels far into the future.

Following Cuthbertson and Taylor (1987), the solution to Equation (4.4) can be expressed as

\[
(4.5) \quad \Delta c_t = \lambda_1 \Delta c_{t-1} + (1 - \lambda_1)(1 - \lambda_1 D) \sum (\lambda_i D)^i E_t y_{p+i},
\]

where \( \lambda_1 \) denotes the stable root of Equation (4.4). As seen in Cuthbertson and Taylor (1987) and Sargent (1989), the roots of Equation (4.4) have the properties that

\[
0 < \lambda_1 < 1 < D^{-1} < \lambda_2 \quad \text{and} \quad \lambda_2^{-1} = D \lambda_1, \quad -\left(\frac{\alpha_0}{\alpha_i} + 1 + D\right)D^{-1} = -\left(\lambda_1 + \lambda_2\right)
\]

Letting \( \lambda = \frac{\alpha_0}{\alpha_i} \), Cuthbertson and Taylor show that
Equation (4.6) provides insight into how the costs $\alpha_0$ and $\alpha_1$ operate in this framework. From this expression it is seen that when the cost associated with deviations from desired consumption ($y_{pt}$) are very large compared to the cost of adjustment the value of $\lambda$ approaches positive infinity, implying that $\lambda_i \to 0$. Using Equation (4.5), this result in turn implies that $\Delta c_i = E_i \Delta y_{pt}$. (How this expression simplifies will be seen below.) Alternatively, when the cost of adjustment is very large compared to deviations from desired consumption, the quantity $\lambda$ approaches zero implying that $\lambda_i \to 1$, revealing through Equation (4.5) that $\Delta c_i = \Delta c_{i-1}$.

These two special cases show the reasonable manner in which the notion of costly adjustment is incorporated into the consumer’s decision process. Intuitively, as the cost associated with deviations from $y_{pt}$ increases relative to the cost of adjustment, the consumer will naturally attempt to consume at the level predicted by the PIH. Alternatively, if the costs associated with altering consumption levels are large compared to deviations between actual and desired consumption, the agent will avoid changing consumption levels and will thus consume at the last period’s level.

As will be seen, the two special cases, $\lambda_i \to 0$ and $\lambda_i \to 1$, may be tested. In particular, the PIH is nested within this more general framework and arises when $\lambda_i \to 0$. Further, a variant of the cost of adjustment model suggested by Cushing (1992) is also seen to be nested within Equation (4.5) and occurs when $\lambda_i \to 1$. In the case of $0 < \lambda_i < 1$, Equation (4.5) must be modified for estimation, which we now turn to.

Transforming Equation (4.5) into an estimable form necessitates analysis of the infinite sum $\sum (\lambda_i D)^i E_i y_{pt+i}$. Under the PIH as specified in Hall (1978) and Flavin (1981), this sum is simply $(1 - \lambda_i D)^{-1} y_{pt}$, because $E_i y_{pt} = y_{pt-1}$. Hall (1978) used this property, discussed in Flavin (1981), to arrive at his random walk hypothesis. Because this property is important in
the work to follow we reproduce a portion of Flavin’s work. In particular, Flavin (1981) showed that
\begin{equation}
E_t y_{pt+1} = (1 + r)y_{pt} - rc_t.
\end{equation}

Our slightly different notion of permanent income produces an analogous result, namely, that
\begin{equation}
E_{t-1} y_{pt+1} = Ry_{pt} - (R - 1)c_t.
\end{equation}
But since \( R = 1 + r \), the two forms are seen to be the same.

Assuming that the PIH is correct and that there is no transitory component to consumption, consumption in period \( t \), \( c_t \), is equal to \( y_{pt} \). Thus
\begin{equation}
E_t y_{pt+1} = (1 + r)y_{pt} - ry_{pt} = y_{pt}.
\end{equation}

More generally, \( E_t y_{pt+j} = y_{pt} \ \forall \ j > 0 \) implying that when the PIH is correct, or equivalently when \( \lambda_1 \to 0 \) in Equation (4.5), the infinite expected sum \( \sum (\lambda_1 D)^i E_t y_{pt+i} \) is equal to \( y_{pt} \).

In the situation where \( \lambda_1 \neq 0 \); that is, when the PIH fails to hold, this simple representation for \( \sum (\lambda_1 D)^i E_t y_{pt+i} \) does not hold, necessitating a tractable method by which the infinite expected sum may be determined. To proceed in this direction, an assumption must be made as to how \( c_t \) evolves over time. In particular, it is assumed that consumption evolves according to
\begin{equation}
c_t = \lambda_1 c_{t-1} + (1 - \lambda_1) y_{pt}.
\end{equation}

Attfield, Demery, and Duck (1992) specify a similar expression for the evolution of consumption. Equation (4.8) assumes that current consumption is a function of lagged consumption and current permanent income where it is noted that when \( \lambda_1 \to 0 \) the agent consumes according to the PIH. By assuming that current consumption can be represented by Equation (4.8), we have the basis of a method with which to solve the infinite expected sum
\begin{equation}
\sum (\lambda_1 D)^i E_t y_{pt+i}.
\end{equation}

We start by considering Equation (4.4) for all future time periods; that is,
\begin{equation}
E_t y_{pt+j} = (1 + r)y_{pt+j-1} - rc_{t+j-1} \quad j = 1, 2, 3, \ldots,
\end{equation}
and use Equation (4.8) to substitute for \( c_{t+j-1} \). We then determine, recursively, Equation (4.9) for all future time periods. Consider, for example, the first few recursions noted below:
\[ E_t y_{pt+1} = (1 + r)y_{pt} - r c_t \]  

[Equation (4.7)]

\[ = (1 + r)y_{pt} - r\lambda_1 c_{t-1} - r(1 - \lambda_1) y_{pt} \]

\[ = \Gamma y_{pt} - r\lambda_1 c_{t-1} \]

where \( \Gamma = [(1 + r) - r(1 - \lambda_1)]. \)

\[ E_t y_{pt+2} = (1 + r)E_t y_{pt+1} + r E_t c_{t+1} \]

\[ = (1 + r)E_t y_{pt+1} - r\lambda_1 c_t - r(1 - \lambda_1) E_t y_{pt+1} \]

\[ = \Gamma E_t y_{pt+1} - r\lambda_1 (\lambda_1 c_{t-1} + (1 - \lambda_1) y_{pt}) \]

\[ = \left[ \Gamma^2 - r\lambda_1 (1 - \lambda_1) \right] y_{pt} - \left[ \Gamma r\lambda_1 + r\lambda_1^2 \right] c_{t-1}. \]

Continuing these recursions allows us to determine all expected future levels of permanent income. An important point to be noted in these recursions for \( E_t y_{pt+j} \) is that the resultant expected quantity can be expressed solely in terms of \( y_{pt} \) and \( c_{t-1} \) for each time period. That is, each expected future level of permanent income is found to be some function of \( y_{pt} \) and \( c_{t-1} \). With these expected levels, derived in the manner just shown, an expression for the infinite expected sum can be formed. Using the expected levels as discussed,

\[ \sum (\lambda_1 D)^i E_t y_{pt+i} \text{ can be represented as some function of current permanent income and lagged consumption because the infinite expected sum of the expected future levels of } y_{pt+i} \text{ are shown to be a function of } y_{pt} \text{ and } c_{t-1}. \]

This relationship can be expressed as

\[ (4.10) \sum (\lambda_1 D)^i E_t y_{pt+i} = g_2(\lambda_1, D, r) y_{pt} - g_1(\lambda_1, D, r) c_{t-1}. \]

where \( g_1(.) \) and \( g_2(.) \) denote some coefficients that depend on \( \lambda_1, D, \) and \( r \) in some complicated fashion.
We now make a short digression to discuss the properties of $g_1(.)$ and $g_2(.)$. Because $g_1$ and $g_2$ represent complicated infinite sums in which the quantity $(1 + r)^t, r > 0, r > 0$, assumes an important role, the question of convergence must be posed. Given that $0 \leq \lambda < 1$, certain conditions must be imposed to ensure that $g_1$ and $g_2$ are finite quantities. In particular, when $\lambda \to 0$ it can be shown that $g_1 = 0$ and $g_2 = 1$. In the case $\lambda \to 1$, however, we must impose the condition that $\frac{1}{D} < (1 + r)$; that is, the time preference of the agent must be strictly less than the real rate of interest. With this condition, it can be shown that $g_1 = B_1$ and $g_2 = B_2$, where $B_1$ and $B_2$ denote some finite positive numbers. In the general case of $0 < \lambda < 1$, we are guaranteed that $g_1$ and $g_2$ are finite and lie in the intervals $(0, B_1)$ and $(1, B_2)$, respectively.

Differencing Equation (4.10) results in the following expression:

$$(4.11) \quad \sum (\lambda, D)^t E_t \Delta y_{pt} = g_2(\lambda, d, r) \Delta y_{pt} - g_1(\lambda, d, r) \Delta c_{t-1}.$$ 

Substituting this expression into Equation (4.5) allows for the change in consumption to evolve as

$$\Delta c_t = \lambda, \Delta c_{t-1} + (1 - \lambda)(1 - \lambda, D)[g_2(\cdot) \Delta y_{pt} - g_1(\cdot) \Delta c_{t-1}]$$

or, defining $q_1 = \lambda, - (1 - \lambda)(1 - \lambda, D)g_1(.) \quad q_2 = (1 - \lambda)(1 - \lambda, D)g_2(.)$.

$$(4.12) \quad \Delta c_t = q_1 \Delta c_{t-1} + q_2 \Delta y_{pt}.$$ 

Adding and subtracting $q_2 E_{t-1} y_{pt}$ to the right-hand side of Equation (4.12) yields

$$\Delta c_t = q_1 \Delta c_{t-1} + q_2 \Delta y_{pt} \pm q_2 E_{t-1} y_{pt},$$

$$\Delta c_t = q_1 \Delta c_{t-1} + q_2 E_{t-1} \Delta y_{pt} + q_2 (E_t - E_{t-1}) y_{pt}.$$ 

With the assumption of rational expectations, the last term may be considered a white noise disturbance term; that is,

$$\Delta c_t = q_1 \Delta c_{t-1} + q_2 E_{t-1} \Delta y_{pt} + \omega_t, \quad \text{where} \quad \omega_t = q_2 (E_t - E_{t-1}) y_{pt}.$$ 

Equation (4.13) is now liable for estimation. Two results should be noted, however, with regard to Equation (4.13). The first deals with the parameters $q_1$ and $q_2$. From its constituent elements, $q_1$ may be either negative or positive, depending on the magnitude of $\lambda$. The parameter $q_2$, however, will always be strictly positive. Second, is that $\lambda_1$ is not identified.
which, is again a byproduct of the method used to simplify $\sum (\lambda, D)^i E_t y_{t+i}$. Estimation of Equation (4.13) would thus only provide estimates of $q_1$ and $q_2$, that depend in some complicated manner on $\lambda_i$. Though unfortunate, this inability to obtain point estimates of $\lambda_i$ is not really important. What is important to this chapter is whether or not the costly adjustment model can be rejected. As discussed later, rejection of this model is contingent on whether $\lambda_i$ is found to be 0. If it can be shown that $\lambda_i \neq 0$, we have sufficient information with which to conclude that the change in consumption must be modeled as a function of the change in permanent income and, possibly, the lagged change in consumption, a conclusion which can be reached even though the precise value of $\lambda_i$ remains unknown. Information can also be obtained by showing that $\lambda_i \neq 1$ because we would be able to reject the notion of costly adjustment presented by Cushing (1992). It remains to be shown, however, that the cases $\lambda_i \to 0$ and $\lambda_i \to 1$ are identified, which we now turn to.

We first consider the case when $\lambda_i \to 0$. Using this condition Equation (4.8) simplifies to $c_t = y_p$. When this notion of consumption is used in the recursive solutions for Equation (4.9), $E_t y_{t+j}$, $j = 1,2,\ldots$, the following will be observed:

$E_t y_{t+1} = (1+r)y_p - r c_t$

$= (1+r)y_p - r(\lambda_i c_{t-1} + (1-\lambda_i)y_p)$

$= (1+r)y_p - ry_p = y_p$.

$E_t y_{t+2} = (1+r)E_t y_{t+1} - rE_t c_{t+1}$

$= (1+r)y_p - r(E_t y_{t+1})$

$= y_p$.

$E_t y_{t+j} = y_p \quad \forall j \geq 1$

Thus, in the case $\lambda_i \to 0$, $E_t \Delta y_p = E_t y_p - y_{p-1} = 0$, implying that Equation (4.13) simplifies to
Equation (4.14) shows that when \( \lambda_1 \to 0 \) consumption evolves as a random walk, as predicted by Hall (1978).

Consider now the case when \( \lambda_1 \to 1 \) where, even though the term \( E_{1,0} \Delta y_{1,t} \neq 0 \), it is multiplied by the parameter \( q_2 \) which is zero, as seen in the derivations leading to Equation (4.12), implying that

\[
\begin{align*}
(4.15) \quad \Delta c_t &= \Delta c_{t-1} + \bar{\omega}_t.
\end{align*}
\]

This result is a variant of the model used in Gushing (1992). (Gushing, however, does not constrain the coefficient on \( \Delta c_{t-1} \) to be 1.)

With the two special cases already noted, Equation (4.13) can be seen to represent the situation \( 0 < \lambda_1 < 1 \) (i.e., the "general" model.) When \( \lambda_1 \to 0 \), Equation (4.14) is obtained, and Equation (4.15) results when \( \lambda_1 \to 1 \). Again, even though \( \lambda_1 \) is not identified in the general model, a test of the costly adjustment model is still possible and would proceed by estimating the model given, \( 0 < \lambda_1 < 1 \) [Equation (4.13)] and also assuming \( \lambda_1 \to 0 \) and \( \lambda_1 \to 1 \), and comparing results using a likelihood ratio test as described later. These tests will provide a definitive statement as to whether \( \lambda_1 \) is in the interval (0, 1) or approaches 0 or 1. Rejection of the hypothesis \( \lambda_1 \to 0 \) provides evidence against the PIH, while failure to accept the notion that \( \lambda_1 \to 1 \) presents evidence against a pure costly adjustment model. By accepting the general model we have sufficient information to conclude that the two types of costs discussed in this chapter have a significant impact on the determination of optimal consumption. Such a conclusion can be reached even though the precise value of \( \lambda_1 \) remains unknown.

Equation (4.13) also makes an implicit assumption that \( \lambda_1 \), \( D \), and \( r \) are time-invariant. With regard to the real interest rate, \( r \), this assumption is dubious. Few would believe in a real rate that has remained constant over the span of the sample. A similar point arises with regards to \( \lambda_1 \). As previously discussed, \( \lambda_1 \) is intimately related to \( D \) and \( \lambda \) (\( \lambda = \alpha_0/\alpha_1 \)). Temporal movements in \( D \) or the two cost parameters will initiate changes in \( \lambda_1 \). It would
appear likely that \( \lambda \), the ratio of costs, would undergo significant change over the sample. Possible reasons for this change include an evolving financial system, deregulation, recessions, and possible equal opportunity legislation. All these situations may cause \( \lambda \) to vary over the sample. If it is believed that \( \lambda \), \( D \), and \( r \) may alter over time, one would expect the parameters in Equation (4.13), by being a function of these variables, also to be time-variant. To examine this possibility, a test will be performed, following estimation, to determine whether the constancy of \( q_1 \) and \( q_2 \) is warranted. Rejection of constancy would suggest that estimation should proceed with time-varying parameters, because explicit movements in these parameters could be modeled. Chapter 5 addresses modeling in this particular framework.

Prior to the estimation of Equation (4.13), the quantity \( E_{t-1} \Delta y_{pt} \) must be evaluated, at least for the general case. As discussed, when \( \lambda \to 0 \) this quantity is simply zero. Under the more general condition \( 0 < \lambda \leq 1 \), this equality no longer holds. The following subsection presents two possible formulations for this quantity.

**Determining \( E_{t-1} \Delta y_{pt} \)**

Estimation of Equation (4.13) requires the replacement of the term \( E_{t-1} \Delta y_{pt} \), which necessitates an explicit formulation of permanent income. In this chapter and the following chapter, an alternative form of permanent income discussed in a section of Chapter 3. In particular, rather than using disposable labor income in deriving permanent income, we use gross labor income and government expenditures. Following Chapter 3, permanent income is defined as

\[
y_{pt} = E_t \left( r \rho \sum \rho^i y_{ki+i} + y_{ki} - r \rho \sum \rho^i g_{ti+i} \right),
\]

where \( r \) is the real interest rate and \( \rho \) is \([1/(1+r)]\), \( y_{ki} \) denotes gross labor income, \( y_{ki} \) is capital income net government bonds, and \( g_{ti} \) denotes government expenditures. Differencing this expression produces
\[ \Delta y_{pt} = \left[ r \sum p^i (E, y_{ht+1} - E, y_{ht+1}) + \Delta y_{kt} - r \sum p^i (E, g_{rt+1} - E, g_{rt+1}) \right] \]

The expected value, with respect to information at \((t-1)\), of the change in permanent income can be expressed as

\[(4.16) \quad E_{t-1} \Delta y_{pt} = E_{t-1} \left( r \sum p^i \Delta y_{ht+1} + \Delta y_{kt} - r \sum p^i \Delta g_{rt+1} \right) \]

Again, this expression can be used in the general case when \(0 < \lambda_t \leq 1\), for in the special case \(\lambda_t \rightarrow 0\), \(E_{t-1} \Delta y_{pt}\) is simply zero. Assuming that \(\{\Delta y_{ht}, \Delta y_{kt}, \Delta g_{rt}\}\) follow some VAR process, \(E_{t-1} \Delta y_{pt}\) can be modeled as some combination of lagged values of these variables, specification of which is presented next.

However, this is not the only way \(E_{t-1} \Delta y_{pt}\) may be modeled. Attfield, Demery, and Duck (1992) proxy \(\Delta y_{pt}\) differently. In their model, PIH the optimal level of consumption is characterized as

\[(4.17) \quad c_t = \theta y_{pt} + (1 - \theta)c_{t-1}, \]

where \(y_{pt}\) is permanent income and \(\theta\) denotes an adjustment parameter \((0 \leq \theta \leq 1)\). Differencing this expression leads to

\[(4.18) \quad \Delta c_t = \theta \Delta y_{pt} + (1 - \theta) \Delta c_{t-1}. \]

Equation (4.18) is of a form similar to Equation (4.12). In Appendix B of Attfield, Demery, and Duck (1992), the authors show that the change in permanent income can be characterized as

\[(4.19) \quad \Delta y_{pt} = r[y_{pt-1} - c_{t-1}] + \omega_t, \]

where \(\omega_t\) is an innovation which is white noise under the assumption of rational expectations and given that their model PIH is correct; that is, given that \(\Delta c_{t-1}\) is controlled for (p. 1209). (For a derivation of this expression the reader is referred to Appendix B of Attfield, Demery, and Duck.) Substituting Equation (4.19) into Equation (4.18) allows the change in consumption to evolve as

\[(4.20) \quad \Delta c_t = (1 + r)(1 - \theta) \Delta c_{t-1} + \theta \omega_t. \]
Testing the legitimacy of their model PIH1, Attfield, Demery, and Duck proceed by estimating Equation (4.20) and examining the residual series. If the model is correct, the residual series should be white noise with $\Delta c_t$ being orthogonal to all information, except $\Delta c_{t-1}$, in time $t-1$.

We choose to model $E_{t-1}\Delta y_{pt}$ by Equation (4.16) rather than using the method of Attfield, Demery, and Duck (1992) because potential problems arise when the latter method is applied to our Model I. To show this problem, we consider an application of their method to our model. We start by defining $\Delta y_{pt}$ according to Equation (4.19), where the innovation term must be redefined so as to be consistent with the expanded notion of permanent income used in this chapter. In particular, we define $\omega_t$ as an innovation which reflects the labor income and government expenditure series as discussed. Next, rather than using Equation (4.17) to substitute for $y_{pt-1}$, we use the expression $c_{t-1} = q_1 c_{t-2} + q_2 y_{pt-1}$, which corresponds to Equation (4.12) in levels. Solving for $y_{pt-1}$ and substituting into Equation (4.19) allows $\Delta y_{pt}$ to be expressed as

$$
\Delta y_{pt} = r \left( \frac{1}{q_2} c_{t-1} - \frac{q_1}{q_2} c_{t-2} - c_{t-1} \right) + \bar{\omega}_t,
$$

which is the notion of the change in $y_{pt}$ when applying the method of Attfield, Demery, and Duck to our model. Substituting this expression of $\Delta y_{pt}$ into Equation (4.12), which is

$$
\Delta c_t = q_1 \Delta c_{t-1} + q_2 \Delta y_{pt},
$$

allows the change in consumption to be expressed as

$$
\Delta c_t = q_1 \Delta c_{t-1} + r q_2 \left[ c_{t-1} - q_1 c_{t-2} - q_2 c_{t-1} \right] + q_2 \bar{\omega}_t,
$$

or,

(4.21) $\Delta c_t = q_1 \Delta c_{t-1} + r q_2 \left[ (1 - q_2) c_{t-1} - q_1 c_{t-2} \right] + q_2 \bar{\omega}_t.$

Note that the expression for the change in consumption found in Attfield, Demery, and Duck, Equation (4.20), does not hold in the context of our Model I. A more serious problem with Equation (4.21) is seen by examining the terms on the right-hand side. As discussed in Chapter 3, consumption is a difference stationary series, implying that $\Delta c_t$ on the left-hand side is stationary, as is the first term on the right-hand side, $\Delta c_{t-1}$. The problem arises with the bracketed term $[(1-q_2)c_{t-1} - q_1 c_{t-2}]$. Unless $(1-q_2) = q_1$, can we be assured that this bracketed
term is stationary. This is an important point because if this equality fails to hold, the regression (4.21) would contain a nonstationary regressor, in which case the disturbance term \( \bar{\omega}_t \) would not necessarily be expected to be white noise, as posited by Attfield, Demery, and Duck in their model PIH1.

With the inability to state a priori whether or not \((1 - \theta_2) = \theta_1\), we assume a more general stance in providing a proxy for \(\Delta y_{pt}\). For the work to follow, Equation (4.16) is used to provide an estimate for \(E_{t-1}\Delta y_{pt}\); that is,

\[
E_{t-1}\Delta y_{pt} = E_{t-1}\left(\theta_p \sum \theta^i \Delta y_{k+i} + \Delta y_{k1} - \theta_p \sum \theta^i \Delta g_{i+1}\right)
\]

Beyond avoiding the potential problems associated with estimating Equation (4.21), characterizing \(E_{t-1}\Delta y_{pt}\) with Equation (4.16) allows for the direct incorporation of information relating to the series \(\{\Delta y_{k1}, \Delta y_{k2}, \Delta g_{i}\}\). It is believed that this additional information will present a more interesting model for studying the change in consumption. Specifying \(E_{t-1}\Delta y_{pt}\) with Equation (4.16) also permits an extension, presented in the following chapter, whereby coefficients in the auxiliary system are allowed to evolve over the sample. Such an analysis permits the expectation-generating mechanism to change, thus reflecting an altering economy. Defining \(E_{t-1}\Delta y_{pt}\) according to Attfield, Demery, and Duck would not permit such an extension. If parameters do change, specifying \(E_{t-1}\Delta y_{pt}\) according to Attfield, Demery, and Duck is not correct because derivation relies upon the implicit assumption of constant coefficients. Because Chapter 5 utilizes time-varying parameters, Equation (4.16) appears to be the most appropriate method of modeling \(E_{t-1}\Delta y_{pt}\).

Using this expression for \(E_{t-1}\Delta y_{pt}\), we can now present a model for Equation (4.13). We start by specifying a VAR system for \(\{\Delta y_{k1}, \Delta y_{k2}, \Delta g_{i}\}\), which will be used in generating expectations. In determining the lag specification for this VAR system we recall the results from Chapter 3. Starting with a VAR(8) system and reducing the number of lags with a likelihood ratio test, lag 3 appeared to be that lag for which no further reduction could be obtained. Thus, a VAR(3) will be used in modeling the auxiliary system comprised of the
variables $\{\Delta y_{it}, \Delta y_{kt}, \Delta g_{iti}\}$. Defining $Z_{2t}$ as $\{\Delta y_{it}, \Delta y_{kt}, \Delta g_{iti}\}$, this VAR(3) process can be represented as

$$Z_{2t} = C_1 Z_{2t-1} + C_2 Z_{2t-2} + C_3 Z_{2t-3} + \eta_t,$$

where the $C_i$'s denote a $(3 \times 3)$ coefficient matrices and $\eta_t = \{\varepsilon^t_i, \varepsilon^y_i, \varepsilon^k_i, \varepsilon^g_i\}$.

Throughout this chapter we assume rational expectations, which necessitates the joint estimation of Equation (4.13) and system (4.22). As will be shown this assumption also imposes formidable cross-equation restrictions. The joint system may be specified as

$$Z_t = B_1 Z_{t-1} + B_2 Z_{t-2} + B_3 Z_{t-3} + \varepsilon_t,$$

where $Z_{t+i} = \{\Delta c_{s+1}, \Delta y_{s+1}, \Delta y_{s+1}, \Delta g_{s+1}\}, \ i = 0, 1, 2, 3$.

The $B_i$'s denote appropriately defined coefficient matrices, and $\varepsilon_t = \{\varepsilon^c_t, \varepsilon^y_t, \varepsilon^k_t, \varepsilon^g_t\}$. We include $\Delta c_{t-2}$ and $\Delta c_{t-3}$ in the variable set so that a redefinition of system (4.23) can be made, however, all parameters associated with these two variables are taken to be zero. These restrictions are necessary for identifying certain key parameters in the model.

Imposition of the cross-equation restrictions is more readily applied by transforming system (4.23) into a VAR(1) system. Such a transformation, which puts the model into a “state-space” form, allows the model to be expressed as

$$
\begin{bmatrix}
Z_t \\
Z_{t-1} \\
Z_{t-2}
\end{bmatrix} =
\begin{bmatrix}
B_1 & B_2 & B_3 \\
I_4 & 0_4 & 0_4 \\
0_4 & I_4 & 0_4
\end{bmatrix}
\begin{bmatrix}
Z_{t-1} \\
Z_{t-2} \\
Z_{t-3}
\end{bmatrix} +
\begin{bmatrix}
\varepsilon_t \\
0_{4x1} \\
0_{4x1}
\end{bmatrix},
$$

or,

$$X_t = AX_{t-1} + \varepsilon_t,$$

with $X_t = \{Z_t, Z_{t-1}, Z_{t-2}\}'$.

Thus, $Z_t = [I_4 \quad 0_4 \quad 0_4] AX_{t-1} + \varepsilon_t$.

With the new system expressed by Equation (4.25), the infinite expected sums found in Equation (4.16);

$$E_{t-1}\left(\rho \sum \rho^t \Delta y_{s+1} + \Delta y_{kt} - \rho \sum \rho^s \Delta g_{s+1}\right),$$
can be determined. Following Keating (1990), and assuming that the coefficient matrix $A$ is nonsingular, these infinite expected sums can be represented as

$$E_{t-1} \left( r \sum p^i \Delta y_{it+1} - r \sum p^i \Delta g_{it+1} \right) = r p \begin{pmatrix} 0 & 1 & 0 & -1 & 0 \end{pmatrix} A (I - p A)^{-1} X_{t-1}.$$ 

Using Equation (4.25), the change in capital income can be expressed as

$$E_{t-1} \Delta y_{kt} = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \end{pmatrix} A X_{t-1}.$$ 

Using these expressions for the components of $E_{t-1} \Delta y_p$, Equation (4.13) can be expressed as

(4.26) \[ \Delta x = \begin{pmatrix} 1 & 0 & 0 & 0 \end{pmatrix} X_{t-1} \]

Using Equation (4.26), the change in capital income can be expressed as

$$E_{t-1} \Delta y_{kt} = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \end{pmatrix} A X_{t-1}.$$ 

Equation (4.26) shows how the change in consumption evolves in this particular model. Note that the assumption of rational expectations presupposes a high degree of structure on the system, the extent of which is found in the restrictions that Equation (4.26) places on the coefficients $a_{11}, a_{12}, \ldots, a_{110}$, which are elements of the coefficient matrix $A$. Derivation of these restrictions is relegated to Appendix B at the end of the paper.

Restrictions imposed upon $a_{12}, a_{13}, \ldots, a_{110}$ are highly nonlinear and are functions of the coefficients $b_{11}, c_{31}, d_{41}, q_1$, and $q_2$. The model of the change in consumption as implied by Model 1 entails estimation of the system

(4.27) \[ \begin{bmatrix} \Delta c_t \\ \Delta y_h \\ \Delta y_{kt} \\ \Delta g_t \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} & a_{17} & a_{18} & a_{19} & a_{110} \\ 0 & b_{22} & b_{23} & b_{24} & b_{25} & b_{26} & b_{27} & b_{28} & b_{29} & b_{210} \\ 0 & c_{32} & c_{33} & c_{34} & c_{35} & c_{36} & c_{37} & c_{38} & c_{39} & c_{310} \\ 0 & d_{42} & d_{43} & d_{44} & d_{45} & d_{46} & d_{47} & d_{48} & d_{49} & d_{410} \end{bmatrix} \begin{bmatrix} Z_{t-1} \\ Z_{t-2} \\ Z_{t-3} \end{bmatrix} + \begin{bmatrix} \varepsilon_{ct} \\ \varepsilon_{yh} \\ \varepsilon_{ykt} \end{bmatrix}, \]

where $a_{11}, a_{12}, \ldots, a_{110}$ reflect the cross-equation restrictions implied by the model, and $Z_{t-1}$, $Z_{t-2}$, and $Z_{t-3}$ are as defined previously (deleting all elements corresponding to $\Delta c_{t-1}$ and $\Delta c_{t-1}$). Identifying $a_{11}$ necessitates the additional restriction that $b_{21} = c_{31} = d_{41} = 0$, which implies that the change in consumption does not Granger-cause the change in labor and capital income, or the change in government expenditure. As seen in Appendix B, imposition of these restrictions implying that $a_{11} = q_1$. Estimation of this system proceeds by imposing all the restrictions
discussed above and obtaining those parameter values \((\bar{a}_{11}, \bar{b}_{22}, \bar{b}_{23}, ..., \bar{c}_{32}, \bar{c}_{33}, ..., \bar{d}_{410})\) that minimize the determinant of the sample covariance matrix. A variable metric minimization routine with numerical derivatives, as developed in Nash (1979), was used for this purpose.

Using system (4.27), we consider the special cases where \(\lambda_1 \to 0\) and \(\lambda_1 \to 1\). As seen from Appendix B, the case \(\lambda_1 \to 0\) implies that \(a_{11} = a_{12} = ... = a_{110} = 0\); that is, consumption follows a random walk. Further, as seen in Appendix B, \(\lambda_1 \to 1\) also implies certain restrictions on \(a_{11}, a_{12}, ..., a_{110}\). By forming these special cases, the costly adjustment model can be tested.

We start by determining the determinant of the sum of squared residuals (SSR) matrix assuming the general (unrestricted) model [i.e., \(\nu_{we} = \left| \sum \varepsilon \varepsilon' \right|\) where \(\varepsilon\) denotes the residual series obtained from estimating system (4.27)]. Next, we can compute the determinant of the SSR matrices under each of the special cases, \(\nu_{\lambda_1 = 0}\) and \(\nu_{\lambda_1 = 1}\). With these quantities, a likelihood ratio test of the form

\[ L = T \left[ \ln(\nu_{\lambda_1 = 0}) - \ln(\nu_{we}) \right] \]

can be computed where \(T\) denotes the sample size (Harvey 1990, pp. 162-166). \(L\) is approximately distributed as \(\chi^2(m)\), where \(m\) denote the number of restrictions. Presently, \(m = 1\) because in deriving the restricted model only one restriction is imposed, namely, \(\lambda_1 \to 0\) (or \(\lambda_1 \to 1\)). A similar test can be used to test the null hypothesis that \(\lambda_1 \to 1\). Rejection of both null hypotheses would provide support for the costly adjustment model. Estimation of this model, along with tests of the two special cases, are presented below. Before turning to such results, however, we examine another possible method of incorporating adjustment costs into consumption determination. The model presented in the following subsection is shown to yield an observational equivalent estimation equation like that utilized above, even though the latter is formed on a very different basis.
Alternative Derivation

We now consider an alternative approach to incorporating adjustment costs which shows that consumption as expressed by Equation (4.13) is not to be uniquely attributable quadratic loss function approach. Specifically, we consider the possibility of incorporating costs of adjustment enter through the utility function. We pursue this material because it is one method by which costly adjustment has been applied to consumption modeling. Papers by Bernanke (1984, 1985), and Cushing (1992) have used the utility function when incorporating costs of adjustment. However, the notions of costs used currently is considerably broader than those used previously. Compared to Cushing (1992) who considers only the costs of altering consumption levels, we explicitly model the effects of the agent being unable to obtain desired levels of consumption. By using this utility function framework within which to present our notion of costly adjustment we are able to make direct comparisons to other studies in this area. Of primary importance in this subsection is that the functional form of the change in consumption is the same whether formulated with the quadratic loss function or a utility function.

In this subsection, we assume that the agent utilizes a utility function of the form

\[ u(c_t) = u_0 + u_1 c_t - \frac{u_2}{2} c_t^2 - \frac{u_3}{2} (c_t - y^*)^2 - \frac{u_4}{2} (c_t - c_{t-1})^2. \]

Note that this utility function, with \( u_3, u_4 \geq 0 \), is a simple quadratic expression to which two additional terms have been added. The first addition, which enters with the marginal utility \( u_3 \), describes the effect that deviations from desired consumption may have on utility. As in the preceding work, we have assumed that desired consumption can be adequately proxied by \( y^* \). It will be shown below that \( c_t = y^* \) only when \( u_3 = u_4 = 0 \). When this equality fails to hold, deviations between actual and desired consumption will have a depressing impact on utility, possibly reflecting the frustration of plans or the forcible breaking of habits. Disutility associated with deviations between actual and desired consumption may arise from the
distaste of broken habits or the dislike of having to reformulate an alternative consumption path, given the inability of the agent to achieve what was desired.

Bernanke (1985) allows for a similar expression; however, a time-invariant (constant) notion of a “bliss point” is used. By representing this bliss point, which we will termed the desired level of consumption, by the quantity $y_{pt}$, we in essence permit this quantity to vary over the sample, offering a very different interpretation as compared to that used by Bernanke (1985). In an evolving economic system, allowing such a quantity to vary over time may be important in providing a more realistic characterization of consumption determination.

Equation (4.28) also incorporates a term which represents the disutility associated with altering consumption plans. Disutility in this instance, measured by $-u_4$, may arise from the displeasure of search (e.g., product research, searching for lowest borrowing rates, etc.) or the need to reformulate consumption levels. This term is the more conventional of the two additions made to the standard utility function. An example of a penalty assigned to altering the level of consumption is found in Cushing (1992).

To formulate the optimal path of consumption implied by this model, we consider the solution of the first-order conditions. Solving reveals that

$$\Delta c_t = \gamma_1 \Delta y_{pt} + \gamma_2 \Delta c_{t-1} \quad \text{with} \quad \gamma_1 = \frac{u_1}{u_2 + u_3 + u_4} \quad \text{and} \quad \gamma_2 = \frac{u_4}{u_2 + u_3 + u_4}. \quad (4.29)$$

Using this expression, we can show that $c_t = y_{pt}$ when $u_1 = u_4 = 0$, which implies that $\Delta c_t = 0$, or $c_t = c_{t-1}$. Substituting this expression into the intertemporal budget constraint from Chapter 3 shows that $c_t = y_{pt}$. Given that $u_3, u_4 \geq 0$, deviations from desired consumption and the previous level of consumption are seen to reduce utility. Utility maximization using such a function would produce a consumption path for which these deviations are balanced.

By adding and subtracting $\gamma_1 E_{t-1} y_{pt}$ to Equation (4.29), the following expression can be obtained:

$$\Delta c_t = \gamma_1 E_{t-1} \Delta y_{pt} + \gamma_2 \Delta c_{t-1} + \omega_t \quad \text{where} \quad \omega_t = \gamma_1 (E_t - E_{t-1}) y_{pt}. \quad (4.30)$$
The special cases previously discussed may also be considered in terms of the utility function approach. Using Equation (4.29), we can study the special case \( \lambda_1 \to 0 \), discussed in relation to the quadratic loss model, which arises when the costs associated with deviations between actual and desired consumption are much greater compared to the costs of altering consumption levels. This special case \( \lambda_1 \to 0 \) must be reinterpreted in terms of the utility function model. We start by considering Equation (4.29) when the marginal utilities associated with altering consumption levels and deviations between actual and desired consumption approach zero. Through Equation (4.29), this situation implies that \( \Delta c_t = 0 \), or \( c_t = c_t \).

Substituting this result into the intertemporal budget constraint of Equation (2.5), shows that \( c_t = y_p \); that is, the agent consumes according to the PIH. Not surprisingly, the situation \( u_3, u_4 \to 0 \) shows that the agent will behave as in Hall's (1978) model.

An analogous situation holds when \( u_3 \to \infty \) with \( u_4 \) being nonzero; that is, the case in which the marginal utility associated with deviations between actual and desired consumption levels approaches infinity. In this case, using Equation (4.29) we see that \( \gamma_1 \to 1 \) and \( \gamma_2 \to 0 \), suggesting that \( c_t = y_p \), again in accord with the PIH. Thus, in the utility function approach, the case in which consumption evolves according to the PIH arises in two situations: \( u_3, u_4 \to 0 \) and \( u_3 \to \infty \) with \( u_4 \) being nonzero, and produces an equivalent estimational form for \( \Delta c_t \) that occurs in the quadratic loss approach under the special case \( \lambda_1 \to 0 \).

Next, we consider the equivalent form of the special case \( \lambda_1 \to 1 \) in the utility function approach. From the discussion of the quadratic loss function, the case \( \lambda_1 \to 1 \) may be interpreted as arising when the costs associated with altering consumption levels is much greater than those arising from deviations between actual and desired levels. Currently, this idea can be incorporated in the utility function derivations when \( u_4 \to \infty \), and \( u_3 \) is some nonzero finite quantity, or when \( u_3 \to 0 \) and \( u_4 \) is nonzero. In the situation where \( u_3 \to 0 \); that is, when costs associated with altering levels do not exist, the term \( \gamma_2 \) in Equation (4.23) will be some term which in general will not equal 1. Thus, the special case \( \lambda_1 \to 1 \) in the utility function approach is slightly different from that used in deriving Equation (4.13) because the
coefficient on $\Delta c_{i,1}$ is not constrained to be 1. However, if $\lambda_2 \to \infty$, the coefficient $\gamma_2$ will approach 1, suggesting that Equation (4.29) will evolve as Equation (4.13) under the case $\lambda_1 \to 1$, because $\gamma_1 \to 0$ while $\gamma_2 \to 1$.

In the general case of $\lambda_1$ and $\lambda_2$ both not being equal to zero, Equation (4.30) provides a model for the changes in consumption that is equivalent to Equation (4.13). This result implies that in the general case in which both costs of adjustment are present, the costly adjustment model can be derived either from the quadratic loss approach or through the utility function. Both Equation (4.13) and Equation (4.30) show that in the case of the general model the change in consumption is a function of other variables. Equivalence between Equations (4.13) and (4.30), at least for the general case, shows that the quadratic loss function is not unique in generating the change in consumption that is a function of the lagged change in consumption and the expected change in permanent income. Given that utility is quadratic, the utility function approach also generates an analogous expression. Note, however, that the parameter interpretation differs between the two approaches.

In all the work to follow, reference will be made solely to the quadratic loss model; that is, Equation (4.13). All interpretations are thus couched in terms of this approach. Analysis of $\lambda_1 \to 1$ will, however, consider both specifications because the expression for the change in consumption is slightly different using the utility function approach.

Finally, we note the finding that the change in consumption is a function of the lagged change in consumption and the change in permanent income is not unique to the modeling of costly adjustment. Deaton (1987, p. 138) shows that when habit formation is considered and given that preferences are fixed over time, the change in consumption can be expressed as

$$\Delta c_t = \alpha \Delta c_{t-1} + (1 - \alpha) \Delta y_{p_t},$$

where $\alpha$ is a measure of habit formation. Adding and subtracting $(1 - \alpha)E_{t-1}y_{p_t}$ we arrive at:

$$\Delta c_t = \alpha \Delta c_{t-1} + (1 - \alpha)E_{t-1}\Delta y_{p_t} + \overline{\omega_t}, \quad \text{where} \quad \overline{\omega_t} = (1 - \alpha)(E_t - E_{t-1})y_{p_t},$$

which is of a form similar to Equation (4.13), albeit subject to the restriction that $q_2$ is equal to $(1 - q_1)$, a restriction also implied by Attfield, Demery, and Duck (1992) in their PIH1 model.
Expressing the change in consumption in such a manner shows that the costly adjustment model is functionally similar to a model which incorporates habit formation. Functional equivalence, however, is not really surprising given that one possible suggestion for the cost associated with the deviation between actual and desired consumption is the forcible breaking of habits coupled with the resultant change in consumption behavior. In this sense, the costly adjustment model produces an estimable result similar to a habit formation model.

Consumption Properties

Before turning to the estimation of Model I, we consider how the notions of orthogonality, excess sensitivity, and excess smoothness relate to the notion of costly adjustment, which provides insight into how Model I compares with the literature. Studying this question allows us to see if this model is consistent with consumption as observed in Chapters 2 and 3. If these properties are found, the model developed in this chapter would appear in accord with the empirical facts. Beyond displaying these properties, however, we have provided a model that offers a theoretical explanation as to how these conditions may arise, namely, through a model of costly adjustment.

Under the assumption of rational expectations and given that $0<\lambda_1\leq 1$, we have shown that the change in consumption is a function of lagged consumption and also of lagged values of $\{\Delta y_t, \Delta y_{t-h}, \Delta g_t\}$ which suggests that the orthogonality condition of Hall (1978) is not met. Orthogonality will only occur when $\lambda_1 = 0$; that is, when $q_1 = 0$ and $q_2 = 1$, implying that consumption follows a random walk.

Similarly, excess sensitivity, where lagged changes in income provide important information about the change in consumption, will occur when $0<\lambda_1<1$. This sensitivity will not occur in either of the special cases $\lambda_1 \to 0$ or $\lambda_1 \to 1$, because in these special cases consumption is either white noise or a function of its lag, respectively. By showing that the change in consumption is a function of changes in labor income, the general model $0<\lambda_1<1$, is thus consistent with the excess sensitivity literature.
We now consider whether Model I displays the smoothness of consumption relative to permanent income that has been observed in the literature. To analyze this question, we study a version of Model I that uses permanent income as derived by Attfield, Demery, and Duck (1992), that is, we impose the condition \((1 - q_2) = q_1\), which is an obvious assumption. We choose to work in this framework since it simplifies the derivations presented below.

Smoothness, as discussed in Chapter 2, smoothness relates to the theoretical and observational fact that for many models of consumption behavior the variance associated with observed consumption is smaller than that associated with consumption as predicted using the PIH. Using Equation (4.21), we can observe under what conditions such smoothness will be observed in Model I. Previously, Equation (4.21) posited, after imposing the restriction \((1 - q_2) = q_1\), that the change in consumption evolved as

\[(4.31) \Delta c_t = q_1 \Delta c_{t-1} + rq_2 (1 - q_2) \Delta c_{t-1} + q_2 \omega_t,\]

where \(\omega_t\) denotes a white noise term. Consider now finding the variance of Equation (4.31):

\[\text{Var}(\Delta c_t) = q_1^2 \text{Var}(\Delta c_{t-1}) + [rq_2 (1 - q_2)]^2 \text{Var}(\Delta c_{t-1}) + q_2^2 \text{Var}(\omega_t) \]

or:

\[\left(1 - q_1^2 - [rq_2 (1 - q_2)]^2\right)\text{Var}(\Delta c_t) = q_2^2 \text{Var}(\omega_t),\]

\[
\frac{\text{Var}(\Delta c_t)}{\text{Var}(\omega_t)} = \frac{q_2^2}{\left(1 - q_1^2 - [rq_2 (1 - q_2)]^2\right)}.
\]

Now, when the PIH is correct there is no excess smoothness implying the \(\text{Var}(\Delta c_t)\) is equal to \(\text{Var}(\omega_t)\). Note that when the PIH is correct, \(\lambda_t \rightarrow 0\), implying that \(q_1 = 0\) and \(q_2 = 1\), with these values the ratio of variances is indeed 1. In the general case of \(0 \leq \lambda_t < 1\) it can be shown that \(\frac{\text{Var}(\Delta c_t)}{\text{Var}(\omega_t)} < 1\), when \(0 < q_2 < 1\). With this parameter specification, Model I will display excess smoothness as noted by Campbell and Deaton (1989), albeit after imposing the restriction \(q_1 = (1 - q_2)\).

By being able to display excess sensitivity, excess smoothness, and non-orthogonality, under various model specifications, Model I is shown to be consistent with much of the
literature. Model I is also consistent with the results of the PIH when the condition \( \lambda_1 = 0 \) is imposed. With these results, we now turn to the actual estimation of the model.

**Estimation**

We start by considering parameter estimates obtained from an unrestricted VAR system \( \{ \Delta c_t, \Delta y_t, \Delta y_k, \Delta g_t \} \). These estimates will allow a comparison to be made between an unrestricted and restricted system while permitting an explicit test of whether the restrictions implied by Model I are warranted. Results for this unconstrained system are

\[
\begin{bmatrix}
\Delta c_t \\
\Delta y_t \\
\Delta y_k \\
\Delta g_t
\end{bmatrix} =
\begin{bmatrix}
0.041 & 0.09** & 0.01 & 0.013 & 0.033 & -0.001 & -0.01 & 0.058* & -0.012 & 0.012 \\
0.022 & 0.013 & 0.11** & 0.24 & -0.12 & 0.04 & 0.08 & 0.211** & 0.04 & 0.3 \\
1.18** & 0.38* & -0.02 & -0.94 & -0.06 & -0.03 & 0.004 & 0.2 & -0.08 & 0.23 \\
-0.11 & 0.01 & 0.02 & 0.21** & 0.004 & -0.02 & 0.08 & 0.05 & -0.02 & 0.134
\end{bmatrix}
\]

where ** denotes significance at the 5 percent level while * is significance at the 10 percent level. The determinant of the SSR matrix associated with this system was found to be 0.34578.

Estimation of the general model proceeds by imposing all the restrictions that the model imposes upon the coefficients \( a^*_{11}, a^*_{12}, a^*_{13}, \ldots, a^*_{110} \) in System (4.27), which are derived and displayed in Appendix B, and minimizing the determinant of the SSR matrix. As mentioned in Chapter 3, the data sample ranges from 1957Q1 through 1993Q4. Parameter estimates of the general model are
with \( q_1 = 0.069 \) and \( q_2 = 0.221 \). It was found that the determinant of the SSR matrix, \( \gamma_{\epsilon\tau} = |\sum \epsilon \epsilon'| \), obtained the value of 0.36614, where \( \epsilon \) [a \((144 \times 4)\) matrix] denotes the resultant residual series. These results reveal a value of \( q_1 \) close to zero. Before any conclusions are made, however, it must be recalled that \( q_1 \) is not equal to \( \lambda_1 \) in the general model. As shown, in the general case \( q_1 \) is a function of \( X, r, \) and \( D \). To make any conclusions with regard to the general model, the two special cases must be analyzed.

With \( \lambda_1 = 0 \), Hall's (1978) random walk is the proper representation of the change in consumption because the expected change in permanent income is zero. To determine the validity of this case, the system \( \{\Delta c_t, \Delta y_{lt}, \Delta y_{kt}, \Delta g_t\} \) is estimated as above but with \( \Delta c_t \) assumed to be a random walk. Following estimation, the determinant of the SSR matrix of the resultant residual series, denoted as \( \gamma_{\Delta c_t=0} \), was found to be 0.4316.

With \( \gamma_{\epsilon\tau} \) and \( \gamma_{\lambda_1=0} \) the null hypothesis that \( \lambda_1 \rightarrow 0 \) can be tested. For this purpose, a likelihood ratio test of the form \( L = 144[\ln(\gamma_{\lambda_1=0}) - \ln(\gamma_{\epsilon\tau})] \) was performed where \( L \sim \chi^2(1) \). One degree of freedom is used because only one restriction is imposed upon the system, namely, that of \( \lambda_1 \rightarrow 0 \). As seen in the derivations that led to Equation (4.13), imposition of this restriction implies that no variables enter into the consumption equation. Computation shows that \( L \) is equal to 23.74 with a \( p \)-value of effectively zero, suggesting rejection of the null hypothesis.
The other special case of interest is $\lambda_i \rightarrow 1$, which implies that $\Delta c_i = \Delta c_{i-1} + \tilde{\varepsilon}_i$. To test this case, the VAR(3) system of $\{\Delta c_i, \Delta y_u, \Delta y_{kt}, \Delta g_t\}$, where $\Delta c_i$ is assumed to evolve as $\Delta c_i = \Delta c_{i-1} + \tilde{\varepsilon}_i$, is estimated and the determinate of the SSR matrix found. Denoting this determinant as $\gamma_{\lambda_i=1}$, estimation provided an estimate of 0.724. A likelihood ratio test similar to that already discussed, produced a value of $L = 98.2$ distributed with one degree of freedom, suggesting that the special case $\lambda_i \rightarrow 1$ can easily be rejected.

A test of $\lambda_i \rightarrow 1$ can also be performed for the case when the utility function derivation of Model 1 is considered. When Model 1 is derived in terms of the utility function, or more specifically when $u_3$ is zero and $u_4$ is nonzero, Equation (4.13) assumes the form $\Delta c_i = \theta \Delta c_{i-1} + \sigma_i$ under the supposition $\lambda_i \rightarrow 1$. Again this approach is a bit more general than the previous test for $\lambda_i \rightarrow 1$ because the coefficient is no longer constrained to be 1. Testing this special case in this alternative model (by a method analogous to that shown above) produces an estimate of the determinant of the SSR matrix of 0.41544, producing a test statistic of 18.2 with a corresponding $p$-value of 0.0002. The special case $\lambda_i \rightarrow 1$ can thus be rejected, regardless of which method is used in deriving Model 1.

Rejection of the two special cases offers support for the general model. Even though a unique value for $\lambda_i$ cannot be assigned in this model, confirmation of the functional form for the change in consumption has been found. Contrary to Hall's (1978) model, the importance of information other than lagged consumption has been found.

Given the adequacy of the general model, at least with regard to the special cases, the ability of such a model to represent nondurable consumption is considered. The most obvious way of achieving this is to plot actual consumption relative to that predicted by the model which appears as Figure 4.1.

As shown in Figure 4.1, estimated consumption tracks the actual level quite well. Similar trends exist between the actual and fitted series. The greatest variability is observed in periods associated with local peaks and troughs. A plot of the residual series for consumption,
Figure 4.2, more clearly reveals this variability and also the possibility of a change in the mean and variance following the first oil crisis.

ACF plots of the residual series associated with $\Delta c_t$, $\Delta y_h$, $\Delta y_{ykt}$, and $\Delta g_t$ are presented in Figures 4.3 through 4.5. These plots provide insight into the appropriateness of Model I. Ninety-five percent confidence intervals are included, offering a basis for which the autocorrelations may be judged for significance. Autocorrelations for a properly specified model would be expected to be statistically insignificant.

These plots reveal residual series which may be considered as being white noise. Large spikes, however, occur at lag 8 for the residual series associated with $\Delta c_t$, $\Delta y_h$, and $\Delta y_{ykt}$. All the autocorrelations for all four series are statistically insignificant at the 5 percent level.

![Figure 4.1. Actual versus fitted consumption: Model I](image)
Figure 4.2. Consumption residuals: Model I

Figure 4.3. ACF $\Delta c_t$ residuals: Model I
Figure 4.4. ACF $\Delta y_t$, residuals: Model I

Figure 4.5. ACF $\Delta y_k$, residuals: Model I

Figure 4.6. ACF $\Delta g_t$, residuals: Model I
A more formal test for white noise is provided with Ljung-Box $Q^*$ statistics, which were described in Chapter 3. Under the null hypothesis of white noise this statistic is distributed as $\chi^2(k)$, with $k$ being the number of lags used. These tests are only approximate, however, because possible correlations among the residual series are ignored. A test which allows for cross-correlations will be presented below. Results of the Ljung-Box test ($Q^*$), as formulated in Chapter 3, are presented in Table 4.1 for various lag specifications. The p-values appear in parentheses.

Table 4.1. Ljung-Box statistics: Model I

<table>
<thead>
<tr>
<th>Lag</th>
<th>$\Delta c_t$</th>
<th>$\Delta y_{ht}$</th>
<th>$\Delta y_{kt}$</th>
<th>$\Delta g_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>1.18</td>
<td>0.32</td>
<td>0.39</td>
<td>0.41</td>
</tr>
<tr>
<td></td>
<td>(0.88)</td>
<td>(0.99)</td>
<td>(0.98)</td>
<td>(0.98)</td>
</tr>
<tr>
<td>8</td>
<td>5.21</td>
<td>5.59</td>
<td>6.24</td>
<td>3.33</td>
</tr>
<tr>
<td></td>
<td>(0.74)</td>
<td>(0.69)</td>
<td>(0.62)</td>
<td>(0.91)</td>
</tr>
<tr>
<td>12</td>
<td>9.64</td>
<td>8.26</td>
<td>12.66</td>
<td>5.87</td>
</tr>
<tr>
<td></td>
<td>(0.65)</td>
<td>(0.77)</td>
<td>(0.39)</td>
<td>(0.92)</td>
</tr>
<tr>
<td>16</td>
<td>10.16</td>
<td>13.52</td>
<td>13.46</td>
<td>7.5</td>
</tr>
<tr>
<td></td>
<td>(0.86)</td>
<td>(0.64)</td>
<td>(0.64)</td>
<td>(0.96)</td>
</tr>
</tbody>
</table>

Note: p-values appear in parenthesis.

Though approximate, these results suggest that each residual series is approximately white noise. The null hypothesis of white noise could not be rejected at normal levels of significance at each of the lag specifications.

A more appropriate test which takes into account cross-correlations that may arise in a multivariate system is the Portmanteau test. Ideally, these residual auto- and cross-correlations should be close to zero for a properly specified model. The Portmanteau test provides a test for the overall significance of the residual autocorrelations up to some specified lag (Lutkepohl 1991, p. 150). Under the null hypothesis of white noise, the test statistic $P^*$ is
distributed as $\chi^2[K^2(h - p)]$, where $K$ is the number of series used in the system, $h$ is the lag specification of the test, and $p$ denotes the order of the VAR system. Following Lutkepohl, the test statistic is

$$P_h = T^2 \sum_{i=1}^{h} (T - i)^{-1} \text{trace}(C_i' C_i^{-1} C_i C_i^{-1})$$

$$C_i = \frac{1}{T} \sum_{t=i}^{T} \mu_t \mu_{t-i},$$

where $T$ is the sample size and $\mu_t$ denotes a column vector of residuals [in the present case of dimension (4 x 1)]. Choosing $h = 24$ produces the test statistic $P_{24} = 330.5$ with a p-value of 0.58, suggesting that the null hypothesis of no correlations among the residuals series cannot be rejected. Thus, univariate, and multivariate analyses of the residual series suggest that Model 1 provides an adequate representation of the data.

Last, we consider a simple test for the constancy of the parameters in Model 1. This test involves dividing the sample into two portions and estimating the model over each portion separately. From the figures of the data presented in Chapter 3 and the results in this section, we choose 1973 quarter 2 as being a division point. The determinant of the SSR matrix for each segment is computed, denoted as $\nu_1$ and $\nu_2$, and compared to that attained when the model is estimated over the whole sample, $\nu_{wc}$. Following Lutkepohl (1991, pp. 400-402), a likelihood ratio test can be used of the form

$$LR = 2 \left\{ -\frac{1}{2} [n_1 (-4\ln(n_1) + \ln(\nu_1))] + n_2 [-4\ln(n_2) + \ln(\nu_2))] + \frac{T}{2} (-4\ln(T) + \ln(\nu_{wc})) \right\},$$

where $n_1$ and $n_2$ denote the number of observations in the first and second segments (61 and 83, respectively) and $T$ refers to the sample size (144 in this case). LR is distributed approximately $\chi^2$ with 29 degrees of freedom, because we are imposing 29 constraints to arrive at the null hypothesis of equal coefficients over the two periods. Upon dividing the sample, the determinant of the SSR matrix, $\nu_1$, was found to be 0.006667, based on the first 64 quarters of data (57Q1 through 73Q2), whereas for the second segment (73Q3 to 93Q4) $\nu_2$ was found to be 0.02588. Calculation produces
LR^* = 2(1539.6 - 1503.6) = 71.97 ,
with 29 degrees of freedom and a p-value of effectively zero. This result suggests that we can confidently reject the null hypothesis of constancy of parameters for Model I. An attempt to model this system in terms of time-varying parameters is made in the following chapter.

Model II

In this section, we introduce a model which can be viewed as a simplistic version of Model I. An alternative representation for the costly adjustment model is sought because of the level of knowledge that the estimation of Model I imposed upon the agent. In estimating system (4.20), several informational assumptions must be made. First is that the agent can correctly conceive and model the infinite expected sums that describe permanent income which, by assuming rational expectations, we take for granted that the individual can perform. Second, given that the agent can determine the necessary infinite expected sums is the requirement that these quantities can be incorporated into the model in the guise of the cross-equation restrictions. Last is the assumption that the agent correctly incorporates the auxiliary system to determine the necessary expectations and that the parameters in this system are known and fixed.

This section seeks to address the first two assumptions by introducing a costly adjustment model which imposes a weaker informational assumption upon the agent. We defer discussion of the third point until the next chapter. The material considered presently is best viewed as a precursor to the work in the succeeding chapter, where the informational requirements of the rational expectations hypothesis are explicitly discussed and a simple learning model presented.

Replacing the infinite expected sums in the manner of Model I requires the agent to have a broad understanding of the system. Beyond assuming that the agent knows the
structure of the PIH, knowledge of the auxiliary system is also presupposed. In estimating Model I, we assumed that the agent proxies the expected change in permanent income as

\[ E_{t-1} \Delta y_{pt} = r(p \begin{pmatrix} 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} A \begin{pmatrix} 1 & -pA \end{pmatrix} X_{t-1} + \begin{pmatrix} 0 & 0 & 1 & 0 \end{pmatrix} AX_{t-1} , \]

where \( A \) and \( X_{t-1} \) are as defined previously. Model II, to be introduced, assumes less analytical skills on the part of the agent. Rather than assuming that the individual fully understands the complicated algebra embodied in Model I, the milder condition that the agent uses a simplistic rule is made. This rule can then be used in deriving a notion of the expected change in permanent income in a manner consistent with an individual who does not possess total knowledge of both the model and the mathematics.

Central to Model II is the assumption that the agent posits the expected change in permanent income as

\[ (4.32) E_{t-1} \Delta y_{pt} = \alpha_1 E_{t-1} \Delta y_{lt} + \alpha_2 E_{t-1} \Delta y_{kt} + \alpha_3 E_{t-1} \Delta g_t . \]

Equation (4.32) appears reasonable because the information embodied in this rule is equivalent to that incorporated by the expression for \( E_{t-1} \Delta y_{pt} \) used in the previous section. This can readily be seen by assuming a VAR(3) auxiliary system for \( \{ \Delta y_{lt}, \Delta y_{kt}, \Delta g_t \} \) which, given the definition of \( \Delta y_{pt} \) discussed previously, shows that \( E_{t-1} \Delta y_{pt} \) is a function of \( \{ \Delta y_{lt-1}, \Delta y_{kt-1}, \Delta g_{t-1}, \Delta y_{lt-2}, \Delta y_{kt-2}, \Delta g_{t-2}, \Delta y_{lt-3}, \Delta y_{kt-3}, \Delta g_{t-3} \} \). Recall that the previous notion for the expected change in permanent income is also a function of these same variables. What differs between the two approaches to determining \( E_{t-1} \Delta y_{pt} \) are the coefficients with which these variables enter into their respective models and the implied cross-equation restrictions. Assuming rational expectations, the coefficients utilized in Model I reflect the necessary infinite sums. In specifying a rule, Model II is able to approximate these complicated coefficients with the \( \alpha_i \)'s of Equation (4.32).

Additionally, Equation (4.32) may be a more realistic approach to representing the expected change in permanent income compared to Model I if it is suspected that the auxiliary system may have changed over the sample. By avoiding the complicated infinite expected
sums, Model II may be less susceptible to any possible misspecification in the formation of expectations. Chapter 5 considers this point in greater detail.

Model II can be explicitly formulated by substituting Equation (4.32) into Equation (4.13); that is

\[ \Delta c_i = q_1 \Delta c_{i-1} + \alpha_1 E_{i-1} \Delta y_h + \alpha_2 E_{i-1} \Delta y_{kt} + \alpha_3 E_{i-1} \Delta g_i + \omega_i, \]

where \( \alpha_1 = q_2 \cdot \alpha_1, \alpha_2 = q_2 \cdot \alpha_2, \alpha_3 = q_2 \cdot \alpha_3. \)

Estimation of Model II proceeds by first specifying a joint system composed of Equation (4.33) and the auxiliary system (4.22). This system may be expressed as

\[
\begin{bmatrix}
\Delta c_{i-1} \\
\Delta y_{h-1} \\
\Delta y_{kt-1} \\
\Delta g_{i-1}
\end{bmatrix}
= 
\begin{bmatrix}
\Gamma_1 & \Gamma_2 & \Gamma_3 & \Gamma_4 & \Gamma_5 & \Gamma_6 & \Gamma_7 & \Gamma_8 & \Gamma_9 \\
0 & b_1 & b_2 & b_3 & b_4 & b_5 & b_6 & b_7 & b_8 & b_9 \\
0 & c_1 & c_2 & c_3 & c_4 & c_5 & c_6 & c_7 & c_8 & c_9 \\
0 & d_1 & d_2 & d_3 & d_4 & d_5 & d_6 & d_7 & d_8 & d_9
\end{bmatrix}
\begin{bmatrix}
\Delta c_{i-1} \\
\Delta y_{h-1} \\
\Delta y_{kt-1} \\
\Delta g_{i-1}
\end{bmatrix}
+ \varepsilon_i,
\]

where \( \Gamma_i = \alpha_i b_i + \alpha_2 c_i + \alpha_3 d_i \) for \( i = 1,2,\ldots,9 \)

and \( \varepsilon_i = \{ \varepsilon_{i1}^{e_i}, \varepsilon_{i2}^{h_i}, \varepsilon_{i3}^{x_i}, \varepsilon_{i4}^{k_i} \}. \)

Note that system (4.34) still requires the incorporation of cross-equation restrictions because an assumption of rational expectations is required by Model II. Also note, however, that the restrictions are much more simple and transparent compared to those utilized in Model I. By assuming a more simplistic manner in which the expected change in permanent income is derived, a system is obtained which allows for the change in consumption to be estimated with less computational burden. Given the complexity of the environment in which the agent operates, this result may be fortuitous because it reduces the amount of knowledge presupposed upon our representative agent.

Model II also permits examination of the special cases noted in the previous section. When \( \lambda_1 \to 0 \), the expected change in permanent income is again 0, implying that \( q_1 \) and \( q_2 \) are
also zero. Equation (4.33) shows that, in this situation, consumption evolves as a random walk. As shown with $\lambda_1 \rightarrow 1$, $q_1 \rightarrow 1$ and $q_2 \rightarrow 0$, which, using Equation (4.33) implies that $\Delta c_t = \Delta c_{t-1}$, again in accord with Model I. A more general test of the special case $\lambda_1 \rightarrow 1$ can be performed where the coefficient on $\Delta c_{t-1}$ is not constrained to be 1. In the general model, $0 < \lambda_1 < 1$, $\lambda_1$ remains unidentified, as in Model I.

As with Model I, we can test the appropriateness of the costly adjustment model in terms of Model II. Testing proceeds by estimating the general model and the two special cases and using a likelihood ratio test, as discussed in the context of Model I. Results of these tests are presented next.

Estimation

We now present estimation results for Model II. These results were obtained by imposing all the restrictions and determining the set of parameter estimates that minimize the determinant of the SSR matrix. Parameter estimates are as follows:

\[
\begin{bmatrix}
\Delta c_t \\
\Delta y_{k_t} \\
\Delta y_{k_{t-1}} \\
\Delta g_t
\end{bmatrix} =
\begin{bmatrix}
0.069 & 0.09 & 0.011 & -0.029 & 0.018 & -0.0018 & -0.0013 & 0.0605 & -0.012 & 0.033 \\
0 & 0.015 & 0.108 & 0.284 & -0.107 & 0.044 & 0.068 & 0.209 & 0.041 & 0.277 \\
0 & 0.49 & -0.001 & -0.364 & 0.153 & -0.023 & -0.079 & 0.217 & -0.074 & -0.01 \\
0 & 0.003 & 0.02 & 0.17 & -0.011 & -0.016 & 0.091 & 0.047 & -0.022 & 0.15
\end{bmatrix}
\begin{bmatrix}
Z_{t-1} \\
Z_{t-2} \\
Z_{t-3}
\end{bmatrix}
\]

\[\bar{\alpha}_1 = 0.0841 \quad \bar{\alpha}_2 = 0.181 \quad \bar{\alpha}_3 = 0.079,\]

where $Z_{t-1} = \{\Delta c_{t-1}, \Delta y_{k_{t-1}}, \Delta y_{k_{t-1}}, \Delta g_{t-1}\}'$,

$Z_{t-2} = \{\Delta y_{k_{t-2}}, \Delta y_{k_{t-2}}, \Delta g_{t-2}\}'$,

$Z_{t-3} = \{\Delta y_{k_{t-3}}, \Delta y_{k_{t-3}}, \Delta g_{t-3}\}'$.

The determinant of the SSR matrix, $\gamma_{ur} = \sum ee'$, associated with these parameter estimates is 0.36476. This quantity is slightly smaller than the corresponding value for Model I because Model II is not as tightly constrained compared to the former. Considerable differences between parameter estimates of the two models are also noted. Equations $\Delta y_{k_t}$, $\Delta y_{k_{t-1}}$, and $\Delta g_t$
reveal the greatest change, whereas parameter estimates in the $\Delta c_i$ equation are broadly consistent between the two models.

Following estimation, tests of the two special cases, $\lambda_i \rightarrow 0$ and $\lambda_i \rightarrow 1$, were performed using likelihood ratio tests. Note that values of the determinant of the SSR matrix for the two special cases are the same as in the previous section because under each special case the structure of $\Delta c_i$ remains the same. Testing whether $\lambda_i \rightarrow 0$ produced a likelihood ratio test statistic of 29.3 with a p-value of effectively zero. Under the restriction $\lambda_i \rightarrow 1$, test statistics of 98.72 and 18.72 were obtained, depending on whether the quadratic cost or utility function method was used in deriving the model, respectively. In either case the hypothesis of $\lambda_i \rightarrow 1$ can easily be rejected. These results suggest that $\lambda_i$ lies within the interval $(0, 1)$, implying that the change in consumption is a function of other variables.

With the acceptance of the model various properties can be analyzed. First, a plot relating actual consumption to estimated consumption is presented in Figure 4.7. As in the case of Figure 4.1 for Model I, Figure 4.7 shows that Model II tracks actual consumption well. Also, increased variability is seen at periods associated with local peaks and troughs of the consumption series. This result is more clearly seen by examining a plot of the residual series, presented in Figure 4.8, which suggests a possible change in mean and variability following the first oil crisis.

ACF plots are also provided for each residual series. Plots of residuals for $\Delta c_i$, $\Delta y_{it}$, $\Delta y_{kt}$, and $\Delta g_i$ appear in Figures 4.9 through 4.12. Again, a properly specified model would imply that autocorrelations in these plots should be small, thus indicating white noise. Ninety-five percent confidence bands are also provided for reference.
Figure 4.7. Actual versus fitted, Model II

Figure 4.8. Consumption residuals Model II
Figure 4.9. ACF $\Delta c_t$ residuals: Model II

Figure 4.10. ACF $\Delta y_{it}$ residuals: Model II

Figure 4.11. ACF $\Delta y_{kt}$ residuals: Model II
Figures 4.9 through 4.12 reveal residual series which may be considered white noise. Large correlations are noted at lag 8 for the residuals associated with $\Delta c_t$, $\Delta y_h$, and $\Delta y_k$, though these correlations are insignificant at the 5 percent level.

More formal tests dealing with the necessary randomness of the residual series were also performed. First, we performed Ljung-Box tests for various lags for each univariate series. Again, this test is only an approximate because cross-correlations are ignored. Results appear in Table 4.2. The p-values are shown in parentheses.

Second, a Portmanteau test was used to test for the significance of auto- and cross-correlations among the four residual series. As with Model I, a lag of 24 was used, which produced a test statistic of $P^{24} = 331$ with an associated p-value of 0.58, suggesting that this null hypothesis could not be rejected. These univariate and multivariate tests suggest the adequacy of Model II.

Last, we examined the possibility of time-varying parameters in Model II. Large deviations between actual and fitted consumption for certain periods may reflect structural change over the sample. To study this possibility we performed a test for the constancy of parameters. As discussed in the previous section, the sample is divided relative to the second quarter of 1973 (73Q2) and Model II is fitted over each segment. The determinant of the SSR
matrix for each segment, \( v_1 \) and \( v_2 \), were computed as 0.00657 and 0.02552, respectively and compared to that obtained when the model is estimated over the whole sample, \( v_{\text{uc}} \).

Calculating a likelihood ratio test, as described previously, provided the estimate that \( LR^* = 77.4 \) with 32 degrees of freedom and an associated p-value of effectively zero. This result suggests that the null hypothesis that parameters are jointly fixed over the two segments can confidently be rejected. Analysis that takes into account the possibility of time-varying parameters is discussed in the following chapter.

<table>
<thead>
<tr>
<th>Lag</th>
<th>( \Delta c_t )</th>
<th>( \Delta y_{it} )</th>
<th>( \Delta y_{kt} )</th>
<th>( \Delta g_t )</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>1.32</td>
<td>0.28</td>
<td>0.50</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td>(0.86)</td>
<td>(0.99)</td>
<td>(0.97)</td>
<td>(0.99)</td>
</tr>
<tr>
<td>8</td>
<td>5.13</td>
<td>5.81</td>
<td>6.28</td>
<td>3.13</td>
</tr>
<tr>
<td></td>
<td>(0.74)</td>
<td>(0.67)</td>
<td>(0.62)</td>
<td>(0.92)</td>
</tr>
<tr>
<td>12</td>
<td>9.68</td>
<td>8.52</td>
<td>12.8</td>
<td>5.76</td>
</tr>
<tr>
<td></td>
<td>(0.64)</td>
<td>(0.74)</td>
<td>(0.38)</td>
<td>(0.93)</td>
</tr>
<tr>
<td>16</td>
<td>10.25</td>
<td>13.98</td>
<td>13.52</td>
<td>7.44</td>
</tr>
<tr>
<td></td>
<td>(0.85)</td>
<td>(0.60)</td>
<td>(0.63)</td>
<td>(0.96)</td>
</tr>
</tbody>
</table>

Note: p-values appear in parentheses.

Cointegration

Estimation of Models I and II relies on the supposition that the series \( \{c_t, y_{it}, y_{kt}, g_t\} \) are difference stationary. Support for this supposition was provided in Chapter 3. Chapter 3, however, also provides evidence that these four variables are cointegrated. As
discussed, an important implication of this property is that VAR estimation carried out solely in first differences is misspecified. Given that the data in this paper are cointegrated, this section attempts to incorporate this relationship in a reestimation of Model I and Model II, thus answering any criticisms based on misspecification. This material has been relegated to an extension because nothing of substance in terms of rejecting or accepting the costly adjustment model is gained when the cointegrating vector is added to the system. Note, however, that such an addition does have a statistically significant impact on both models.

To study cointegration and consumption, this section is divided as follows. First, we examine the importance of the cointegrating vector in the unrestricted VAR model by estimating an error correction model. Next, Model I is rederived to take into account the cointegrating vector. We assume that this vector enters through the auxiliary system. Following the estimation results of Model I, a redefined version of Model II is presented.

**Error Correction**

Error correction models are nothing new in consumption theory. Davidson, Hendry, Sbra, and Yeo (1978) applied such a model in response to the poor performance of so-called "theory" based models. This subsection will fit an error correction model to the relevant data, thus providing insight into what may be expected when Models I and II are reestimated using this larger information set.

In Chapter 3, we determined that the variables \{c_t, y_t, y_{kt}, g_t\} possessed one cointegrating vector which, following normalization, may be expressed as

\[ Cl_t = - c_t + 6.13 - 0.16y_t - 1.4y_{kt} + 1.88g_t. \]

By definition \( Cl_t \) is a stationary variable and is termed the equilibrium error by Engle and Granger (1987a). To formulate an error correction model, we simply augment the term \( \alpha Cl_{t-1} \) to the VAR(3) system that we have utilized throughout this paper. The vector \( \alpha \) can be interpreted as a measure of the speed by which the system corrects last period's equilibrium error (Campbell and Perron 1991, page 30). This system may be represented as
(4.35) \( \Delta Z_t = \alpha \Delta Z_{t-1} + A_1 \Delta Z_{t-1} + A_2 \Delta Z_{t-2} + A_3 \Delta Z_{t-3} + \epsilon_t \)

using the notations from previous sections. Estimating Equation (4.35) using the data from Chapter 3 produces the following results.

\[
\begin{bmatrix}
\Delta c_1 \\
\Delta y_{y1} \\
\Delta y_{yt} \\
\Delta g_{t1}
\end{bmatrix} = \begin{bmatrix}
0.032 & 0.022 & 0.093** & 0.009 & 0.013 & 0.031 & -0.0002 & -0.009 & 0.055* & -0.013 & 0.015 \\
0.11 & -0.19* & 0.02 & 0.12** & 0.21 & -0.1 & 0.05 & 0.024 & 0.25** & 0.051 & 0.24 \\
1.65** & -1.09** & 0.38* & 0.035 & -0.94* & 0.003 & 0.03 & -0.08 & 0.37** & 0.002 & 0.01 \\
-0.17 & 0.12** & 0.008 & 0.018 & 0.23** & -0.01 & -0.02 & 0.12 & 0.026 & -0.028 & 0.17*
\end{bmatrix}
\]

where * denotes significance at the 10 percent level and ** indicates significance at the 5 percent level. The determinant of the SSR associated with the residuals from this system was found to be 0.28433. Significance of the error correction terms can be tested by comparing the value 0.28433 obtained from the error correction model to the determinant of the SSR matrix obtained when the restrictions that all error correction terms are zero (\( \alpha = 0 \)) are imposed. Under these restrictions, estimation proceeds with the unconstrained VAR model noted above, whereas the determinant of the SSR matrix was found to be 0.34578. Computing a likelihood ratio test provides a test statistic of 28.2 with 4 degrees of freedom and a p-value of 0.0001, suggesting that the error correction terms have a significant effect on the system. We now proceed to incorporate these error correction terms into Models I and II.

**Model I**

Cointegration is incorporated into Model I by expanding the auxiliary system as follows:

(4.36) \( Z_t = \beta C_{t-1} + C_1 Z_{t-1} + C_2 Z_{t-2} + C_3 Z_{t-3} + \epsilon_t \),

where \( Z_t = (\Delta y_{y1}, \Delta y_{yt}, \Delta g_{t1}) \).
The column vector $\beta$ measures the importance of equilibrium error in its respective equation. Estimation proceeds with joint estimation of Equation (4.13) and the modified auxiliary system (4.36). As in the preceding section, the model suggests certain cross-equation restrictions that must be imposed, derivations of which are presented in Appendix C. Parameter estimates of this rederived Model I can be expressed as

$$\begin{bmatrix}
\Delta c_i \\
\Delta y_{t-1} \\
\Delta y_{u-1} \\
\Delta g_{t-1}
\end{bmatrix} = 
\begin{bmatrix}
0.172 & 0.038 & 0.005 & -0.039 & -0.058 & 0.012 & 0.0013 & -0.008 & 0.003 & -0.001 & 0.002 \\
0 & -0.004 & 0.12 & 0.24 & -0.222 & -0.1 & 0.054 & 0.05 & 0.23 & 0.06 & 0.26 \\
0 & 0.59 & 0.07 & -0.062 & -0.91 & 0.172 & 0.016 & -0.13 & 0.458 & -0.017 & 0.031 \\
0 & -0.02 & 0.013 & 0.173 & 0.082 & -0.022 & -0.02 & 0.102 & 0.017 & -0.025 & 0.15
\end{bmatrix}$$

$q_t = 0.063$.

The determinant of the SSR matrix associated with this system was found to be 0.331227. The coefficients are quite similar when compared to those obtained under Model I with no cointegration. A large difference, however, appears in the coefficient of $\Delta c_{t-1}$. To test whether each univariate residual series is individually white noise, Ljung-Box tests were performed, results are presented in Table 4.3.

A Portmanteau test was utilized to test whether the four residual series are jointly white noise. Using a lag of 24, this test produced the result that $P^{24} = 329.15$ with a p-value of 0.6, suggesting that the null hypothesis of white noise cannot be rejected. The ACF of the residuals associated with the change in consumption appears as Figure 4.13.
Figure 4.13. ACF $\Delta c_i$ Model I with cointegration

Table 4.3. Ljung-Box statistics: Model I with cointegration

<table>
<thead>
<tr>
<th>Lag</th>
<th>$\Delta c_i$</th>
<th>$\Delta y_{it}$</th>
<th>$\Delta y_{kt}$</th>
<th>$\Delta g_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>4.06</td>
<td>0.87</td>
<td>0.79</td>
<td>0.47</td>
</tr>
<tr>
<td></td>
<td>(0.4)</td>
<td>(0.93)</td>
<td>(0.93)</td>
<td>(0.98)</td>
</tr>
<tr>
<td>8</td>
<td>8.74</td>
<td>5.23</td>
<td>3.31</td>
<td>3.39</td>
</tr>
<tr>
<td></td>
<td>(0.37)</td>
<td>(0.73)</td>
<td>(0.91)</td>
<td>(0.91)</td>
</tr>
<tr>
<td>12</td>
<td>13.32</td>
<td>9.42</td>
<td>9.08</td>
<td>6.33</td>
</tr>
<tr>
<td></td>
<td>(0.35)</td>
<td>(0.67)</td>
<td>(0.7)</td>
<td>(0.9)</td>
</tr>
<tr>
<td>16</td>
<td>13.51</td>
<td>15.43</td>
<td>9.65</td>
<td>8.13</td>
</tr>
<tr>
<td></td>
<td>(0.64)</td>
<td>(0.5)</td>
<td>(0.88)</td>
<td>(0.95)</td>
</tr>
</tbody>
</table>

Note: p-values appear in parentheses.

These results show that this cointegration version of Model I produces good results. Univariate and multivariate tests suggest that residual series are reasonably clean. A likelihood ratio test can also be performed to examine whether incorporating the cointegrating vector has an appreciable effect on Model I. Performing such a test with the null hypothesis of no error correction term produces a test statistic of 14.43 with 3 degrees of freedom and a p-value of 0.002 which can be rejected at any reasonable level of significance, suggesting that cointegration has a significant impact upon Model I.
Results of this error correction version of Model I also imply that the two special cases of $\lambda_1 \rightarrow 0$ and $\lambda_1 \rightarrow 1$ can both be rejected because the determinant of the SSR matrix is smaller compared to Model I while estimated quantities under the special cases remain unchanged. This result suggests that even though the incorporation of error correction provides a superior model, nothing of substance is gained in terms of accepting or rejecting the cost of adjustment model.

Model II

The cointegrating term is incorporated into Model II by expanding the auxiliary system which may be expressed as

\[
(4.37) \quad Z_t = \beta C_{t-1} + C_1 Z_{t-1} + C_2 Z_{t-2} + C_3 Z_{t-3} + \epsilon_t,
\]

where $Z_t = \{\Delta y_{it}, \Delta y_{it}, \Delta g_{it}\}$.

Estimation proceeds with joint estimation of Equation (4.37) with the modified auxiliary system. As with the original model, cross-equation restrictions must be imposed that are of the same format as the antecedent Model II. Parameter estimates of Model II incorporating the cointegrating vector are expressed as

\[
\begin{bmatrix}
\Delta c_{t-1} \\
\Delta y_{it} \\
\Delta y_{it} \\
\Delta g_{it}
\end{bmatrix}
= \begin{bmatrix}
0.042 & 0.092 & 0.008 & 0.014 & 0.009 & 0.03 & -0.008 & 0.01 & 0.055 & -0.02 & 0.039 \\
0 & 0.023 & 0.12 & 0.26 & -0.19 & -0.01 & 0.04 & 0.08 & 0.25 & 0.04 & 0.31 \\
0.054 & 0.08 & -0.67 & -0.9 & 0.16 & 0.06 & -0.25 & 0.42 & -0.04 & -0.136 \\
0 & -0.002 & 0.02 & 0.18 & 0.13 & -0.018 & -0.01 & 0.07 & 0.023 & -0.018 & 0.12
\end{bmatrix}
\]

\[\bar{a}_1 = -0.186, \quad \bar{a}_2 = -0.183, \quad \bar{a}_3 = 1.042.\]

The determinant of the associated variance/covariance matrix was found to be 0.31885. Coefficient estimates are noted to be quite similar compared to those for Model II.
when cointegration is ignored. To test whether each univariate residual series is individually white noise, Ljung-Box tests were performed. The results of these tests appear in Table 4.4.

A Portmanteau test was utilized to test whether the four residual series are jointly white noise. Using a lag of 24, this test produced the result that $P^{24} = 328$ with a p-value of 0.62, suggesting that the null hypothesis of white noise cannot be rejected. The ACF of the residuals associated with the change in consumption appear in Figure 4.14.

Table 4.4. Ljung-Box statistics: Model II with cointegration

<table>
<thead>
<tr>
<th>Lag</th>
<th>$\Delta c_t$</th>
<th>$\Delta y_n$</th>
<th>$\Delta y_{kt}$</th>
<th>$\Delta g_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>1.13</td>
<td>0.84</td>
<td>1.39</td>
<td>1.63</td>
</tr>
<tr>
<td></td>
<td>(0.89)</td>
<td>(0.93)</td>
<td>(0.85)</td>
<td>(0.80)</td>
</tr>
<tr>
<td>8</td>
<td>5.91</td>
<td>5.42</td>
<td>4.01</td>
<td>3.91</td>
</tr>
<tr>
<td></td>
<td>(0.66)</td>
<td>(0.71)</td>
<td>(0.86)</td>
<td>(0.87)</td>
</tr>
<tr>
<td>12</td>
<td>11.56</td>
<td>8.45</td>
<td>10.5</td>
<td>7.04</td>
</tr>
<tr>
<td></td>
<td>(0.49)</td>
<td>(0.75)</td>
<td>(0.57)</td>
<td>(0.86)</td>
</tr>
<tr>
<td>16</td>
<td>12.51</td>
<td>14.36</td>
<td>11.2</td>
<td>8.46</td>
</tr>
<tr>
<td></td>
<td>(0.71)</td>
<td>(0.57)</td>
<td>(0.8)</td>
<td>(0.93)</td>
</tr>
</tbody>
</table>

Note: p-values appear in parentheses.

Figure 4.14. ACF $\Delta c_t$ Model II with cointegration
With these results, the obvious question arises as to whether the addition of error correction has a statistically significant effect on the model. To provide an answer, a likelihood ratio test for the significance of the \( \beta \) vector produced a test statistic of 19.37 with 3 degrees of freedom and a p-value of 0.0002, implying that the cointegration vector significantly contributes to this model. Analysis of the residual series also reveals a model that adequately represents the data.

Because the determinant of the SSR matrix of the cointegration model is less than the corresponding value of the previous version of Model II, we can also conclude that the two special cases of \( \lambda_1 \to 0 \) and \( \lambda_1 \to 1 \) can both be rejected in this instance. Thus, even though cointegration provides a possibly superior model, nothing of substance changes with regard to the notion of costly adjustment. In both versions of Model II, the general case of \( 0 < \lambda_1 < 1 \) is relevant, implying that costs of adjustment as defined in this chapter are found to exist, regardless to whether or not cointegration is considered.

**Policy Implications**

Before concluding this chapter, we consider the effects of government policy in the costly adjustment framework. An important implication with the work presented in this chapter relates to the ability of the government to influence consumption. Under the PIH, as formulated by Hall (1978), consumption follows a random walk implying that the best prediction of next period's level of consumption is today's level. In this environment government intervention cannot alter expected levels since current consumption is orthogonal to all lagged information except the lagged level of consumption.

The situation, however, is very much different under the existence of costs of adjustment. Theoretically, the existence of these costs show that the current change in
consumption is a function of its lagged change and the expected change in permanent income. Altering the agent’s perceptions in regards to changes in permanent income provides a means of influencing consumption determination.

With the acceptance of the general model \((0<\lambda_t<1)\) found in the preceding sections, expected changes in gross labor income, capital income net government bonds, and government expenditures will all impact upon the consumption path of the representative agent. Consumption may thus be expanded or contracted through expected movements of the variables \(\{\Delta y_t, \Delta y_k, \Delta g_t\}\). Being a function of these variables an important avenue is open to the government in influencing consumption levels.

Policies may alter consumption in several possible ways. First is through expected changes in gross labor income and capital income. By boosting the confidence of agents in terms of what the future may have in store, increases in consumption may be realized today. Similarly, if expectations of the future are pessimistic, current consumption would be predicted to decrease. Many possibilities exist in influencing expected gross labor income and capital income. For example, a reduction in unemployment and a devaluation of the dollar may lead to an expected increase in gross labor income, while an increase in the interest rate or a decrease in the CPI may bring about an expected increase in capital income. These possibilities suggest that by altering key macroeconomic series and indicators, through various policies, expected income may change, allowing for some degree of control of consumption.

Consumption may also be influenced through expected changes in government expenditures. With the assumption of Ricardian equivalence the sum of expected, discounted government expenditures is equal to the expected sum of discounted taxes, minus current bonds, implying that consumption will decrease under the perception of an increased size of government. Advocacy of the downsizing of government offers one possible means of increasing consumption levels. Not surprisingly, it is also noted that increases in government debt reduces consumption through a decrease in \(y_k\).

A final approach in directing consumption is through the costs terms \(\alpha_0\) and \(\alpha_1\). By altering the costs of adjustments a degree of control over consumption is achieved. It is noted
that if the government did a superb job of removing all imperfections that may confront the agent, consumption will ultimately evolve as a random walk, eliminating any possibility of influencing consumption levels. It would thus seem in the best interest of the government to allow capital markets to possess some degree of rigidities, ensuring some discretion in altering consumption levels. Beyond this, it is also possible for the government to alter $\alpha_0$ and $\alpha_1$ through deregulation, greater dissemination of information, or greater minority lending, for example, so as to change the relative magnitudes associated with the lagged change in consumption and the expected change in permanent income ($q_1$ and $q_2$ of Model I) which comprise the current change in consumption. Changing such quantities offers a means of altering the composition of current consumption. Possible alterations in the composition of the change of consumption are important since while the expected change in permanent income is liable to some degree of control, the lagged change in consumption is not. Influence on current consumption levels may be best served by increasing the importance associated with the expected change in permanent income.

What seems interesting in this discussion is the efficacy of government intervention even under the assumption of rational expectations. As opposed to much of the work of the 1970s and the early 1980s which purported to show the futility of meddling on the part of the government, this chapter produces results that are quite different. The basis of these results arise from the possibility that the perfect capital market assumption is unfounded. Rejecting such an assumption as acceptance of the general model implies, suggests that consumption does not evolve as a random walk opening up the possibility of altering consumption levels through government policies. While the agent still looks towards the future in formulating a consumption path, rigidities introduced by imperfect capital markets permit the government to affect consumption through the perceptions of the future and with changes in the costs of adjustment.
Conclusions

We end this chapter by considering the various conclusions that have been gained from the costly adjustment framework. Beyond evaluating Models I and II, the implications that these models have on the estimation of consumption is also discussed. The most significant conclusion of this chapter is that the costly adjustment framework could not be rejected.

We begin by considering Model I. Before attempting to interpret parameters, the adequacy of the model was ascertained. The resultant residual series were found to be reasonably close to white noise as tested using univariate and multivariate methods. One cause for concern is the possibility of time-varying parameters discussion of which, however, is deferred until Chapter 5. Testing the adequacy of Model I also necessitated testing whether the restrictions imposed by the model are warranted. If the implied restrictions are too strong (i.e., rejected by the data) any inference made by this model about the evolution of consumption would be meaningless. Testing whether the restrictions implied by Model I are statistically legitimate proceeds by comparing the determinant of the SSR matrix obtained from Model I to that found when no restrictions are imposed, which is the unconstrained VAR system discussed in the second section of this chapter. A likelihood ratio test produced a test statistic of 8.24 with 13 degrees of freedom and a p-value of 0.83, suggesting that the restrictions implied by Model I cannot be rejected.

With adequacy of Model I in terms of diagnostics and valid restrictions we can proceed to study the implications of the model. As discussed, it is not so much the point estimates of $q_1$ and $q_2$ that are important, but rather whether the two special cases, $\lambda_1 \rightarrow 0$ and $\lambda_1 \rightarrow 1$, can be accepted or rejected. Again, if these two scenarios can be rejected, evidence is provided for the costly adjustment model as presented in this chapter. Following estimation, both special cases were strongly rejected, suggesting that $0 < \lambda_1 < 1$.

The primary result of Model I is that our representative agent encounters costs when attempting to formulate a consumption path. One explanation of these costs is the existence of
imperfect capital markets. These costs may arise from two distinct sources: the first cost is associated with the agent's inability to attain desired levels of consumption as predicted by the PIH, and the second cost is incurred as the agent attempts to alter his consumption path. Taken together, these two costs are found to have a significant impact on consumption. Proof of this significance is provided by our inability to reject the general model, $0 < \lambda_i < 1$.

Observationally, Model I suggests that the change in consumption will be a function of its lagged level and the expected change in permanent income. Results of Model I also show that this model displays traits similar to those found in the consumption literature, namely, lack of orthogonality, excess sensitivity, and excess smoothness. Model I thus produces results in keeping with observed facts, while offering an explanation as to how these properties may arise.

Even though Model I produced a strong case for the existence of costly adjustment, the level of knowledge embodied in this model is considered high. Impetus for Model II arose from the desire to introduce a model that imposes less knowledge upon the agent. As discussed, Model II assumes a simpler formulation for the expected change in permanent income, which is utilized in determining the change in consumption. Central to Model II is the assumption that the agent uses a simple linear rule in determining expected permanent income. The primary advantage of specifying a rule is that the formidable cross-equation restrictions that were implied by Model I can be eliminated, as can the extreme implicit informational assumptions. Model II represents the view that the agent may not have complete knowledge of the economic system and that consumption would be formulated given this possible ignorance.

Estimation of Model II produces results similar to those attained using Model I. Diagnostics show proper specification, although parameters may not be constant over the sample. We also note that the restrictions implied by Model II cannot be rejected. As in Model I, tests of $\lambda_i \to 0$ and $\lambda_i \to 1$ show that the general model cannot be rejected, implying that costs have a significant effect on consumption. As with Model I, the agent is seen in Model II to operate in an imperfect world. An important result of Model II is that its results are
consistent with those obtained using Model I, suggesting that even though less information is utilized by the agent, the assumption of some degree of naivety is not injurious in this particular framework.

Models I and II together suggest that the perfect market assumption used by the PIH is strong. Even though the precise sources of these imperfections remain unknown, the effect is clearly seen. Both models produce results found elsewhere in the literature, namely, that the change in consumption does depend on lagged information and that consumption is excessively smooth. However, these conclusions have been reached in a framework that allows for the presence of costs of adjustment. Obtaining an \( \lambda_1 \) in the interval \((0, 1)\) suggests the significant impact that these costs of adjustment have on the modeling of consumption. Model results suggest that market imperfections must be incorporated to ensure a true characterization of consumption.

Inclusion of the error correction terms also fails to alter the basic conclusion of the existence of costly adjustment. Although these terms are statistically significant, the two special cases, \( \lambda_1 \to 0 \) and \( \lambda_1 \to 1 \), can easily be rejected under both reformulations of Model I and Model II. The existence of costly adjustment as laid forth in this chapter is thus found regardless of whether we account for cointegration.

An important implication of the costly adjustment model deals with the possibility of altering consumption levels through government intervention. As discussed in the previous section, consumption levels may be altered through expected changes in gross labor income, capital income, and government expenditures. By changing the perceptions of the agent in regards to the future, government intervention may have a certain affect upon consumption. In a forward looking model as presented in this chapter, a wide range of policies may be found to impact upon the optimal consumption path. Additional latitude in influencing consumption may also exist by altering the costs of adjustment, \( \alpha_0 \) and \( \alpha_1 \), and thus the composition of current consumption.

The costly adjustment framework presented in this chapter was used to model market imperfections as generally as possible. This approach is believed to be important because
precise knowledge of imperfections may not be possible in a complex economy. By offering a framework that avoids specifying particulars we believe that a more realistic view of consumption determination is achieved. Estimation of the costly adjustment models, Models I and II, however, relied heavily upon the assumption of rational expectations that imposed some strong assumptions on our representative agent. Foremost of these assumptions is that the agent knows the complete expectation-generating system, including parameter values. Such an assumption seems incongruous to the work of this chapter, which seeks to allow the agent to possess less than perfect information about the economic environment. It seems contradictory to use the notion of costly adjustment as an acknowledgment that the agent possesses imperfect knowledge about the complex economy, only to implicitly assume that the agent knows the correct manner in which expectations are formed. Chapter 5 seeks to alleviate this contradiction by assuming that the agent may not possess complete knowledge on how expectations are generated.

In the following chapter, imperfect information is incorporated into the analysis of consumption by acknowledging that the agent may have less than perfect information of the system. It is assumed that knowledge is gained only as the agent learns about the system. Greater flexibility may be achieved in this case because the agent is able to recognize gradual changes in the economy, allowing changing perspectives to be incorporated into the consumption decision. By casting agents in a dynamic environment and allowing them to be cognizant of it, a more realistic view of consumption may be obtained.

Time-varying parameters are of central importance in the notion of learning presented in Chapter 5. Beyond providing a more realistic environment within which the agent operates, additional insights are offered into consumption determination. Both Models I and II strongly suggest that the general model \(0<\lambda<1\) is appropriate in modeling consumption. However, both models also display parameters that may not be constant over the sample. An interesting question thus arises in whether this strong acceptance reflects the true validity of the costly adjustment or the effects when model parameters are forced to be time-invariant in spite of large changes in the economic environment. One subsection in the following chapter studies
this important question. By permitting parameters to change and allowing such change to affect the determination of expectations and the coefficients in Equation (4.13), a more realistic characterization of consumption determination in the costly adjustment framework may be achieved.
CHAPTER 5: LEARNING

Introduction

In much of the work in consumption modeling, as well as in the material presented in the previous chapter, the rational expectations hypothesis (REH) was assumed for estimation. Although such an assumption is common, there are those who demur the ease with which it is made and invoked. The primary objection, which forms the basis of this chapter, deals with the level of knowledge the hypothesis presupposes upon the agent. Specifically, invocation of the REH implicitly assumes that the agent knows the complete system including parameter estimates. In deriving Model I, this condition implies that, beyond knowing the auxiliary system and its parameters, the agent is also cognizant of the cross-equation restrictions. Explanation as to how this knowledge is achieved is not offered.

Although this informational assumption may be appropriate for a system that has been operating indefinitely, it seems questionable when viewed in conjunction with sample sizes typical of macroeconomic series. The alternative approach for expectation formation explored in this chapter does not assume some a priori level of knowledge on the part of the agent; rather, we allow the agent to “learn” about the system over time.

To proceed in this direction, we assume that the agent uses a simple linear rule to proxy the more complicated rational expectation equilibrium (REE). Based on the work in the previous chapter, Model II may be thought of as providing such a rule as an alternative to estimating Model I. Estimation thus avoids the complex structure associated with the REE, requiring less analytical skills on the part of the agent. By allowing coefficients of the linear rule to vary over the sample, learning is incorporated by assuming that the agent uses some
type of recursive least squares estimation technique. Parameter estimates is formulated contingent upon the level of information available at that particular moment.

Modeling consumption within this learning framework potentially allows for more interesting and realistic results. By permitting model parameters to be time-varying the effects of an altering economic system can be incorporated into the analysis of consumption. Given the tumultuous change which the U.S. economy displayed between 1957 through 1993, this generalization appears potentially important to this particular analysis.

As modeled in the previous chapter, costly adjustment lends itself to the notion of learning because the model parameters $q_1$ and $q_2$ (or $q_1, \alpha_1, \alpha_2, \alpha_3$ in the case of Model II) would be expected to vary as the sample progressed. An evolving financial system and changing real interest and discount rates are possible explanations. By explicitly permitting these parameters to vary, we allow the agent is allowed to incorporate information relative to the consumption decision as it becomes available.

Learning also appears highly relevant given the evidence against time-invariant models found in the previous chapter. Models I and II both relied heavily on the generation of expectations derived from a time-invariant auxiliary system. If this system is in fact time-varying, application of the REH as in the previous chapter is incorrect because substitution of the unknown expectations requires coefficients to be constant over the sample. Once coefficients are permitted to vary, the REE itself will be observed to change over the sample. The learning model introduced in this chapter allows for the possibility of a time-varying REE and incorporates it into the modeling of consumption. By explicitly modeling expectations in a manner that accounts for possible change, analysis can be performed in a more flexible environment.

To best develop and implement this notion of learning, this chapter is organized as follows. The next section, entitled “Rational Expectations,” presents a brief review of the REH, where informational requirements and econometric implications are discussed. It is the informational assumptions, and not the implications, made by the REH that are the point of
contention in this chapter. Two different approaches to learning are offered as alternatives to the informational assumptions of the REH.

Implementing some degree of learning in the modeling of consumption, provides the basis of the third section, entitled “Learning.” Following the work of Friedman (1979), Hall (1993), and Bullard (1992) a rule with time-varying parameters is specified and estimation proceeds with a form of recursive least squares. Learning is incorporated in this estimation procedure because for any given period estimated coefficients will only reflect information up to that point. Beyond reviewing previous approaches at formulating learning models, this paper offers an alternative view by allowing the system to be subject to continual change. Recursive least squares estimation, however, will be inappropriate in this type of environment. Implementation of the alternative learning rule considered in this chapter relies on the Kalman filter to provide optimal updating for parameters that are subject to continual change. A brief review of the Kalman filter is presented in a subsection.

The section, entitled as “The Learning Model,” lays forth an explicit formulation of the learning model. Estimation results and diagnostics reveal the appropriateness of the model in describing consumption behavior. Given an appropriate model, as determined through the diagnostics, various results are considered. First, we study the behavior of estimated coefficients. Plots of these estimates reveal that although a number of coefficients converge to certain levels given enough time, others do not. This situation casts doubts as to whether any learning rule will converge to the REE. More important, however, this coefficient instability raises the question of whether the notion of steady REE is even relevant. If some coefficients are truly time-varying, the REE itself would shift through time, casting doubt upon the assumption of rational expectations as used in the previous chapter. The conclusions reached in this chapter will prove very useful in ascertaining the legitimacy of previous work.

A second result of this chapter deals with testing the two special cases, discussed in the previous chapter, while allowing for a time-varying framework. Given the possibility of changes in $q_1$ and $q_2$, which are coefficients of the consumption equation, and the possibility that expectations utilized in estimation are wrong, the strong acceptance of the general model
found in the previous chapter is questioned. We consider one simplistic test of the general model in the section “Testing Special Cases.” Conclusions from implementing the learning model are offered in the final section.

Rational Expectations

The Rational Expectation Hypothesis (REH) has had a lasting impact since its introduction by Muth (1961). It was Muth’s belief that the agent forms expectations based on the “true” structural model of the economy (Pesaran 1988). In particular, the assumption is made that the individual’s perception of the probability distribution of future outcome, conditional on the available information, coincides with the actual distribution, conditional on that information.

Two assumptions about how the agent acts in this paradigm are made. First, agents exploit all information until the marginal cost of acquiring it equals the marginal benefits of improving forecasts. Second, certain information, such as the structure of the model and the values of certain exogenous variables, is freely available. These two assumptions imply that the rational agent uses all available information on structure and exogenous variables to make forecasts (Demery 1983, p. 243).

Under this expectations scheme, the economic agent will forecast the value of a variable in accordance with the actual process by which that variable is determined, using all the information that is available at the time forecasts are made (Demery 1983, p. 236). Application is found in the previous chapter, where the assumption of rational expectations was utilized by assuming that the agent understands the structure of the joint model and the necessary parameters, which enabled the unknown expectations to be solved out. This is the standard method by which expectations have been handled in contemporary macroeconomic models.
Implications of the REH can be noted by allowing $\Omega_{t-1}$ to denote the information set at time $t-1$, which is assumed to contain all relevant data for the economic model. Conditional on this information set, the joint probability distribution of the variable entering the agent's perceived model is given by $f(x_t|\Omega_{t-1})$, where the subjective conditional distribution on $x_t$ coincides exactly with the objective conditional distribution of $x_t$ (Pesaran 1988, p. 24). With these assumptions and given that $S_{t-1}$ is a subset of $\Omega_{t-1}$ ($S_{t-1} \subseteq \Omega_{t-1}$), Pesaran derives the following statistical properties of the rational expectations hypothesis (pp. 25-26):

a) $E(\epsilon_t|\Omega_{t-1}) = 0$  
   "unbiasedness property"

   where $\epsilon_t = x_t - E(x_t|\Omega_{t-1})$

b) $E(\epsilon_t|S_{t-1}) = 0$  
   "orthogonality property"

c) $E(\epsilon_t | x_{t-1}, ... ) = 0$  
   "efficiency property"

d) $E(\epsilon_t \epsilon_{t-i}) = 0$  $i \geq 1$

   $E(\epsilon_t \epsilon_{t-i}) = 0$  $i \leq 1$  
   "lack of serial correlation"

Orthogonality means that the agent will use all available information to make forecasts and will not ignore any information that may be useful in forecasting. Serial independence implies a zero correlation between forecast errors; that is, agents cannot learn from past forecast errors (Irwin and Thraen 1994, p. 137).

In studying the REH, the efficient incorporation of information is not a point of contention. What is questionable is how the agent attains the level of information which the REH presupposes. Consider, for example, Model 1 from the previous chapter. By supposing that the agent knows the system and the necessary parameters the unknown expectations can be determined. How this level of information is achieved and how the necessary cross-equation restrictions are formulated and implemented are not stated. Barring an omniscient being, it is hard to explain how an agent could be posited as simply possessing this knowledge. As Friedman (1979) states:
Recently many macroeconomists have turned in this context to Muth (1961) conception of "rational" expectations, according to which economic agents form their expectations as if they know the process which will ultimately generate the actual outcomes in question. . . . What is typically missing in rational expectations models, however, is a clear outline of the way in which economic agents derive knowledge which they then use to formulate expectations meeting this requirement. (p. 23)

Much of the theoretical and empirical work on the macroeconomic applications of the REH has approached this question by simply assuming that the agent has been operating within the economy sufficiently long so as to have discovered the coefficients which determine its structure (Demery 1983, page 245). With knowledge of the model structure, expectations could be solved as in Chapter 4. Friedman (1979) and Hall (1993) consider such an informational situation as constituting a steady state since no additional knowledge is required on the part of the agent, while Pesaran (1988) calls this state an rational expectations equilibrium (REE). Regardless of title, no incentive exists on the part of the agent to alter his beliefs about the economic environment (Pesaran 1988, p. 33).

The work in this chapter approaches expectation generation given the assumption that the system has not been operating ad infinimum. This generalization will permit the introduction of an alternative method by which the necessary expectations can be generated. A method that will allow the agent to "learn" about the system as the sample progresses.

Beyond providing a more pragmatic view of the economy in general, this alternative also appears more in keeping with the data used in this paper. As recalled from the previous chapter, the hypothesis of constant parameters was strongly rejected for both Models I and II. These changes in parameters may be attributed to structural changes in the U.S. economy (oil shocks, for example) or possibly to the gradual change associated with an evolving economy. Whatever the causes, economic agents must learn the new structure as the sample unfolds
Simply assuming that agents know a time-invariant system is potentially misleading because if agents confronted with a structural change continue to use a previous REH forecasting rule, the optimality properties discussed previously will not necessarily hold. The optimality properties of the REH are conditional on a properly specified model (Irwin and Thraen 1993, p. 137). Expectations formed from a model that is misspecified relative to parameterization would clearly not be rational.

The alternative methodology presented next allows the agent to formulate parameter estimates as the sample progresses, allowing for the incorporation of gradual change into the determination of optimal consumption. It should be noted, however, that the agent is assumed to know the correct functional form of the system. Potential uncertainty only arises with parameter estimates.

The Learning Model

Contrary to the informational assumptions made by the REH, we now consider the situation where a steady state is not assumed a priori. In this situation, the exact manner in which the agent is assumed to learn is of utmost importance. Bray and Savin (1986) define two types of learning (p. 1130):

1) Rational Learning

Under this scheme, the agent is assumed to formulate expectations based on a correctly specified model, but where a finite number of parameters may be unknown. Rational learning then centers on estimating these unknown parameters in an interactive setting where there is feedback from the incorrectly estimated parameters to the actual parameters (Pesaran 1988, pp. 34-36). Following Pesaran, however, two provisos exist. First, except for the unknown parameters, the true economic model is taken to be common knowledge. Second,
although it is assumed that agents know the true equilibrium relations in the economy, no explanation is offered as to how this information was obtained.

(2) Bounded Rational Learning

This notion of learning is generally weaker than the preceding type. Here, agents are not required to know the structural equilibrium relations. Instead agents use “reasonable” rules of learning to which they remain committed over the whole period that learning takes place (Pesaran 1988, p. 35). The rule used is often appropriate in the REE but misspecified when there is learning. Pesaran notes two shortcomings connected with this approach. First is that what constitutes a reasonable rule is not explained, nor is it explained how agents collectively choose this rule. Second, the approach does not permit for revision of the learning rule. Most researchers who use bounded rational learning models tend to specify a rule a priori. The learning rule adopted in these models necessitates that agents know the reduced form equations of the true model, except for unknown parameters (Pesaran 1988, p. 37).

The primary difference between rational learning and bounded rational learning models is related to the amount of a priori information agents are assumed to possess. Under the former, the agent is assumed to know the true structural relations of the economy, and under the latter only the reduced-form equations are required to be known. It is in the sense of using knowledge of reduced-form relations instead of the structural-form relations that bounded rational learning makes a less demanding assumption about the level of knowledge and analytical abilities that the agent possesses (Pesaran 1988, p. 36).

To incorporate learning in this chapter, a variant of bounded rational learning is utilized. In particular, attention is focused on adaptive learning where agents are assumed to use an intuitive procedure for making and changing their choices on the basis of past outcomes. Uncertainty arises only in terms of parameter estimates, while functional forms are assumed to be known. For this purpose, we start with the assumption of a statistical or econometric procedure that is used for estimating a perceived law of motion for the variables we seek to forecast. Forecasts are then computed using the law of motion and compared to the actual value once it becomes available, thus providing a new data point. Agents then
reestimate the perceived law of motion using the additional observation in their data set and so on. The system is seen to operate "on line" (Honkapohja 1993, p. 588). The mechanics behind this approach will become evident in the work to follow.

Honkapohja (1993, p. 588) classifies these procedures as bounded rational in that the agent uses a model that is misspecified while he is learning; that is, outside the REE. The REE will only be achieved if the sequence of parameter values converges to some appropriate value.

We now present an example from Honkapohja to more clearly illustrate how adaptive learning has been implemented in economic modeling. Honkapohja examines the system (p. 589):

\[ y_t = \mu + A E_{t-1} y_t + C w_t \]
\[ w_t = S w_{t-1} + v_t \]

where \( y \) is an \((n \times 1)\) endogenous vector, \( w \) is an observed \((p \times 1)\) vector of exogenous variables, and \( v \) is a \((p \times 1)\) vector of white noise shocks. Assuming rational expectations this system can be expressed as:

\[ y_t = \bar{a} + \bar{b} w_{t-1} + \eta_t, \quad \text{where} \]
\[ \bar{a} = (I - A)^{-1} \mu, \quad \bar{b} = (I - A)^{-1} CS, \quad \text{and} \quad \eta_t = (I - A)^{-1} Cv_t. \]

By assuming rational expectations we take for granted that the agent understands the manner in which the terms \( \bar{a}, \bar{b}, \text{and} \eta_t \) (which represent the REE) are formulated when \( y_t \) is determined. As discussed, no explanation is offered as to how the agent achieves this level of knowledge, or even whether it is plausible that such matrix manipulations could be performed. Honkapohja, however, approaches the system in terms of a bounded rational learning problem by assuming that the agent has the following perceived law of motion at time \( t \)

\[ y_t = a_{t-1} + b_{t-1} w_{t-1} + v_t, \]

which is used in generating \( E_{t-1} y_t \). Coefficients are allowed to vary, meaning that the determination of \( E_{t-1} y_t \) requires estimates of \( a_{t-1} \) and \( b_{t-1} \), which are not known a priori. Learning is explicitly incorporated into this model by allowing the agent to build up knowledge about these unknown coefficients as the sample progresses. One practical method
of incorporating this notion is with some form of recursive least squares, where coefficient estimates at a given period are formed using only information available at that particular time. Estimation in this manner formalizes the idea that expectations formed at time $t$ should only be based on information available at time $t$ and before. This approach compares to the rational expectations equilibrium of Equation (5.2), as represented by $\bar{a}$ and $\bar{b}$, which in essence assumes that expectations, at any point in the sample are formed using parameter estimates derived from the complete sample.

Estimation of Equation (5.3) proceeds with the recursive least squares algorithm

\begin{align*}
\theta_t &= \theta_{t-1} + \frac{1}{t} R_{t-1}^{-1} Z_{t-1} e_t, \\
R_t &= R_{t-1} + \frac{1}{t} (Z_{t-1}' Z_{t-1} - R_{t-1}),
\end{align*}

(5.4)

where $\theta_t = (a_t, b_t)$, $Z_t = (1 w_t')$, and $e_t = y_t - \theta_{t-1}' Z_{t-1}$. Parameter estimates of the perceived law of motion $\theta_t$ are noted to be updated at time $t$ using the latest vector of forecast errors, as measured by $e_t$ (Honkapohja 1993, p. 590). Estimates of $a_t$ and $b_t$ are seen to incorporate information only available at time $t$, implying that expectations are more consistent with available information as compared to expectations formed using the REH, at least as discussed by Marcet and Sargent (1989).

A question of importance with this example, and to learning rules in general, is whether the rule expressed by Equation (5.3) will ultimately achieve the REE, as summarized in Equation (5.2), given sufficient time. In this particular example, the learning rule will approach the solution as implied by the REH if $a_t \rightarrow (I-A)^{-1} \mu$ and $b_t \rightarrow (I-A)^{-1} CS$ as $t \rightarrow \infty$.

Convergence theorems, as will be discussed, provide requirements under which convergence of the learning rule to the REE may occur. If it is found that $a_t$ and $b_t$ fail to approach some stable level, the learning rule will not lead the agent to the REE.

In studying convergence, Honkapohja (1993) applies results from Marcet and Sargent (1989). These authors examine a class of updating schemes where convergence to the REE is guaranteed, if certain conditions are met. Building on the work of Ljung (1977) and Ljung and
Soderstrom (1983), Marcet and Sargent show that a particular type of recursive least squares procedure, used in obtaining parameter estimates, possesses the required convergence property. Without going into too much detail, this technique can be discussed by supposing that an agent attempts to estimate:

\[ y_i = \theta^* \varphi(t) + \nu(t), \]

where \( \varphi(t) \) denote all the right-hand side variables, \( \nu(t) \) is some error term, and \( \theta \) represents a column vector of parameters. Ljung and Soderstrom (1983) specify the following recursive algorithm that allows for the generation of a time-varying estimate of the vector \( \theta \):

\[
\tilde{\theta}(t) = \tilde{\theta}(t - 1) + \frac{1}{t} R^{-1}(t) \alpha_t [y(t) - \tilde{\theta}^* (t - 1) \varphi(t)] \\
R(t) = R(t - 1) + \frac{1}{t} \alpha_t \varphi(t) \varphi(t)' - R(t - 1),
\]

(5.5)

where \( t \) denotes the particular period, \( \alpha_t \) is some positive nondecreasing sequence with the property that \( \alpha_t \rightarrow 1 \) as \( t \rightarrow \infty \), and \( R(t) = \sum_{i=1}^{t-1} \alpha_i \varphi(i) \varphi(i)' \). The product \( \frac{\alpha_t}{t} R^{-1}(t) \) is called the gain, a term which achieves significance once the Kalman filter is introduced. Marcet and Sargent (1989) show that one condition that must be fulfilled to ensure convergence is for the gain term to approach zero as the sample progresses, which is very important to the analysis. As \( t \rightarrow \infty \) system (5.5) shows that \( \tilde{\theta}(t) = \tilde{\theta}(t - 1) \), implying that \( \tilde{\theta}(t) \) will ultimately become time-invariant. Convergence theorems as discussed in Ljung and Soderstrom (1983) and Marcet and Sargent (1989) depend crucially on this assumption in order to present a tractable methodology in which the possibility of convergence can be studied. One simple application of this idea is presented in Appendix D. Further discussion dealing with the convergence of system (5.5) will be considered. It is also noted that the recursive least squares algorithm as specified by Marcet and Sargent differs from that used by Friedman (1979) by the addition of the term \( \{1/t\} \) and an \( \alpha_t \) that approaches 1.

Bullard (1992) offers an alternative view of system (5.5) that permits this method to be directly comparable to the one used later in this paper. [Ljung and Soderstrom (1983) present
an analogous modification). Bullard (1992, p. 163) shows that system (5.5) can be reexpressed as

\[ \theta_{t-i} = \theta_{t-i-1} + P_{t-i} \phi_{t-i} f_{t-i}^{-1} [ y_t - \phi_{t-i} \theta_{t-i} ] \]

\[ P_t = P_{t-i} - P_{t-i} \phi_{t-i} \phi_{t-i} P_{t-i} f_{t-i}^{-1} \]

where the scalar \( f_{t-i} = \left( \frac{1}{\alpha_i} \right) + \phi_{t-i} P_{t-i} \phi_{t-i} \) and \( P_i \) is defined as \( P_i = \frac{1}{t} R_i^{-1} \).

This modification expresses system (5.5) in terms of the Kalman filter recursions to be discussed later.

With these preliminaries, we now consider how the notion of adaptive learning may be applied to the modeling of consumption as presented in Chapter 4. Using Equation (4.9), the change in consumption can be expressed as

\[ \Delta c_t = q_1 \Delta c_{t-i} + q_2 \epsilon_{t-i} \Delta y_{pt} + \omega_t \]

where \( y_{pt} \) denotes permanent income and \( \omega_t \) is a white noise disturbance. From the previous chapter, the rational expectations equilibrium for Model I can be expressed as

\[ \Delta c_t = q_1 \Delta c_{t-i} + q_2 [0 1 0 -1 0] (1-pA)^{-1} AX_{t-i-1} + q_2 [0 0 1 0 0] AX_{t-i-1} + \epsilon_t \]

where \( A \) and \( X_t \) are as previously defined. However, the change in consumption as expressed by Equation (5.8) relies heavily upon the assumption that the agent correctly understands the system. To specify a learning rule to mimic the behavior of Equation (5.8), note that the right-hand side is a function of the variables \{ \Delta c_{t-i}, \Delta y_{pt}, \Delta y_{xt-1}, \Delta y_{xt-2}, \Delta y_{xt-3}, \Delta y_{kt-1}, \Delta y_{kt-2}, \Delta y_{kt-3}, \Delta g_{t-1}, \Delta g_{t-2}, \Delta g_{t-3} \}. This can be seen by carrying out all the implied matrix operations and recalling from Chapter 4 that the vector \( X_t \) is comprised of these variables. Since \( \Delta c_t \) in Equation (5.8) is a function of such variables we specify the following perceived law of motion

\[ \Delta c_t = a_{01} + a_{11} \Delta y_{x-1} + a_{21} \Delta y_{x-2} + a_{31} \Delta y_{x-3} + a_{41} \Delta y_{xt-1} + a_{51} \Delta y_{xt-2} + a_{61} \Delta y_{xt-3} + a_{71} \Delta g_{t-1} + a_{81} \Delta g_{t-2} + a_{91} \Delta g_{t-3} + \epsilon_t \]
where all coefficients are permitted to vary and a time-varying intercept has been added to pick up any changes in the mean. Equation (5.9) is of a form analogous to that used by Honkapohja (1993). However, derivation of Equation (5.9) is a bit more complex than that used by Honkapohja. Equation (5.9) may be thought of as arising from three separate rules used to generate \( \{E_{t+1}\Delta y_t, E_{t+1}\Delta y_k, E_{t+1}\Delta g_t\} \), which may be formed from a time-varying auxiliary system, Equation (5.16) below. For example, we might specify the following three rules to determine the three expected changes:

\[
\Delta y_t = b_{20t} + b_{22t}\Delta y_{t-1} + b_{23t}\Delta y_{t-2} + \ldots + b_{25t}\Delta y_{t-k} + \ldots + b_{210t}\Delta g_{t-3},
\]

\[
\Delta y_k = b_{30t} + b_{32t}\Delta y_{t-1} + b_{33t}\Delta y_{t-2} + \ldots + b_{35t}\Delta y_{t-k} + \ldots + b_{310t}\Delta g_{t-3},
\]

\[
\Delta g_t = b_{40t} + b_{42t}\Delta y_{t-1} + b_{43t}\Delta y_{t-2} + \ldots + b_{45t}\Delta y_{t-k} + \ldots + b_{410t}\Delta g_{t-3}.
\]

By substituting these rules into Model I (since the infinite expected sums can be expressed as some functions of \( \{E_{t+1}\Delta y_t, E_{t+1}\Delta y_k, E_{t+1}\Delta g_t\} \) of the previous chapter while reorganizing and changing notation, Equation (5.9) can be determined. We make explicit statement of this result because such knowledge would be necessary in discussing convergence of this learning rule to the REE, which is to be presented below.

Estimation would proceed by specifying an auxiliary system comprised of a VAR(3), using the variables \( \{\Delta y_t, \Delta y_k, \Delta g_t\} \), and utilizing joint estimation of Equation (5.9) with this VAR(3) system. Joint estimation is believed to be appropriate in accounting for any correlations that may exist between \( \Delta c_t \) and \( \{\Delta y_t, \Delta y_k, \Delta g_t\} \). If parameters are estimated using the method suggested by Marcet and Sargent (1989), and summarized by system (5.5), the rule expressed by Equation (5.9) may converge to the REE as the sample progresses.

Convergence, which depends on the model and the data, has been described by Sargent (1993) as a fixed point in the mapping from the perceived law of motion to the actual. Marcet and Sargent (1989), building on the work of Ljung and Soderstrom (1983), show that the existence of such a point can be determined by studying the behavior of the system of ordinary differential equations, assuming system coefficients are stable. Convergence to the REE is obtained if a rest point exists in this system of differential equations Sargent (1993, p. 126).
Details of this procedure are found in Ljung and Soderstrom (1983) and Sargent (1993). A simple application is presented in Appendix D at the end of the paper.

Convergence of Equation (5.9) to the REE can be shown by substituting the three rules noted above into Model I [Equation (4.13)] and assuming that the coefficients in these rules are stable. Following Honkapohja (1993) and Marcet and Sargent (1989) we may study convergence by examining the stability of a system of differential equations composed of the system coefficients. Examining convergence in terms of differential equations depends critically upon the stability of the coefficients. Avoiding details, it can be shown that once this nonlinear system of differential equations is linearized (see Chapter 2, Leonard and Van Long (1992)) around the coefficient estimates of Chapter 4, convergence may be easily studied. In particular, assuming a real interest rate of 2 percent ($p = 0.98$), as in Chapter 4, stability of the system is found when $q_2$ is in the range (0 0.4). It is recalled that in Chapter 4 $q_2$ was found to be 0.221, suggesting that under the assumption of constant coefficients the learning rule (5.9) will ultimately converge to the REE. We emphasize the assumption of coefficient stability since, as will be shown below, many of the rule coefficients are anything but stable.

Explicit discussion of a system of differential equations in studying convergence to the REE may, however, not be necessary. Sargent (1993, p. 131) states in proposition (b) that if estimation using the approach set forth in system (5.5) produces parameter estimates [and estimates of $R(t)$] which converge to some level, they will in fact converge to the rest points of the differential equation. Thus, parameter convergence would imply attainment of the REE, at least when the system is estimated using the approach of Marcet and Sargent (1989), suggesting that the evolution of coefficients is very important to the analysis. Because of this importance, we now focus our attention on how the estimation method of Marcet and Sargent (1989) performs under differing circumstances. We start by considering system (5.5), which shows that as $t \to \infty$ coefficients will ultimately evolve as

\[ \beta_t = \beta_{t-1}, \]
\[ R_t = R_{t-1}. \]
Marcet and Sargent (1989, p. 341) further show that as \( t \) becomes large the estimate of \( \beta_t \) will be the weighted OLS estimator

\[
\beta_t = \left( \sum_{i=1}^{t-1} \alpha_i x_{t-i} y_{t-i} \right)^{-1} \left( \sum_{i=1}^{t-1} \alpha_i x_{t-i} y_{t} \right)
\]

where \( \alpha_i \to 1 \) as \( i \to \infty \) [e.g., \( \alpha_i = \frac{i}{1+i} \)].

Coefficients estimated in this manner, however, are seen to allow only for gradual evolution in the estimated coefficient. Continual change may thus be missed if this type of estimate is used. To examine this point, we present a simple simulation which assumes that the true parameter follows a random walk process and compare how estimation utilizing Equation (5.11) performs in such a situation. In this exercise, we assume that the true parameter evolves as

\[ \beta_t = \beta_{t-1} + \eta_t, \]

where \( \eta_t \) is iid N(0,0.01).

It is further assumed that \( x_t \) is randomly drawn from a distribution which is N(0,16), and that \( y_t \) is formed as:

\[ y_t = \beta_t x_t + \epsilon_t \]

\[ \epsilon_t \text{ iid N(0,0.005)}. \]

With the generated vectors \( y_t \) and \( x_t \), we now use Equation (5.11) to obtain coefficient estimates for each period (1000 observations are used). Estimation shows that in this dynamic environment, estimates of \( \beta_t \) could not change fast enough; that is, coefficient estimates are only able to incorporate change in the system very slowly. This distinction between the true coefficient, \( \beta_t \), and its estimate \( \beta_t \) is important. In the following work \( \beta \) represents that correct, yet unknown, coefficient that we seek to estimate. To visualize these differences, we plot the actual parameters, \( \beta_t \), and estimates denoted as MS, \( \beta_t \), are presented in Figure 5.1.
This plot shows that although the weighted recursive OLS estimates vaguely follow the variation in the true coefficients, the large movements that occurred in $\beta_t$ over the sample are missed. Although this is not surprising given the manner in which $b_t$ is estimated, it does raise the question about how such an estimation procedure would perform if parameters are subject to continual change. In particular, by specifying $\beta_t$ as a random walk process, we would expect the coefficient to wander throughout the sample. For the particular draws used in establishing the plot in Figure 5.1, however, the recursive least squares estimator of Marcet and Sargent (1989) seems to converge to some level. This appears to be problematic because if this coefficient estimate $b_t$ is used in a learning rule, we may conclude that it converged, implying that the learning rule has attained the REE. In actuality, such an equilibrium may itself shift because the true coefficient, $\beta_t$, will be changing continuously throughout the sample.

To make this discussion a bit more concrete, assume that the simulation equation $y_t$ is actually a learning rule for the simple model

$$y_t = gE_{t-1} y_t + x_t + \varepsilon_t.$$  

The REE corresponding to this model is simply
\[ y_t = \frac{1}{1-g} x_t + \epsilon_t, \]

which is liable to OLS estimation using all the data in the sample. Performing this calculation shows that the coefficient estimate is approximately 5. In Appendix D, we show under what conditions this learning rule will converge to the REE of the model. Again, these conditions are contingent upon estimation using system (5.5). The important result of Appendix D is that given \( g < 1 \) or \( \frac{1}{1-g} > 1 \), convergence to the REE is possible.

The simple simulation exercise just presented suggests that coefficient estimates \( b_i \) converge to some level represented by the REE, \( \frac{1}{1-g} \), which is again the OLS estimate of \( y_t \) on \( x_t \). Coefficient convergence thus implies that the learning rule ultimately will obtain the rational expectations solution. What is curious in this result is that we know that the true relationship between \( y_t \) and \( x_t \) is contingent upon a coefficient, \( \beta_t \), which follows a random walk and is thus likely to wander throughout the sample, implying that the REE would also shift through time. We would expect the REE to alter through time because we know that the correct relationship between \( y_t \) and \( x_t \) is \( y_t = \beta_t x_t + \epsilon_t \), implying that \( \frac{1}{1-g} \) is time-varying because \( \beta_t = \frac{1}{1-g} \).

These results above suggest that when coefficients that are liable to continuous change, the convergence results based on Marcet and Sargent (1989) may be inappropriate. Application of this estimation technique may force convergence of coefficient estimates, given enough time, when in fact the true coefficients (denoted as \( \beta_t \)) are not stable. To some extent, using the simple example discussed earlier, estimation with system (5.5) appears to presuppose the answer. While such techniques may be appropriate for slowly evolving systems (\( \beta_t = \beta_{t-1} \)), estimation and analysis of convergence may give misleading results in the situation when coefficients evolve as random walks.
Alternatively, we can also consider how estimation using Equation (5.11) may perform when the variance associated with the evolution of parameters is very small; that is, when coefficients evolve approximately as $\beta_t = \beta_{t-1}$. To study this situation, a simple simulation exercise will be carried out as before, but where it is assumed that $\text{Var}(\eta) = 1E-06$.

Generating $x_t$ and $\varepsilon_t$ as previously described and using Equation (5.11) to produce coefficient estimates, we can generate Figure 5.2, which plots actual and estimated parameters.

![Figure 5.2. Coefficient plots: $y_t=\beta_0 x_t+\varepsilon_t$, $\beta_t=\beta_{t-1}+\eta_t$, $\text{Var}(\eta)=1E-06$](image)

Figure 5.2 shows that when coefficients evolve slowly, estimates produced by Equation (5.11) tend to track the true coefficient quite well. Calculating the observed REE of this simple model produces an OLS estimate of approximately 5.005, indicating that the estimated coefficients approach those associated with the REE. Contrary to the previous situation, learning ultimately will reach the REE, which itself will be approximately steady through time because the change in $\beta_t$ is slight. This result is not really surprising, however, because in the present case $\beta_t = \beta_{t-1}$.

These simulation exercises reveal that the evolution of the true coefficients is very important in any discussion of convergence. If true coefficients evolve such that $\beta_t = \beta_{t-1}$,
coefficient estimates will obtain some stable level, implying that the learning rule will converge to the REE. The REE in this situation will also be steady through time, suggesting that the REH is appropriate. Alternatively, if the variance associated with $\beta$ is relatively large, these true parameters will wander extensively throughout the sample. As shown, however, there is a possibility that estimation using the method of Marcet and Sargent (1989) may produce coefficient estimates that appear to converge to some level. Although such convergence may seem to imply that the learning rule has found the REE, this conclusion is not correct. If the true coefficients are truly time-varying, the REE itself will shift through time. Thus, even though coefficient estimates may obtain the REE as implied when coefficients are assumed to be time-invariant, this result is meaningless because in reality the REE itself is not constant once coefficients are allowed to vary. Estimation using Marcet and Sargent in this particular case seems to impose some naivety upon agents because they will continue to believe that they have attained the REE, in spite of large estimation errors. Any changes in such beliefs will come very slowly.

Because of the possibility that model parameters may be subject to a great deal of change over the sample, we offer an alternative estimational rule and estimation method. As discussed, if the coefficients do not reach some stable level, convergence to the REE will not be achieved, at least as studied by Marcet and Sargent (1989). Also as mentioned, Equation (5.9) may be thought of as arising from three separate learning rules that may be used in determining $\{E_{t-1} \Delta y_t, E_{t-1} \Delta y_{kt}, E_{t-1} \Delta g_t\}$. If any elements of these rules fail to achieve a steady state, the change in consumption as expressed by Equation (5.9) will not achieve the REE. Explicit discussion of testing whether the coefficients in the auxiliary system display such variability is deferred for now; currently we will use only the results from this analysis. As will be seen, a number of coefficients in these rules (which are the equations of the auxiliary system previously discussed) may evolve as random walks, suggesting that they are subject to continual change over the sample. By having coefficients that behave in such a way, Equation (5.9) will not converge to the REE because many of the coefficients fail to achieve the required stable level. If some coefficients are truly time-varying, however, convergence to the
REE is meaningless because such an equilibrium would shift throughout the sample. In considering Model I from the previous chapter, for example, if some of the variables in the auxiliary system are really time-varying, the REE of the model will change every time these underlying coefficients change. More important, application of the REH in this environment is questionable because substituting for the unknown expectations as performed in the previous chapter relied on the constancy of coefficients. The simplicity and tractability of the REH are lost once the possibility of time-varying coefficients is accepted.

In the situation where some coefficients may be time-varying, the generation of expectations is of vital importance. Because the REH is not strictly appropriate, an alternative method must be used which will explicitly allow for changes in the expectation generating mechanism. For this purpose, we offer an alternative learning rule which incorporates expectations directly; in particular, we will use the rule

$$\Delta c_t = q_{1t} + q_{2t} \Delta c_{t-1} + \alpha_{1t} E_{t-1} \Delta y_{ht} + \alpha_{2t} E_{t-1} \Delta y_{kt} + \alpha_{3t} E_{t-1} \Delta g_t + \epsilon_t^e.$$  

This rule is noted to be a time-varying version of Model II from the previous chapter, augmented with a time-varying intercept. Change is incorporated into the evolution of $\Delta c_t$ through the parameters $q_{1t}, q_{2t}, \alpha_{1t}, \alpha_{2t},$ and $\alpha_{3t}$ and also through $\{E_{t-1} \Delta y_{ht}, E_{t-1} \Delta y_{kt}, E_{t-1} \Delta g_t\}$. A time-varying auxiliary system, specified later, is used in generating the necessary expectations. By explicitly incorporating expectations that are generated by a time-varying system, any changes in the economy can be directly utilized in estimating the change in consumption.

It should be noted, however, that Equation (5.9) and Equation (5.12) may behave quite closely. To see this assume that the time-varying auxiliary system can be adequately represented by a VAR(3), which appears as system (5.16). Substituting these expectations into Equation (5.12) shows that:

$$\Delta c_t = (q_{0t} + \alpha_{1t} b_{20t-1} + \alpha_{2t} b_{30t-1} + \alpha_{3t} b_{40t-1}) + q_{1t} \Delta c_{t-1} + (\alpha_{1t} b_{21t-1} + \alpha_{2t} b_{31t-1} + \alpha_{3t} b_{41t-1}) \Delta y_{ht-1} + (\alpha_{1t} b_{22t-1} + \alpha_{2t} b_{32t-2} + \alpha_{3t} b_{42t-1}) \Delta y_{kt-2} + \ldots + (\alpha_{1t} b_{28t-1} + \alpha_{2t} b_{38t-2} + \alpha_{3t} b_{48t-1}) \Delta g_{t-1},$$

or, with $\gamma_t$ suitably defined as

$$\gamma_{ht} + q_{1t} \Delta c_{t-1} + \gamma_{2t} \Delta y_{ht-1} + \gamma_{3t} \Delta y_{kt-2} + \ldots + \gamma_{8t} \Delta g_{t-3},$$
implying that when system coefficients are stable (that is, $b_{ji} = b_{jk}$ for large enough $t$) rules (5.9) and (5.12) will not be very different.

With the possibility of large changes in the system, we also specify an alternative method in which Rule (5.12) may be estimated. As discussed, when the system is subject to continual changes, recursive least squares will not be appropriate. Instead, we assume that coefficients evolve as random walks and utilize the Kalman filter in estimation. The appropriateness of the random walk will be discussed below, where tests are performed to see whether this assumption is warranted. Following a discussion of the Kalman filter, we find that rule (5.12) is very amenable to estimation.

By specifying the change in consumption as rule (5.12), we explicitly allow the coefficients $q_{0t}$, $q_{1t}$, $\alpha_{1t}$, $\alpha_{2t}$, and $\alpha_{3t}$ to vary over time. In the previous chapter, it was noted that the parameters of Model II were not deep; that is, the estimated parameters $\{q_{0t}, q_{1t}, \alpha_{1t}, \alpha_{2t}, \alpha_{3t}\}$ are functions of the underlying parameters $\lambda$, $r$, and $D$ (the deep parameters) and also the expectational parameters from matrix $A$. This implies that the model is liable to the Lucas critique (Lucas 1976) because the parameters in this particular specification will alter every time the marginal processes of $\Delta y_{lt}$, $\Delta y_{kt}$, $\Delta g_{lt}$, $\lambda$, $r$, and $D$ are subject to structural breaks (Favero 1993, p. 458). The coefficients $q_{1t}$, $\alpha_{1t}$, $\alpha_{2t}$, and $\alpha_{3t}$ may then be expected to vary, reflecting change in the underlying parameters. Such a possibility casts some doubt about the strong acceptance of the costly adjustment model found in the previous chapter because acceptance may have had more to do with altering coefficients than with the validity of the model. Analysis of the model in a time-varying framework is also considered.

By permitting the coefficients in Equation (5.12) to vary, the model is seen as going some distance in satisfying the Lucas critique. As mentioned, once the parameters of this equation are permitted to vary, the effects of change among the underlying parameters and the auxiliary system can efficiently be incorporated into the costly adjustment model. These rule parameters are of interest because they provide valuable information about how the complete system varies over the sample.
Equation (5.12) necessitates the determination of \( \{E_{t-1} \Delta y, E_{t-1} \Delta y_{kt}, E_{t-1} \Delta g_{t}\} \). For the work in this chapter, a time-varying auxiliary system is used which appears appropriate given the rejection of time-invariant coefficients found in Chapter 4 and following an explicit test of the stability of the auxiliary system noted below. These results suggest that the expectation-generating mechanism should be allowed to vary over the sample. The auxiliary system to be used may be expressed as:

\[
\Delta y_{lt} = b_{20, l} + b_{22, l} \Delta y_{lt-1} + b_{23, l} \Delta y_{lt-2} + b_{24, l} \Delta y_{lt-3} + b_{25, l} \Delta y_{kt-2} + b_{26, l} \Delta g_{t-1} + b_{27, l} \Delta g_{t-2} + b_{28, l} \Delta g_{t-3},
\]

\[
\Delta y_{kt} = b_{30, l} + b_{32, l} \Delta y_{lt-1} + b_{33, l} \Delta y_{lt-2} + b_{34, l} \Delta y_{lt-3} + b_{35, l} \Delta y_{kt-2} + b_{36, l} \Delta g_{t-1} + b_{37, l} \Delta g_{t-2} + b_{38, l} \Delta g_{t-3}, \quad \text{and}
\]

\[
\Delta g_{t} = b_{40, l} + b_{42, l} \Delta y_{lt-1} + b_{43, l} \Delta y_{lt-2} + b_{44, l} \Delta y_{lt-3} + b_{45, l} \Delta y_{kt-2} + b_{46, l} \Delta g_{t-1} + b_{47, l} \Delta g_{t-2} + b_{48, l} \Delta g_{t-3}.
\]

This method of specifying such a time-varying system in this manner are found in papers by Struth (1984) and Hall (1993). Note that \( \Delta y_{kt-1} \) and \( \Delta y_{kt-3} \) have been dropped from the system to reduce the computational burden. These two variables were found to be jointly insignificant in the time-invariant version of Model II.

Two comments need to be made about this system. First, a time-varying constant is specified to pick up any change in the mean over the sample. Second, unlike Honkapohja (1993), we assume that the agent uses the parameter values which embody information up till period \( t \). [Honkapohja (1993) uses parameters specified at time \( (t-1) \).] If we assume coefficients follow a random walk, however, both specifications will be the same.

With system (5.14) the expectations for Equation (5.12) can be determined. In particular;
\[ E_{t-1} \Delta y_{t} = b_{20,t-1} + b_{22,t-1} \Delta y_{t-1} + b_{23,t-1} \Delta y_{t-2} + b_{24,t-1} \Delta y_{t-3} + b_{25,t-1} \Delta y_{t-2} + b_{26,t-1} \Delta y_{t-1} + b_{27,t-1} \Delta y_{t-2} + b_{28,t-1} \Delta y_{t-3} + b_{29,t-1} \Delta y_{t-2} + b_{30,t-1} \Delta y_{t-1} + b_{31,t-1} \Delta y_{t-2} + b_{32,t-1} \Delta y_{t-3} + b_{33,t-1} \Delta y_{t-2} + b_{34,t-1} \Delta y_{t-3} + b_{35,t-1} \Delta y_{t-2} + b_{36,t-1} \Delta y_{t-1} + b_{37,t-1} \Delta y_{t-2} + b_{38,t-1} \Delta y_{t-3} + b_{39,t-1} \Delta y_{t-2} + b_{40,t-1} \Delta y_{t-1} + b_{41,t-1} \Delta y_{t-2} + b_{42,t-1} \Delta y_{t-3} + b_{43,t-1} \Delta y_{t-2} + b_{44,t-1} \Delta y_{t-3} + b_{45,t-1} \Delta y_{t-2} + b_{46,t-1} \Delta y_{t-1} + b_{47,t-1} \Delta y_{t-2} + b_{48,t-1} \Delta y_{t-3} + b_{49,t-1} \Delta y_{t-2} + b_{50,t-1} \Delta y_{t-1} + b_{51,t-1} \Delta y_{t-2} + b_{52,t-1} \Delta y_{t-3} + b_{53,t-1} \Delta y_{t-2} + b_{54,t-1} \Delta y_{t-3} + b_{55,t-1} \Delta y_{t-2} + b_{56,t-1} \Delta y_{t-1} + b_{57,t-1} \Delta y_{t-2} + b_{58,t-1} \Delta y_{t-3} + b_{59,t-1} \Delta y_{t-2} + b_{60,t-1} \Delta y_{t-1} + b_{61,t-1} \Delta y_{t-2} + b_{62,t-1} \Delta y_{t-3} + b_{63,t-1} \Delta y_{t-2} + b_{64,t-1} \Delta y_{t-3} + b_{65,t-1} \Delta y_{t-2} + b_{66,t-1} \Delta y_{t-1} + b_{67,t-1} \Delta y_{t-2} + b_{68,t-1} \Delta y_{t-3} + b_{69,t-1} \Delta y_{t-2} + b_{70,t-1} \Delta y_{t-3} + b_{71,t-1} \Delta y_{t-2} + b_{72,t-1} \Delta y_{t-3} + b_{73,t-1} \Delta y_{t-2} + b_{74,t-1} \Delta y_{t-3} + b_{75,t-1} \Delta y_{t-2} + b_{76,t-1} \Delta y_{t-3} + b_{77,t-1} \Delta y_{t-2} + b_{78,t-1} \Delta y_{t-3} + b_{79,t-1} \Delta y_{t-2} + b_{80,t-1} \Delta y_{t-3} + b_{81,t-1} \Delta y_{t-2} + b_{82,t-1} \Delta y_{t-3} + b_{83,t-1} \Delta y_{t-2} + b_{84,t-1} \Delta y_{t-3} + b_{85,t-1} \Delta y_{t-2} + b_{86,t-1} \Delta y_{t-3} + b_{87,t-1} \Delta y_{t-2} + b_{88,t-1} \Delta y_{t-3} \]

(5.15)

where, for example, \( b_{22,t-1} = E[b_{22,t-1} | \mathcal{I}_{t-1}] \); that is, the agent's expected value of \( b_{22,t-1} \) given information at time \( (t-1) \). Assuming that coefficients follow random walk processes, these expected values may be rewritten as, with justification to follow:

\[ R_{1}^{t-1} = E_{t-1} \Delta y_{t} = b_{20,t-1} + b_{22,t-1} \Delta y_{t-1} + b_{23,t-1} \Delta y_{t-2} + b_{24,t-1} \Delta y_{t-3} + b_{25,t-1} \Delta y_{t-2} + b_{26,t-1} \Delta y_{t-1} + b_{27,t-1} \Delta y_{t-2} + b_{28,t-1} \Delta y_{t-3} + b_{29,t-1} \Delta y_{t-2} + b_{30,t-1} \Delta y_{t-1} + b_{31,t-1} \Delta y_{t-2} + b_{32,t-1} \Delta y_{t-3} + b_{33,t-1} \Delta y_{t-2} + b_{34,t-1} \Delta y_{t-3} + b_{35,t-1} \Delta y_{t-2} + b_{36,t-1} \Delta y_{t-1} + b_{37,t-1} \Delta y_{t-2} + b_{38,t-1} \Delta y_{t-3} + b_{39,t-1} \Delta y_{t-2} + b_{40,t-1} \Delta y_{t-1} + b_{41,t-1} \Delta y_{t-2} + b_{42,t-1} \Delta y_{t-3} + b_{43,t-1} \Delta y_{t-2} + b_{44,t-1} \Delta y_{t-3} + b_{45,t-1} \Delta y_{t-2} + b_{46,t-1} \Delta y_{t-1} + b_{47,t-1} \Delta y_{t-2} + b_{48,t-1} \Delta y_{t-3} + b_{49,t-1} \Delta y_{t-2} + b_{50,t-1} \Delta y_{t-1} + b_{51,t-1} \Delta y_{t-2} + b_{52,t-1} \Delta y_{t-3} + b_{53,t-1} \Delta y_{t-2} + b_{54,t-1} \Delta y_{t-3} + b_{55,t-1} \Delta y_{t-2} + b_{56,t-1} \Delta y_{t-1} + b_{57,t-1} \Delta y_{t-2} + b_{58,t-1} \Delta y_{t-3} + b_{59,t-1} \Delta y_{t-2} + b_{60,t-1} \Delta y_{t-1} + b_{61,t-1} \Delta y_{t-2} + b_{62,t-1} \Delta y_{t-3} + b_{63,t-1} \Delta y_{t-2} + b_{64,t-1} \Delta y_{t-3} + b_{65,t-1} \Delta y_{t-2} + b_{66,t-1} \Delta y_{t-1} + b_{67,t-1} \Delta y_{t-2} + b_{68,t-1} \Delta y_{t-3} + b_{69,t-1} \Delta y_{t-2} + b_{70,t-1} \Delta y_{t-1} + b_{71,t-1} \Delta y_{t-2} + b_{72,t-1} \Delta y_{t-3} + b_{73,t-1} \Delta y_{t-2} + b_{74,t-1} \Delta y_{t-3} + b_{75,t-1} \Delta y_{t-2} + b_{76,t-1} \Delta y_{t-1} + b_{77,t-1} \Delta y_{t-2} + b_{78,t-1} \Delta y_{t-3} + b_{79,t-1} \Delta y_{t-2} + b_{80,t-1} \Delta y_{t-1} + b_{81,t-1} \Delta y_{t-2} + b_{82,t-1} \Delta y_{t-3} + b_{83,t-1} \Delta y_{t-2} + b_{84,t-1} \Delta y_{t-3} + b_{85,t-1} \Delta y_{t-2} + b_{86,t-1} \Delta y_{t-1} + b_{87,t-1} \Delta y_{t-2} + b_{88,t-1} \Delta y_{t-3}. \]

(5.16)

The \( R_{i} \)'s have a superscript to stress the fact they are composed of coefficient estimates and variables available at time \( t-1 \) or before. Using these expectations, Equation (5.12) may be rewritten as

\[(5.17) \Delta c_{t} = q_{0,t} + q_{1,t} \Delta c_{t-1} + \alpha_{1,t} R_{1}^{t-1} + \alpha_{2,t} R_{2}^{t-1} + \alpha_{3,t} R_{3}^{t-1} + \epsilon_{t}^{c}. \]

By acknowledging the Lucas critique, Equation (5.17) embodies a learning rule which permits the agent to incorporate changes in the economic system in a flexible manner. Implementation of learning in this particular framework is considered in the following section.

**Estimation**

Estimation proceeds by combining Equation (5.12) and system (5.14) to form a joint system, permitting correlations to exist among the variables of the auxiliary system and the change in consumption. This specification differs from Hall (1993), who assumes that
expectations are formed as a separate step, an approach which in essence ignores cross-correlations between $\Delta c_t$ and $\{\Delta y_{t1}, \Delta y_{t2}, \Delta g_t\}$. By explicitly modeling these correlations, a more general view on determining consumption is obtained. A multivariate Kalman filter, to be discussed, will be used in providing optimal estimates for all unknown coefficients. Before presenting this particular estimation technique, however, alternative estimation methods and parameter specifications are discussed.

We first consider the work of Friedman (1979) because important insights can be realized by examining this simplest case. Friedman allows the agent to learn about some unknown parameter, $\beta$, via recursive least squares. With such an estimation technique, parameter estimates are updated each period as new information is obtained. Forming expectations with such a system allows the expectation-generating mechanism to alter as the agent obtains more information about the system. As Friedman states:

As time passes, economic agents who exploit optimally whatever information is available will in general revise their expectations generation process, and will in general therefore revise any associated predictions which are still relevant.... This expectation formation procedure is fully optimal in the sense of meeting the informational exploitation assumption of Muth's rational expectation hypothesis. Economic agents forming expectations in this way at time period $t$ take full advantage of all available information, in the most efficient way to yield the minimum variance predictor within the class of all linear unbiased predictors. (p. 30)

Insight into Friedman's procedure can be gained by considering the simple example where an agent attempts to formulate $E_{t1}y_{t1}$ given that $y_t = \beta y_{t-1} + \epsilon_t$. Rational expectations implies that $E_{t1}y_t = b y_{t-1}$, where $b$ is a parameter estimate. Friedman, however, lets the agent learn about $\beta$ over time by using recursive least squares to obtain estimates of $\beta$ at time $t$. In particular, the estimator of $\beta$, $b_t$, is assumed to evolve as
where $b_i$ is updated each period as new information becomes available. [Derivation of this is found in (Harvey 1992, pp. 98-99).] This estimation technique can also be shown to be analogous to the approach used by Marcet and Sargent (1989) [system (5.5)] when $\alpha_i = 1 \forall i$. Using this estimation procedure, it is seen that $E_{t+1}Y_t = b_{t+1}Y_{t+1}$, which provides an explanation as to why this learning procedure is a member of the bounded rational class. At time $t$, $b_i$ will generally not be equal to $\beta$; thus, the resultant expectation may not be equal to that obtained by assuming rational expectations. As Friedman (1979) shows, the optimality properties of the REH only arise when the sequence $\{b_t\}$ approaches $\beta$ as $t$ increases. Comparing the recursion system just presented to that used by Honkapohja (1993) shows that we are not guaranteed convergence because the term $\frac{\alpha_i}{t}$ (i.e., the gain term) need not approach zero as the sample progresses, because the term $\left(\frac{\alpha_i}{t}\right)$ is excluded in this particular model.

Convergence as discussed by Marcet and Sargent (1989) and Ljung and Soderstrom (1983) requires that a term that “looks like” $1/t$ (where $t$ denotes time) is included in the gain term [proposition (e) (Sargent 1993, p. 131)]. As shown in Appendix D, a requirement before applying the convergence theorem is for $b_t = b_{t+1}$; currently, this only occurs if the gain term approaches 0 as $t \to \infty$. We will show, however, that the sequence $\{b_t\}$ may still approach $\beta$.

The work in this chapter differs from the approaches used by Friedman (1979), Marcet and Sargent (1989), and Honkapohja (1993) by assuming that the economic system is continuously changing over time. Under such a regime, least squares learning, whether following the approach of Marcet and Sargent (1989) or Friedman (1979), is no longer optimal. In this circumstance an application of the Kalman filter is the appropriate way of proxying the rational agent’s learning process (Kim and Nelson 1989, p. 434). To model
continual change, we allow the coefficients of the system to follow a random walk process; that is,
\[ \beta_t = \beta_{t-1} + \eta_t, \]
where \( \beta_t \) denotes some vector of true coefficients and \( \eta_t \) is an appropriately dimensioned vector of disturbances. Engle and Granger (1987b) argue that random walk parameters are also plausible in the instance of structural change. For example, if a policy regime is changed agents will adjust their behavior. Coefficients that held in the previous regime will change to reflect the altered environment. Plausibly, one would change the estimate of the parameter vector only when new information becomes available, thus suggesting a unit root (Engle and Granger 1987b, page 249).

The assumption regarding the possibility of change in the system seems reasonable from the vantage point of the representative agent participating in an evolving economy. Recent economic history shows that this assumption of change is quite plausible. As will be seen, the method used in this chapter is sufficiently flexible so as to encompass the least squares learning approach found in Friedman (1979).

Following Kim and Nelson (1989), the assumption of continual change can be tested. First, the system must be shown to be time-varying. In Chapter 4, parameters of both Models I and II were shown to be time-varying. Currently, we focus on the VAR(3) auxiliary system comprised of \( \{ \Delta y_{it}, \Delta y_{kt}, \Delta g_{it} \} \), because it is this system which generates expectations. This approach will offer explicit conclusions about whether the expectation-generating mechanism used in Equation (5.12) does indeed change over the sample. As discussed in Chapter 4, the sample is split into two portions, before and after 73Q2, and the generalized variances of each segment are compared to that obtained when the VAR(3) system is estimated over the whole sample. A likelihood ratio test, described previously, is then performed to test whether the coefficients are equal over the two periods. Denoting \( v_1 \), \( v_2 \), and \( v_0 \) as the determinants of the sample variance/covariance matrices for the first segment, second segment, and the whole segment, respectively, the following results are obtained: \( v_1 = 5.02E-07 \), \( v_2 = 7.84E-07 \), and \( v_0 = 8.89E-07 \). Performing a likelihood ratio test produces a test statistic of 45.12 with 24
degrees of freedom and an associated p-value of 0.005. Thus, the null hypothesis of invariant coefficients over the two segments of the sample can be rejected at the 1 percent level.

Given that parameters are not constant, we can test whether parameters follow a random walk following Kim and Nelson (1989) and Bohara and Sauer (1992). Under the alternative hypothesis of random walk coefficients (the null hypothesis is simply that coefficients are time-varying), residuals from an OLS regression have a particular heteroscedastic form which depends on $t^*X^2$, where $t$ denotes time and $X$ is a vector of regressors including a constant. From Breusch and Pagan (1979), one-half times the explained sum of squares from a regression of $\frac{\hat{u}_t^2}{\sigma_u^2}$ on $t^* X^2$, where $\hat{u}_t^2$ denotes the residuals from the OLS regression and $\sigma_u^2$ is the associated variance, is distributed as $\chi^2(k)$, where $k$ is the number of regressors (Kim and Nelson 1989, p. 434). Tests may be performed on each variable individually (with $k = 1$) to determine whether the coefficient associated with that particular variable evolves as a random walk, or with all variables to determine whether all coefficients jointly follow a random walk. A significant test statistic would imply that the coefficients, either individually or jointly, follow a random walk. Performing these tests on the univariate series that comprise system (5.14) produces the results presented in Table 5.1. These results show that for many coefficients the alternative hypotheses of random walk parameters are accepted. This result seems to occur most strongly in the $\Delta y_t$ equation.

By assuming that the parameters evolve as a random walk, we implicitly allow the agent to be cognizant of the time-varying nature of the system. Parameters are permitted to vary over time with no particular direction or trend. Such a parameter specification allows considerable scope for systematic variation in the parameters (Cuthbertson, Hall, and Taylor 1992). Bullard (1992, p. 163) shows how the recursive estimation technique of Marcet and Sargent (1989) is modified once the system parameters are allowed to follow a random walk. Building on the work of Ljung and Soderstrom (1983), the following system can be specified:
\[(5.18) \, y_i' = \varphi_{t-1} y_i' + \eta_i'. \]
\[(5.19) \, \beta_i' = \beta_{t-1}' + \omega_i \quad \text{where:} \]
\[E(\eta_i, \eta_i') = H, \, E(\eta_i) = 0, \]
\[E(\omega_i, \omega_i') = \Omega, \, E(\omega_i) = 0, \]
\[E(\eta_i, \omega_i') = 0. \]

Table 5.1. Coefficient tests

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Calculated $\chi(k)$</th>
<th>Parameter</th>
<th>Calculated $\chi(k)$</th>
<th>Parameter</th>
<th>Calculated $\chi(k)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b_{02}$</td>
<td>13.77**</td>
<td>$b_{03}$</td>
<td>17.95**</td>
<td>$b_{04}$</td>
<td>15.9**</td>
</tr>
<tr>
<td>$b_{22}$</td>
<td>10.25**</td>
<td>$b_{32}$</td>
<td>11.50**</td>
<td>$b_{42}$</td>
<td>4.20*</td>
</tr>
<tr>
<td>$b_{23}$</td>
<td>0.95</td>
<td>$b_{33}$</td>
<td>1.78</td>
<td>$b_{43}$</td>
<td>0.01</td>
</tr>
<tr>
<td>$b_{24}$</td>
<td>5.45*</td>
<td>$b_{34}$</td>
<td>3.30</td>
<td>$b_{44}$</td>
<td>5.85*</td>
</tr>
<tr>
<td>$b_{25}$</td>
<td>3.79</td>
<td>$b_{35}$</td>
<td>4.10*</td>
<td>$b_{45}$</td>
<td>2.70</td>
</tr>
<tr>
<td>$b_{26}$</td>
<td>6.33*</td>
<td>$b_{36}$</td>
<td>5.00*</td>
<td>$b_{46}$</td>
<td>10.10**</td>
</tr>
<tr>
<td>$b_{27}$</td>
<td>1.94</td>
<td>$b_{37}$</td>
<td>4.16*</td>
<td>$b_{47}$</td>
<td>4.70*</td>
</tr>
<tr>
<td>$b_{28}$</td>
<td>3.17</td>
<td>$b_{38}$</td>
<td>2.95</td>
<td>$b_{48}$</td>
<td>4.80*</td>
</tr>
</tbody>
</table>

all 8 parameters 17.67* all 8 parameters 21.11** all 8 parameters 17.97*

Note: * denotes significance at the 5% level and ** denotes significance at the 1% level.

Equation (5.19), the state equation, shows that $\beta$ evolves as a random walk. Bullard (1992) shows that this system can be estimated recursively using the system

\[(5.20) \, \beta_i' = \beta_{t-1}' + P_{t-1} \varphi_{t-1} ' E^{-1}_{t-1} \{ y_i - \varphi_{t-1} \beta_{t-1}' \}, \]
\[P_t = P_{t-1} + \Omega - P_{t-1} \varphi_{t-1} ' P_{t-1} E^{-1}_{t-1}, \]

where $\varphi_i = H + \varphi_{t-1} ' P_{t-1} \varphi_{t-1}$.

When $\Omega = 0$ and $H = 1/\alpha$, this system corresponds to that analyzed by Marcet and Sargent (1989) and Honkapohja (1993). By estimating a more general model where these conditions...
are not imposed, convergence to the REE is not guaranteed (Bullard 1992, p. 164). Ljung and Soderstrom (1983) provide an explanation

This term \([\Omega]\) prevents \(P_t\) from tending to zero consequently it keeps up the gain vector \([P_t|\Gamma_t|G_t|^{-1}]\). This is very natural from an intuitive point of view. When the system is time-varying the algorithm must be more 'alert'. The price for being persistently alert \([P_t|\Gamma_t|G_t|^{-1}\) not tending to zero] is, of course, that the estimates are always sensitive to the random disturbances in the measurements. The estimates will not converge to their true values, the covariance matrix does not tend to zero. There is, as always, a compromise between alertness and noise sensitivity. (p. 56)

Thus, a cost is incurred by permitting the system parameters to evolve in the manner just discussed, namely, convergence cannot be guaranteed. Although such a situation would be unfortunate if the true parameters were fixed, this inability to converge is moot if the true parameters are in fact time-varying. The notion of convergence to some fixed value would be chimerical because even if this learning process succeeded in converging to the true model, it may in effect be chasing a moving target and not converge to a stable set of parameters (Hall 1993, p. 273).

Ljung and Soderstrom (1983) also suggest that even though convergence is not guaranteed, as described by Marcet and Sargent (1989), it may in practice occur if the gain vector gets small (as it frequently does in estimation) as the sample increases. This can easily be checked once the system is estimated. A decreasing gain vector is largely dependent upon the data and thus cannot be presupposed a priori. One would expect that the sequence of elements of the gain vector corresponding to a parameter which shows little variability to get small while that associated with a volatile parameter to remain large as the sample progresses. Estimation which allows parameters to follow a random walk may thus represent a more general approach compared to that used by Marcet and Sargent (1989), because the results of the latter may arise "automatically," if the data warrants it. Stability of coefficients could be ascertained by studying time-plots of coefficient estimates and by examining estimated variances which comprise the matrix \(Q\) \([E(\eta,\eta)]\).
We now explicitly consider the manner in which learning is incorporated in modeling consumption. Central to this work is the Kalman filter, which provides an efficient method by which time-varying parameters may be modeled. Estimation with this method, however, provides much more. As a byproduct, Kalman filter estimation provides optimal one-step-ahead predictions conditional on all information up till the preceding period. These predictions can be used as proxies for the unknown values \( \{E_{t-1}, \Delta y_h, E_{t-1}, \Delta y_{kt}, E_{t-1}, \Delta g_t\} \) in Equation (5.12). In this framework, agents are assumed to assess their model continuously and revise it as soon as persistent, or systematic, forecast errors become evident. It is thus possible to model how the agents attain rationality and how they manage to identify and adjust to breaks in the economic structure (Struth 1984, p. 211). Expectations generated with the Kalman filter also possess two nice qualities which are important if the agent is rational, namely, that forecasts are unbiased and that prediction errors are serially uncorrelated (Harvey 1989, pp. 110-112). Recall that these are requirements for the REH. Following Hall (1993) tests can be performed to see if the learning model fulfills these properties.

We now offer a brief review of the Kalman filter. More detailed expositions may be found in Harvey (1989, 1992) and Cuthbertson, Hall, and Taylor (1992).

The Kalman filter

The material in this section draws heavily from Harvey (1989, 1992). To maintain consistency, Harvey's notation will be followed.

As witnessed by a growing number of applications, the Kalman filter has been recognized as an efficient solution technique in estimating state-space models. The general state space form relates an observable variable \( y_t \) with the state vector \( \alpha_t \) (Harvey 1989, pp. 100-101). This equation, called the measurement equation, can be expressed as

\[
y_t = Z_t \alpha_t + \varepsilon_t, \quad t = 1, \ldots, T,
\]
where $e_i$ reflects the fact that $Z_t \alpha_t$ measures $y_t$ only with error. The vectors $y_t$ and $e_i$ are $(n \times 1)$, $\alpha_t$ is $(m \times 1)$, and $Z_t$ is $(n \times m)$. In all the work to follow, we assume that $E(e_i) = 0$ and $\text{Var}(e_i) = H$.

The elements of $\alpha_t$, however, are not observable and are assumed to be generated by the process

$$\alpha_t = T \alpha_{t-1} + \eta_t \quad t = 1, \ldots, T,$$

which is the transition equation where it is assumed that $E(\eta_t) = 0$, $\text{Var}(\eta_t) = Q$, and $\text{Cov}(\eta_t, e_i') = 0$. Taken together $Z_t$, $T$, $H$, and $Q$ are referred to as the system matrices.

Once the model has been placed in a state-space form, the Kalman filter can be applied. Assuming the $e_i$ and $\eta_t$ are normally distributed, the filter gives the mean and covariance matrix of the state vector conditional on the information available at that particular point in time (Harvey 1989, p. 101). That is, given the assumptions above, the state-space system can be estimated to provide optimal estimators for the state vector $\alpha_t$ based on all the observations up till $y_t$. The covariance matrix of the associated error [i.e., $E[(\alpha_t - a_t)(\alpha_t - a_t)']$ where $a_t$ is an estimate of $\alpha_t$] is denoted as $P_t$. Conditional on information at time $t-1$, estimates of $a_t$ and $P_t$ can be produced by the prediction equations

\begin{align*}
(5.21) \quad a_{t|t-1} &= T a_{t-1} \quad \text{and} \\
(5.22) \quad P_{t|t-1} &= T P_{t-1} T' + Q,
\end{align*}

where $a_{t|t-1}$ denotes the estimate of $a_t$ conditional on information up till period $t-1$. An analogous interpretation holds for $P_{t|t-1}$. Using this notion of $a_{t|t-1}$, an estimate for $y_t$ conditional on information in time $t-1$ ($L_{t-1}$) can be expressed as

$$E(y_t|L_{t-1}) = y_{t|t-1} = Z_t a_{t|t-1}.$$

The mean square error of the prediction error, $v_t$, where $v_t = y_t - y_{t|t-1}$ can be expressed as

$$v_t = Z_t (\alpha_t - a_t) + e_t.$$

The variance/covariance associated with this error is

$$F_t = Z_t P_{t|t-1} Z_t' + H.$$
Assuming that \( \epsilon_t \) and \( \eta_t \) are normally distributed, \( \nu_t \) will be distributed as \( \text{MVN}(0, F_t) \), with \( F_t \) being a \((n \times n)\) positive definite variance/covariance matrix. Using the previously defined expressions, the estimator of the state vector, \( \alpha_t \), can be updated as each new observation becomes available. These updating equations are:

\[
\begin{align*}
\alpha_t &= \alpha_{t-1} + P_{t-1}Z_tF_t^{-1}(y_t - Z_t\alpha_{t-1}) \quad \text{and} \\
P_t &= P_{t-1} - P_{t-1}Z_tF_t^{-1}Z_t^tP_{t-1}.
\end{align*}
\]

Together, Equations (5.21) through (5.24) comprise the Kalman filter. Given initial conditions \( \alpha_0 \) and \( P_0 \) and the elements of \( Q \) and \( H \), the Kalman filter provides the optimal estimator of the state as each new observation becomes available. When all \( \tau \) observations have been processed, the estimator, \( \hat{\alpha}_\tau \), contains all the information necessary to make predictions of future observations (Harvey 1992, p. 86).

Equation (5.23) shows precisely how information is incorporated into the estimator of \( \alpha_t \). Struth (1984) terms this equation the learning rule, where forecast errors (\( \nu_t \)) provide revisions in the state variable estimates through two avenues. First is through the presence of the forecast error itself. Second is through the adjustment coefficient \( P_{t-1}Z_tF_t^{-1} \) (that is, the gain), which is a function of the inverse covariance matrix of the prediction errors. Taken together, prediction errors provide information with which agents may revise estimates of \( \alpha_t \) from the previous period.

As mentioned, utilizing the recursions implied by Equations (5.21) through (5.24) to achieve optimal estimates of \( \alpha_t \) requires initial estimates of \( \alpha_0 \) and \( P_0 \) along with the variances and covariances that comprise \( Q \) and \( H \). In this chapter the assumption of a diffuse prior will be used in providing initial conditions for \( \alpha_0 \) and \( P_0 \). Definition and rationale for this choice are discussed later. As for the necessary elements of \( Q \) and \( H \), these quantities must be estimated, conditional on the initial conditions. For tractability, it is common to assume that \( Q \) is a diagonal matrix, which precludes correlation among elements of the state vector, implying that only \( m \) variances need to be determined. The matrix \( H \) is left unrestricted, meaning that \( n(n-1)/2 \) variances and covariances need to be determined. Even with the assumptions placed
on \( Q \), estimation of these unknown parameters (called hyperparameters) is not an easy task. Estimation of these quantities uses the maximum likelihood procedure laid forth in Harvey (1989).

Following Harvey (1989), the likelihood function may be obtained from the state space model via the prediction error decomposition. This likelihood function can be expressed as:

\[
\lambda = \ln L = -\frac{n\pi}{2} + \ln 2\pi - \frac{1}{2} \sum_{i=1}^{x} \ln|F_i| - \frac{1}{2} \sum_{i=1}^{x} v'_i F^{-1}_i v_i,
\]

where \( \tau \) is the number of periods in the sample and \( n \) denotes the number of series. This likelihood function can be evaluated, given appropriate initial conditions and estimates of \( Q \) and \( H \), by running through Equations (5.21) through (5.24) from time 1 to time \( \tau \). Estimation of the components of the matrices \( Q \) and \( H \) proceeds by attempting to minimize \(-\ln L\) using a variable metric minimization routine with numerical derivatives.

**Application of the Learning Model**

Now that we have provided a brief introduction to the Kalman filter, an application to the learning model is considered. The previously specified learning model necessitates the joint estimation of Equations (5.14) and (5.17). Of primary importance in this system is the determination of the expectations \( \{E_{t-1}\Delta y_{it}, E_{t-1}\Delta y_{ut}, E_{t-1}\Delta g_t\} \), which are generated by a time-varying auxiliary system. As discussed, we assume that all parameters evolve as random walks; that is, \( \beta_t = T\beta_{t-1} + \eta_t \), where \( T \) is an identity matrix and \( \text{Var}(\eta_t\eta_t') = Q \), implying that \( b_{it-1} = b_{t-1} \). With this assumption the necessary expectations can be generated via system (5.15). Substituting these expectations into Equation (5.12) allows the change in consumption to evolve as

\[
\Delta c_t = q_{0t} + q_{1t}\Delta c_{t-1} + \alpha_{1t} R^{-1}_1 + \alpha_{2t} R^{-1}_2 + \alpha_{3t} R^{-1}_3 + \varepsilon_t,
\]

with the \( R^{-1}_1 \) defined as in system (5.16). Equations (5.26) and system (5.14) may thus be combined into a single system by appropriately stacking the model parameters in the form
\(G_t = (q_{it}, q_{it}, \alpha_{it}, \alpha_{2it}, \beta_{it}, \beta_{it}, \beta_{23it}, \beta_{24it}, ..., \beta_{28it})^\prime\), where \(q_{it}, \alpha_{it}, \) and \(\beta_{it}\) denote the true unknown parameters \((q_{it}, \alpha_{it}, \) and \(b_{it}\) are estimates of these quantities). Using this parameter vector, the measurement equation can be expressed as:

\[
\begin{bmatrix}
\Delta c_t \\
\Delta y_{lt} \\
\Delta y_{kt} \\
\Delta g_t
\end{bmatrix} = \begin{bmatrix}
1 & \Delta c_{t-1} & R_{1t}^{-1} & R_{2t}^{-1} & R_{3t}^{-1} & 0_{1x8} & 0_{1x8} & 0_{1x8}
0 & 0 & 0 & 0 & 0 & \gamma_t & 0_{1x8} & 0_{1x8}
0 & 0 & 0 & 0 & 0 & 0_{1x8} & \gamma_t & 0_{1x8}
0 & 0 & 0 & 0 & 0 & 0_{1x8} & 0_{1x8} & \gamma_t
\end{bmatrix}
\begin{bmatrix}
\varepsilon_t^c \\
\varepsilon_t^{y_l} \\
\varepsilon_t^{y_k} \\
\varepsilon_t^g
\end{bmatrix},
\]

where \(\gamma_t = [1 \Delta y_{h-1} \Delta y_{h-2} \Delta y_{h-3} \Delta y_{kt-2} \Delta g_{t-1} \Delta g_{t-2} \Delta g_{t-3}]\), or as

(5.27) \(Y_t = Z_t G_t + \varepsilon_t\).

The associated state equation is:

(5.28) \(G_t = G_{t-1} + \eta_t\),

where \(G_t\) and \(\eta_t\) are \((29 \times 1)\) vectors. Equations (5.27) and (5.28) together form a state-space model which is liable for estimation with the Kalman filter. Note that if \(\text{Var}(\eta_t\eta_t')\) is equal to a zero matrix, estimation of this system is analogous to that of recursive least squares.

Having the measurement matrix \(Z_t\) dependent on information available at time \(t-1\), however, requires explanation. Harvey (1989) terms this formulation a “conditionally Gaussian model”, which means that

even though the system matrices may depend on observations up to and including [time \(t-1\)], they may be regarded as being fixed once we are at time \(t-1\)... derivation of the Kalman filter goes through exactly as [above] with \(a_{l,t-1}\) and \(P_{l,t-1}\) now interpreted as the mean and covariance matrix of the distribution of \(\alpha_t\) conditional on the information at time \(t-1\). (p. 156)

Because \(R_1, R_2,\) and \(R_3\) are known in period \(t\), the measurement matrix can be taken as being fixed. However, a measurement matrix that varies with time may have an affect on estimation. Unlike models with a fixed measurement matrix, there is no guarantee that the Kalman gain will converge to some level as the sample progresses. As discussed, this condition is important in determining whether the learning model converges to the REE
because even though convergence of the sequence of gain terms is not guaranteed a priori, convergence may be obtained in practice. Ljung and Soderstrom (1983) suggest that as the sample progresses the gain terms may in fact approach some limit, even though the measurement matrix is time-varying. Thus, it is possible for the learning model used in this chapter to converge to the REE if warranted by the data.

By allowing parameters to evolve as a random walk while using the Kalman filter for estimation, estimated coefficients are derived differently compared to Friedman (1979) and Marcet and Sargent (1989). The difference in estimation can be seen by studying Equation (5.23) and by letting $b_t = a_t$. With this change of notation, Equation (5.23) can be expressed as:

$$a_t = a_{t-1} + (P_{t-1} + Q)Z_t'F_t^{-1}(y_t - Z_t a_{t-1})$$

because $P_{t-1} = P_{t-1} + Q$.

Thus, the vector $a_t$ will only behave as in the recursive least square approach if $Q = 0$. In the absence of this condition, such behavior will not be observed. An appealing aspect of estimation as suggested in this chapter is that $Q$ will be determined by the data. If some coefficients do in fact evolve in a manner consistent with the approach of Marcet and Sargent (1989), the associated element of $Q$ will be zero, or very close to zero. By assuming that coefficients evolve as a random walk and with maximum likelihood estimation using the Kalman filter, a more general approach is achieved because behavior related to estimation following the methods of Friedman (1979) and Marcet and Sargent (1989) may arise as special cases.

To implement this learning model, we must determine the elements of $Q$ and $H$. The variance/covariance matrix $Q$ is assumed to be a diagonal matrix of dimension 29, implying that covariances among the coefficients are ignored. Each diagonal element of $Q$ is the variance for the associated coefficient. The variance/covariance matrix $H$ is left unrestricted, thus necessitating the estimation of four variances and six covariances. In total, 39 hyperparameters need to be determined.

Estimation proceeds by specifying $a_0$ and $P_0$ and some initial values for $Q$ and $H$ and utilizing the Kalman recursions in an attempt to minimize Equation (5.25), which is the likelihood function. The terms $a_0$ and $P_0$ are provided by assuming a diffuse prior which
acknowledges that the agent has a lack of prior information as the system starts. Thus, the agent approaches the system in complete ignorance. Statistically, this translates into assuming that $a_0$ is a $(29 \times 1)$ zero vector and $P_0$ a $(29 \times 29)$ identity matrix which is multiplied by some large constant (Harvey (1989, p. 212); Ljung and Soderstrom 1983, p. 21). We assume that the agent uses the first 15 quarters to obtain information about the system and thus exclude these data in evaluating the likelihood function.

With initial estimates of $Q$ and $H$ and the diffuse prior specifications for $a_0$ and $P_0$, determining the 39 unknown variances and covariances that minimize Equation (5.25) can proceed. The only aspect of this estimation that will be commented on here is the need for care when the minimization routine iterates to some minimum. It was found in searching the parameter space that the $F_1$ matrix had a tendency to violate its assumed positive definitiveness. In such an event, the recursions break down. To alleviate this possibility, the eigenvalues of $F_1$ are computed and compared to 0 at each level of recursion (that is, when Equations (5.21) through (5.24) are evaluated at time $t$). Negative eigenvalues violate positive definiteness [Rule 4, Lutkepohl (1991, p. 457)] and a large penalty is affixed to that particular combination of hyperparameters that generated such a result. Iterations are continued until no further decrease in $-\ln \lambda$, from Equation (5.25), can be achieved. Though convergence is slow, different initial values tend to produce similar values of the likelihood function, suggesting robustness of the results reported in the following subsection.

**Estimation Results**

Using the method just described, estimation produces the maximum likelihood estimates reported in Table 5.2. Again, elements of $Q$ denote variances associated with the state equation. Zeros denote estimated variances that are very small.
Table 5.2. Estimated Elements of $Q$

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimated Variance</th>
<th>Parameter</th>
<th>Estimated Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>$q_0$</td>
<td>0.5E-6</td>
<td>$b_{33}$</td>
<td>0.0</td>
</tr>
<tr>
<td>$q_1$</td>
<td>0.0</td>
<td>$b_{34}$</td>
<td>0.0005</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>0.0</td>
<td>$b_{35}$</td>
<td>0.002</td>
</tr>
<tr>
<td>$\alpha_2$</td>
<td>0.5E-4</td>
<td>$b_{36}$</td>
<td>0.0</td>
</tr>
<tr>
<td>$\alpha_3$</td>
<td>0.0</td>
<td>$b_{37}$</td>
<td>0.039</td>
</tr>
<tr>
<td>$b_{02}$</td>
<td>0.0</td>
<td>$b_{38}$</td>
<td>0.0</td>
</tr>
<tr>
<td>$b_{22}$</td>
<td>0.0</td>
<td>$b_{44}$</td>
<td>0.1E-5</td>
</tr>
<tr>
<td>$b_{23}$</td>
<td>0.0</td>
<td>$b_{42}$</td>
<td>0.003</td>
</tr>
<tr>
<td>$b_{24}$</td>
<td>0.0</td>
<td>$b_{43}$</td>
<td>0.0006</td>
</tr>
<tr>
<td>$b_{25}$</td>
<td>0.0</td>
<td>$b_{44}$</td>
<td>0.0</td>
</tr>
<tr>
<td>$b_{26}$</td>
<td>0.0</td>
<td>$b_{45}$</td>
<td>0.0</td>
</tr>
<tr>
<td>$b_{27}$</td>
<td>0.002</td>
<td>$b_{46}$</td>
<td>0.2E-4</td>
</tr>
<tr>
<td>$b_{28}$</td>
<td>0.0002</td>
<td>$b_{47}$</td>
<td>0.0002</td>
</tr>
<tr>
<td>$b_{03}$</td>
<td>0.0</td>
<td>$b_{48}$</td>
<td>0.0</td>
</tr>
<tr>
<td>$b_{32}$</td>
<td>0.052</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Maximum likelihood estimation also produces estimates of the variance/covariance matrix $H$. These estimates of the matrix $H$ are

$$
\begin{bmatrix}
0.00136 & 0.0014 & 0.0022 & 0.0004 \\
0.0014 & 0.0105 & 0.007 & 0.0005 \\
0.0022 & 0.007 & 0.047 & -0.0015 \\
0.0004 & 0.0005 & -0.0015 & 0.00166
\end{bmatrix}
$$

As a point of comparison the sample variance/covariance matrix of Model II, estimated in Chapter 4, is provided. This matrix was calculated by forming the product $\epsilon'\epsilon$, where $\epsilon$ denotes the residual series, and dividing by $T - n - 1$, where $T$ is the sample size and $n$ is the number of parameters in the model ($n$ is equal to 31 for Model II) (Lutkepohl 1991, page 72). Performing these calculations provides the estimates
Note that appreciable differences in variances occur between the learning model and the time-invariant Model II. Covariances are seen to be similar between both models. A decrease in variance obtained with the learning model is not surprising given that its parameters are allowed to vary over the sample.

**Diagnostics**

Following estimation, the appropriateness of the model must be determined. One useful way to ascertain whether the model is properly specified is to examine the one-period ahead forecast errors \( (y_t - y_{t-1}) \). Following Engle and Watson (1981), Bohara and Sauer (1989) utilize heteroskedasticity adjusted one-step-ahead forecast errors in testing the model. These standardized residuals may be calculated as

\[
\bar{v}_i = v_i F_i^{-\frac{1}{2}} \quad \text{where} \quad v_i = y_t - y_{t-1}.
\]

When the model is correctly specified these residuals have the property that \( \bar{v}_i \) is distributed \( \text{MVN}(0_{(4 \times 1)}, \text{I}_{(4 \times 4)}) \) (Harvey 1989, p. 256).

Given that we are using a finite sample and estimates of the unknown parameters, such a condition would hold only approximately. Using the standardized residuals generated from the model the following results are obtained:

\[
\bar{v} = \begin{bmatrix} -0.05 \\ -0.023 \\ -0.05 \\ -0.1 \end{bmatrix}, \quad S^2_v = \begin{bmatrix} 1.01 & 0.008 & -0.02 & 0.085 \\ 0.983 & -0.008 & 0.056 \\ 0.997 & 0.056 \\ 1.052 \end{bmatrix}
\]
None of the mean terms were found to be statistically significant. The estimated variance/covariance matrix is broadly consistent with what would be expected with a properly specified model.

An additional question that needs to be addressed is whether these standardized residuals are indeed approximately normally distributed. To provide an answer, a test of non-normality is performed based on Harvey (1989, p. 260). Denoting the standardized third and fourth moments (i.e., skew and kurtosis) of the standardized residuals about the mean

\[
\begin{align*}
\sqrt{b_1} &= \bar{\sigma}^{-3} \sum (\bar{v}_i - \bar{v})^3 / T^* \\
\bar{b}_2 &= \bar{\sigma}^{-4} \sum (\bar{v}_i - \bar{v})^4 / T^*
\end{align*}
\]

where \( \bar{\sigma} \) denotes the sample standard deviation, \( \bar{v} \) denotes the sample mean, and \( T^* \) is taken to equal 130, the following test may be performed:

\[
N = (T^*/6)b_1 + (T^*/24)(b_2 - 3)^2,
\]

which is distributed as \( \chi^2(2) \). A significant test statistic would imply the rejection of normality. Such a test, however, is very approximate given the small sample used in this paper. Performing this test on each standardized residual series generates the results presented in Table 5.3.

<table>
<thead>
<tr>
<th>Series</th>
<th>N</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \bar{v}_{ct} )</td>
<td>0.77</td>
<td>0.68</td>
</tr>
<tr>
<td>( \bar{v}_{ylt} )</td>
<td>6.58</td>
<td>0.037</td>
</tr>
<tr>
<td>( \bar{v}_{yl} )</td>
<td>0.59</td>
<td>0.75</td>
</tr>
<tr>
<td>( \bar{v}_{gt} )</td>
<td>5.91</td>
<td>0.052</td>
</tr>
</tbody>
</table>
Normality cannot be rejected at the 5 percent level for the series $\bar{V}_{\text{yt}}$, $\bar{V}_{\text{yt}}$, and $\bar{V}_{\text{gt}}$, and at the 3 percent level for the series $\bar{V}_{\text{yt}}$. Acknowledging the small sample utilized in estimation, these results are taken to be broadly consistent with the necessary normality.

These standardized residuals may also be used in testing for serial correlation. Univariate and multivariate Portmanteau tests for serial correlation are now considered. With a properly specified model we would expect the absence of auto- and cross-correlations among the residual series. We start by considering ACF plots with 95 percent confidence intervals for each of the adjusted residual series, which appear in Figures 5.3 through 5.6.

Figure 5.3. ACF corrected $\Delta c_t$ residuals

Figure 5.4. ACF corrected $\Delta y_{ht}$ residuals
These plots suggest that the model presented is acceptable. One caveat exists with the $\Delta y_{kt}$ equation because a significant correlation, at the 5 percent level, is found at lag 8. We choose to minimize the importance of this "spike" because of the temporal distance of the lag and also because a large correlation at lag 8 was also present in the original series $\Delta y_{kt}$. We modeled $\Delta y_{kt-8}$ in an attempt to account for any possible autocorrelation at this lag, although it is difficult to give any reasonable explanation as to why the agent would utilize such a term in determining consumption. A further problem with expanding the model is the loss of an additional year of data, which is entailed when this new term is included. Because of these difficulties, this extension will not be considered further and the significant correlation at lag 8 is simply accepted.
Ljung-Box Q statistics were also determined using the adjusted residuals. Table 5.4 presents results for each series with various lag specifications. Last, we consider a multivariate Portmanteau test to examine whether the auto- and cross-correlations among the adjusted residual series are jointly zero. As discussed in the previous chapter, calculating this statistic with a lag of 24 produced a test statistic of 339.6 with a p-value of 0.43. Regardless to whether univariate or multivariate tests are used it appears that the null hypothesis of lack of serial correlation among the residual series cannot be rejected. Again, the Q statistics related to the $\Delta y_{kt}$ equation reflect the significant spike at lag 8; however, the null hypothesis of white noise cannot be rejected at the 5 percent level. We use these results to argue for the adequacy of the learning model.

Table 5.4. Ljung-Box statistics: time-varying model

<table>
<thead>
<tr>
<th>lag 4</th>
<th>$\Delta c_4$</th>
<th>$\Delta y_a$</th>
<th>$\Delta y_{kt}$</th>
<th>$\Delta g_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.95</td>
<td>1.02</td>
<td>1.91</td>
<td>4.58</td>
</tr>
<tr>
<td></td>
<td>(0.79)</td>
<td>(0.91)</td>
<td>(0.75)</td>
<td>(0.33)</td>
</tr>
<tr>
<td>lag 8</td>
<td>6.62</td>
<td>4.79</td>
<td>13.39</td>
<td>8.28</td>
</tr>
<tr>
<td></td>
<td>(0.58)</td>
<td>(0.78)</td>
<td>(0.10)</td>
<td>(0.41)</td>
</tr>
<tr>
<td>lag 12</td>
<td>13.93</td>
<td>7.59</td>
<td>17.35</td>
<td>8.5</td>
</tr>
<tr>
<td></td>
<td>(0.31)</td>
<td>(0.82)</td>
<td>(0.14)</td>
<td>(0.75)</td>
</tr>
<tr>
<td>lag 16</td>
<td>16.09</td>
<td>11.88</td>
<td>20.18</td>
<td>15.6</td>
</tr>
<tr>
<td></td>
<td>(0.45)</td>
<td>(0.75)</td>
<td>(0.21)</td>
<td>(0.48)</td>
</tr>
</tbody>
</table>

Note: p-values appear in parentheses.

Given the adequacy of the model, an analysis of coefficients can be considered. Estimation shows that many of the elements of the matrix $Q$, that is, that variance matrix associated with the state equation, are very small or effectively zero. As discussed, coefficients which possess a very small variance would evolve as in a recursive least squares model. Nonzero variances, however, would indicate that these particular coefficients would evolve in
a very different manner. The fact that a number of coefficients possessed relatively large variances indicates the advantage of estimation under this more general approach because estimation with Friedman (1979) or Marcell and Sargent (1989) in essence constrains such variances to be zero. From Equation (5.20) when variances are nonzero, coefficients will not converge to a stable level.

Plots of how each coefficient evolves over the sample are provided in Figures 5.7 through 5.35. Beyond presenting valuable information as to whether the relevant coefficient converges to some level, these plots also allow us to see how each coefficient changes over the sample. It will prove useful in this study to see during what period large changes occur because variability may be tied to certain economic events. As discussed, one advantage of the learning model over time-invariant models is the incorporation of parameters that are able to change. Analysis of these changes may prove useful in the analysis of consumption.

![Figure 5.7. Coefficient q_{0t}](image)

![Figure 5.8. Coefficient q_{1t}](image)
Figure 5.9. Coefficient $\alpha_{1t}$

Figure 5.10. Coefficient $\alpha_{2t}$

Figure 5.11. Coefficient $\alpha_{3t}$
Figure 5.12. Coefficient $b_{02}$

Figure 5.13. Coefficient $b_{22}$

Figure 5.14. Coefficient $b_{23}$
Figure 5.15. Coefficient $b_{24}$

Figure 5.16. Coefficient $b_{25}$

Figure 5.17. Coefficient $b_{26}$
Figure 5.18. Coefficient $b_{27}$

Figure 5.19. Coefficient $b_{28}$

Figure 5.20. Coefficient $b_{33}$
Figure 5.21. Coefficient $b_{32}$

Figure 5.22. Coefficient $b_{33}$

Figure 5.23. Coefficient $b_{34}$
Figure 5.24. Coefficient $b_{35}$

Figure 5.25. Coefficient $b_{36}$

Figure 5.26. Coefficient $b_{37}$
Figure 5.30. Coefficient $b_{43}$

Figure 5.31. Coefficient $b_{44}$

Figure 5.32. Coefficient $b_{45}$
Figure 5.33. Coefficient $b_{46}$

Figure 5.34. Coefficient $b_{47}$

Figure 5.35. Coefficient $b_{48}$
In terms of general comments, what is most conspicuous about these plots and the estimated variances is that coefficient stability is associated with small or zero variance estimates. Given sufficient time, these coefficients settle to some level. This result is compared to the alternative of a nonzero variance, where coefficients appear volatile over the whole sample. Convergence in this situation does not appear likely. Note, however, that a number of coefficients display prominent downward trends, patterns which are not affected by large coefficient variances. Examples of this result are seen with \( b_{04} \) and \( b_{03} \).

As discussed, parameter convergence is of central importance in determining whether the bounded rational learning model would converge to the REE. Our results suggest that although many coefficients do indeed converge to some level, others do not. Those that do not have a tendency to wander. Such a situation suggests that the learning model specified by Equation (5.9) would not converge to the REE associated with Model 1. As discussed, Equation (5.9) is comprised of three rules which allows the agent to learn about the infinite expected sums associated with \( \Delta y_k \) and \( \Delta g_k \) and also \( E_{t-1} \Delta y_k \). A number of the coefficients in these rules are noted not to approach stable levels, suggesting the inability of \( \Delta c_t \) to approach the REE. This result provides impetus for specifying Equation (5.12) as the learning rule to be used in this chapter because of the inability of achieving the REE with Equation (5.9). As discussed previously, this rule explicitly allows expectations to be formed and utilized in the modeling of \( \Delta c_t \), regardless of whether system coefficients converge.

The possibility of nonconvergence to the REE may not, however, be of any significance. If the parameters of Model 1 are truly time-varying as suggested by the tests and results previously noted, a time-invariant notion of the REE is meaningless. Expectations generated as in Model I of Chapter 4 are wrong because the system is misspecified under the assumption of time-varying parameters. Discussion of convergence to some REE is irrelevant because such an "equilibrium" would vary over time. When viewed in this context the learning model represented by Equation (5.12) seems a more realistic alternative by which the agent would generate and utilize expectations in a dynamic economy. Application of rational expectations as utilized in the previous chapter is thus questionable.
Explicit discussion of model results may now be considered. We start by examining the parameters that directly reflect notions of costly adjustment, namely $q_{it}$, $\alpha_{1t}$, $\alpha_{2t}$, and $\alpha_{3t}$. As discussed above each of these parameters is a function of terms that may alter over the sample, which may explain changes over the sample. Estimated variances and plots of each of these coefficients reveals, however, that convergence is achieved in all cases except $\alpha_{2t}$.

While $q_{it}$ appears to settle to a certain level by the end of the sample, it is noted that it reached this level following a gradual increase that occurred during the late 1970s and early 1980s. As discussed, this parameter is a complex function of the adjustment cost ratio; the subjective discount rate, $D$; and the real rate of interest, $r$; which makes it difficult to ascertain which underlying factor may be responsible for the observed shift. Possible explanations for the shift include the advent of financial deregulation, which would alter the adjustment cost ratio, or high rates of inflation, which may have affected $r$ and $D$.

The behavior of $q_{it}$ also reveals the importance of permitting this coefficient to vary over time. As discussed, an obvious change in level occurred during the sample. Attempts at modeling $q_{it}$ as being time-invariant (as discussed in Chapter 4) may be hazardous given the structural change in one (or possibly all) of the deep parameters. Doubts may also be cast on the results of the previous chapter in testing the general model because this structural change was ignored.

The $\alpha_{it}$ coefficients are more difficult to interpret because they reflect a diverse combination of elements. Beyond incorporating the key costly adjustment term $\lambda$, coupled with $D$ and $r$, these terms also represent notions of the infinite expectations that were generated in Model I. Thus, beyond reflecting changes in the terms $\lambda$, $D$, and $r$, these $\alpha_i$'s also incorporate changes in the auxiliary system. This result offers one possible explanation as to why $\alpha_{2t}$ is much more volatile than either $\alpha_{1t}$ or $\alpha_{3t}$. By examining the estimated variances and coefficient plots, we note that the constituent elements of $\Delta y_{kt}$ are much more variable compared to the $\Delta y_{kt}$ and $\Delta g_{kt}$ equations. This result would suggest that the coefficient that
weights the expectation of this series, $\alpha_{2t}$, would also be more volatile compared to the other weighting coefficients, $\alpha_{1t}$ and $\alpha_{3t}$.

The constant term, $q_{ot}$, of the change in consumption equation also produces an interesting result. Examination of the associated estimated variance and coefficient plot reveals a coefficient that does not converge; rather, it displays a downward trend as the sample progresses. This condition suggests that the drift term associated with the change in nondurable consumption is not fixed and that attempts to impose a fixed mean level, or in the case of the previous chapter, subtracting such a term to alleviate constants is potentially misleading. It appears more appropriate to allow the drift term to vary as the sample progresses, thus admitting the dynamics of this coefficient.

Interesting results are also found in the auxiliary system. We begin by considering the constant terms $b_{02t}$, $b_{03t}$, and $b_{04t}$, which model the mean of each respective series. The term $b_{02t}$ displays a large decrease after the first oil crisis, followed by a more gradual decrease from the late 1970s through the early 1980s. Note that this drift term settles to a stable level by 1984. This time path suggests that the mean of $\Delta y_{it}$ decreases following economic adversity. The term $b_{03t}$ displays a similar pattern, at least until the early 1980s, with large decreases followed by a gradual increase until the late 1980s. Last, we consider the mean of the $\Delta g_{it}$ series, $b_{04t}$, which possesses an interesting time path in that it declines throughout the sample. A long decline is noted until 1976, followed a large drop in 1980 and a precipitous fall during 1991. Such behavior reveals the hazards of simply assuming that such a coefficient is constant, as in the time-invariant models. By allowing this term to alter, a truer image arises.

Finally, we consider parameters of the lagged variables in the auxiliary system. Most of these terms converge to certain levels as the sample progresses. An interesting exception is found with the coefficients associated with $\Delta g_{it-2}$ in each of the three equations; that is, $b_{27t}$, $b_{37t}$, and $b_{47t}$. All these terms display behavior which alters considerably throughout the sample, with a pronounced decrease occurring after the first oil crisis. The time plot of $b_{24t}$ also reveals an example of a coefficient that wanders extensively throughout the sample.
We also note that the coefficients of the $\Delta y_{kt}$ equation appear much better behaved compared to those of $\Delta y_{kt}$ and $\Delta g_{kt}$. Time plots reveal that most of the constituent coefficients, except $b_{20t}$, $b_{27t}$, and possibly $b_{28t}$, remain at constant levels quite early in the sample. This result compares to the parameters of $\Delta y_{kt}$, which appear more volatile. The change in government expenditure, $\Delta g_{kt}$, seems to lie between these two points. These results suggest that $\Delta y_{kt}$ is subject to greater change over the sample compared to $\Delta y_{kt}$. Obvious implications also hold for $E_{t-1} \Delta y_{kt}$, which would be expected to display greater evolution over the sample compared to $E_{t-1} \Delta y_{kt}$ and $E_{t-1} \Delta g_{kt}$. One possible explanation is the continual increase in government expenditure and greater reliance upon deficit financing, which would affect the variable $\Delta y_{kt}$ because, as discussed in Chapter 3, the variable $y_{kt}$ reflects capital income net of government bonds.

Two important results emerge from this time-varying analysis. First is that a number of coefficients in the auxiliary system fail to converge, suggesting that the REE will not be achieved. As discussed, coefficient stability is a necessary condition for studying convergence. More important, coefficient instability may suggest that the REE itself varies with time. The modeling of expectations as described in the previous chapter would thus be incorrect because replacement of expectations relied upon coefficient stability. A second result deals with the evolution of the coefficients $q_{1t}$ and $\alpha_{2t}$. The coefficient $q_{1t}$ displays a sharp increase following the late 1970s suggesting some type of structural change, whereas $\alpha_{2t}$ possesses a distinctive upward trend. Attempts at modeling these terms as time-invariant are inappropriate because each is subject to appreciable change over the sample.

These results are important when discussing the strong acceptance of the general model found in the previous chapter. Given the possibility that the REE is time-varying and certain coefficients in the consumption equation may change over the sample, this strong acceptance is questionable. A more realistic alternative is to test the general model in the time-varying/learning framework. Estimation within this environment explicitly allows for changing expectations and coefficients, thus providing a more realistic view of the general, costly
adjustment model. Tests of the general model assuming the time-varying/learning framework are considered in the following subsection.

**Testing Special Cases**

We now consider the validity of the general model in the context of this chapter. By performing tests of the special cases in a time-varying/learning framework we can incorporate changes in the system directly into the analysis. Ignoring the time-varying nature of the system, as is done in the previous chapter, may have had a significant effect in the strong acceptance of the general model.

To study the general model in the context of the time-varying/learning framework, we consider a simplistic approach in testing the special cases $\lambda_1 \rightarrow 0$ and $\lambda_1 \rightarrow 1$. The case $\lambda_1 \rightarrow 0$ is considered first. As noted in Chapter 4, imposition of this case implies that consumption can be adequately represented by a random walk. Tests performed in the previous chapter operated by comparing the value of the log likelihood under the unrestricted model (the general model) to that obtained when $\{\Delta y_{t-1}, \Delta y_{t-2}, \Delta y_{t-3}, \Delta g_{t-1}, \Delta g_{t-2}, \Delta g_{t-3}\}$ are omitted from the model (i.e., consumption follows a random walk) and using a likelihood ratio test. Such an approach, unfortunately, does not hold in the time-varying framework. In the time-varying model presented in this chapter, the supposition that consumption follows a random walk implies that the true coefficients, $q_t, \alpha_{1t}, \alpha_{2t},$ and $\alpha_{3t}$ (the coefficients of $\{\Delta c_{t-1}, \Delta y_{t-1}, \Delta y_{t-2}, \Delta y_{t-3}, \Delta g_{t-1}\}$) are equal to zero for all time periods. Testing this hypothesis is not straightforward because the restriction must be imposed over all periods. Application of the standard likelihood ratio test in this case appears to be unclear. Although the imposition of restrictions over all periods has been addressed in the literature [for example, Doran (1992)], testing these restrictions has yet, it is believed, to be done.

One possible method of imposing and testing the supposition that consumption follows a random walk is to allow the relevant coefficients to evolve as
while assuming that all the other coefficients in the model evolve as a standard random walk, as was done previously. Tests may then be performed by estimating the model with the $a_i$’s unconstrained and making a comparison to the likelihood achieved when the $a_i$’s are constrained to be zero. This constrained case would imply that the relevant parameter would on average be equal to zero. However, comparing such likelihoods with a standard likelihood ratio test may not be appropriate in this case because the distribution of this test statistic is not clear. These issues are discussed in section 5.1.2 of Harvey (1989).

Because of these problems in explicitly testing whether consumption follows a random walk in this time-varying framework, we suggest a more indirect approach. In particular, we examine 95 percent confidence bands around the coefficient estimates of $q_{1t}$, $\alpha_{1t}$, $\alpha_{2t}$, and $\alpha_{3t}$, which may be denoted as $q_{1t}$, $\alpha_{1t}$, $\alpha_{2t}$, and $\alpha_{3t}$. Variances used in generating these bands are from the relevant diagonal element of the $P_t$ matrix previously discussed, which is the variance/covariance matrix of the coefficient estimates. Coefficient estimates used in these tests are the minimum mean square linear estimators (MMSLE) of the true coefficients based on information up to and including time $t$. The estimators are unconditionally unbiased when the unconstrained covariance matrix of the estimation errors is $P_t$ (Harvey 1992, p. 90). Using these estimates, the 95 percent confidence intervals would provide some indication as to whether the true coefficients are significantly different from zero at any particular time period.

This approach is more heuristic compared to explicit tests because we cannot say whether all the coefficients are jointly equal to zero as required by the random walk hypothesis. Rather, these plots will give a rough idea of whether each coefficient estimate is individually different from zero. Plots of the coefficient estimates $q_{1t}$, $\alpha_{1t}$, $\alpha_{2t}$, and $\alpha_{3t}$, along with confidence intervals, appear in Figures 5.36 through 5.39.
Figure 5.36. Coefficient estimates and confidence intervals: $q_{1t}$

Figure 5.37. Coefficient estimates and confidence intervals: $\alpha_{1t}$

Figure 5.38. Coefficient estimates and confidence intervals: $\alpha_{2t}$
In analyzing these plots, what is most obvious is the insignificance of the estimates for many time periods. The estimator $q_{1t}$, the coefficient of $\Delta c_{t-1}$, is seen to be significantly different from zero at the 5 percent level most of the time following the early 1980s, but this estimator is insignificant prior to the 1980s. Coefficient estimate $a_{2t}$ also displays relatively brief periods of significance, though over most of the sample this estimate is insignificant. The coefficient estimates $a_{t1}$ and $a_{3t}$ each appear insignificantly different from zero over the whole sample; however, the insignificance associated with $a_{t1}$ is marginal.

Although these plots are not able to state whether all coefficient estimates are jointly different from zero, such a strong statement may not be necessary. To reject the notion that consumption follows a random walk all that is needed is for some variable to enter significantly into the consumption equation. If at least one coefficient can be found to be significantly different from zero for at least one period, evidence against the random walk hypothesis over the complete sample would be obtained. Note, however, that if the t-statistics suggest that each coefficient is insignificantly different from zero we cannot infer that these four coefficients are jointly zero.

In studying the possibility that consumption follows a random walk for all periods, we concentrate on the coefficient $q_{1t}$. The plot in Figure 5.36 suggests that in periods following the early 1980s this coefficient estimate is significantly different from zero at the 5 percent
level. This result provides sufficient information to conclude that consumption is not a random walk process, at least when viewed relative to the complete series. Such a statement can be made regardless of the significance, or insignificance, of the estimates $\alpha_{1t}$, $\alpha_{2t}$, and $\alpha_{3t}$.

Rejection of the special case $\lambda_i \rightarrow 0$ allows us to conclude that $0 < \lambda_i < 1$, which suggests that consumption relies on more information than simply its lagged level when a time-varying/learning framework is considered.

A similar exercise can be performed in testing the special case $\lambda_i \rightarrow 1$. As discussed in the previous chapter, this case suggests that the change in consumption is a function of the lagged change in consumption. We utilize the special case as derived from the utility function derivation which does not constrain the coefficient on $\Delta c_{t-1}$ to be 1; that is,

$$\Delta c_t = q_0 + q_1 \Delta c_{t-1} + \varepsilon_t.$$  

Now, if additional variables were added to this model, we would expect the associated coefficients to be insignificantly different from zero. In studying this possibility, we can use the plots for the estimates $\alpha_{1t}$, $\alpha_{2t}$, and $\alpha_{3t}$ (Figures 5.37 through 5.39) because if the special case $\lambda_i \rightarrow 1$ is appropriate the coefficients associated with $\{E_{t-1} \Delta y_{ht}, E_{t-1} \Delta y_{at}, E_{t-1} \Delta g_t\}$ should be jointly zero. As discussed, the 95 percent confidence bands related to $\alpha_{2t}$ appear to contain nonzero values for at least some periods in the sample, implying that the special case $\lambda_i \rightarrow 1$ does not hold over the whole sample. It may thus be concluded that the general model $0 < \lambda_i < 1$ cannot be rejected when the complete sample is considered.

We must stress that the analyses just described are relative to a particular portion of the sample. Thus, while it would seem that the special case $\lambda_i \rightarrow 0$ can be rejected over the entire sample, there may be some periods in which it may be appropriate, for example, before the 1980s. We are purposely imprecise here because we cannot conclude that all coefficient estimates are jointly zero on the basis of the t-statistics already presented. However, this result suggests that it may be inappropriate to discuss consumption patterns as some time-invariant concept. Economic reality may necessitate consumption models that may change over the sample to reflect the altering environment. Consumption models that ignore these possibilities,
such as those discussed in the previous chapter, may present results that mask the true behavior of the agent. In particular, such an approach ignores the possibility that consumption may evolve as a random walk for some periods during the sample. Estimation as performed in this chapter allows for such possible evolution in consumption, although tests of whether this evolution actually occurred cannot be handled with the tools on hand.

By explicitly allowing the parameters $q_{rt}$, $\alpha_t$, $\alpha_{rt}$, and $\alpha_{rt}$ to vary to reflect possible changes in $r$, $D$, and the ratio of costs ($\lambda$), as done with the time-varying/learning model, the effects on consumption of an evolving economic system can be realized. Given the amount of change which the US economy has incurred over the sample, changes in these parameters would be expected. It may thus be quite possible for consumption to evolve as a random walk during one segment of the sample, only to revert to the general model ($0<\lambda_t<1$) at some later date. The conclusion drawn in the previous chapter that the agent is subject to costs of adjustment while formulating a consumption path may thus be considered an overly strong statement. Whereas it is true that the general model cannot be rejected relative to the whole sample, there may in fact be some periods in which it does not hold. Unfortunately, the $t$-statistic approach utilized above cannot be used in identifying these periods.

Conclusions

We believe that learning is important in modeling consumption because of the assumptions upon which the REH is constructed and given the possibility that the parameters of Model I and Model II are time-varying. Rational expectations as used in Chapter 4 assumed that the agent knows the structure of the model and the necessary model parameters. Required knowledge relied upon the supposition that the agent has been operating in the economy sufficiently long to have discovered the model structure, including coefficient estimates. In estimation, this supposition allowed for unknown expectations to be solved out because
coefficients could be considered fixed over the sample. What we question in this chapter is how such an assumption can be made given the sample sizes that are typical in macroeconomic analysis. Recent history has revealed an economic system that has undergone significant change in the past twenty years. Viewed relative to economic reality, the assumption of understanding appears exceedingly strong.

Expectations in this chapter have been generated in a more realistic manner. Instead of assuming that the agent "knows" the system, we will allow the agent to build up knowledge as the sample proceeds thus allowing for fundamental change to be incorporated as it occurs. Expectations are formed by allowing the agent to learn about the system over time. Learning as utilized currently applies some type of updating scheme to obtain parameter estimates using only information available at that time. Conditions under which the learning rule may converge to the rational expectations equilibrium were also discussed.

Specifying Equation (5.9) as the learning rule was found to be inappropriate because a number of coefficients that comprise the rule were found to be quite volatile, suggesting that convergence to the REE will not occur. However, the possibility of time-varying coefficients renders the discussion of an equilibrium pedantic, for such a state would change as the underlying coefficients changed. Given this situation, an alternative method of incorporating learning was discussed. This alternative, represented by Equation (5.12), allowed for the direct incorporation of expectations, acknowledging a changing expectation-generating mechanism. In particular, the change in consumption was assumed to evolve as

$$\Delta c_t = q_0 + q_{11} \Delta c_{t-1} + \alpha_{11} E_{t-1} \Delta y_{t-1} + \alpha_{21} E_{t-1} \Delta y_{kt} + \alpha_{31} E_{t-1} \Delta g_t + \epsilon_t.$$ 

A time-varying auxiliary system was also specified to determine the unknown expectations. Estimation with this model was noted to be very flexible and to offer a reasonable explanation as to how consumption may evolve in an altering system. By permitting the auxiliary system to be time-varying, a constant REE is not presupposed, a situation which forces the agent to learn about the model parameters over time. To add more generality to the model, coefficients were assumed to evolve as random walks.
Specifying model parameters as evolving as random walks represents a more general approach at specifying learning compared to the methods of Friedman (1979) and Marcet and Sargent (1989). Using these processes and estimating with the Kalman filter allowed the notion of continual change to be introduced into the system. In specifying this process, however, other estimation methods may arise in the case when the estimated variances associated with coefficients approach zero. Thus, parameter evolution as presented by Friedman (1979) and Marcet and Sargent (1989) will only arise if warranted by the data. It is recalled that following estimation a number of estimated variances were found to be relatively large and generate significant movements in the relevant parameters over the sample. By ignoring these variances, as in essence Friedman (1979) and Marcet and Sargent (1989) do, important effects related to coefficient estimation may be missed.

Beyond providing for a more realistic generation of expectations, the incorporation of time-varying parameters also seems theoretically possible in the consumption equation because of the nature of the coefficients $q_{1t}$, $\alpha_{1t}$, $\alpha_{2t}$, and $\alpha_{3t}$. As shown in Chapter 4, these coefficients are not “deep,” because they are complicated functions of $\lambda$, $D$, and $r$. Changes in any of these quantities would result in movements in the coefficients. With an evolving economy, it would thus seem likely that these estimated coefficients would change over the sample, which we explicitly allow in this chapter. Following estimation, significant changes were noted for $q_{1t}$ and $\alpha_{2t}$. Attempts at forcing these coefficients to be constant, as was done in the previous chapter, may have a detrimental impact upon estimation results.

Estimation utilized a Kalman filter, which provides an efficient solution technique once coefficients are allowed to evolve as random walks. Recursive least squares may arise as a special case when the estimated variance of a particular coefficient is very small. Maximum likelihood estimation reveals that a number of coefficient variances were relatively large, suggesting the importance of a more general approach in modeling coefficients. After using various diagnostics to determine the adequacy of the model, estimated coefficients could be studied.
Parameters of primary concern are those which are elements of the rule, namely, \( q_0, \) 
\( q_1, \alpha_1, \alpha_2, \) and \( \alpha_3 \) from Equation (5.12). The terms \( q_1, \alpha_1, \) and \( \alpha_3 \) all are seen to converge, while the terms \( q_0 \) and \( \alpha_2 \) displayed a downward trend and upward trend, respectively. The evolution of the costs associated with consumption are most clearly displayed in the time path of \( q_1, \) a plot which reveals a large jump during the early 1980s possibly reflecting the start of financial deregulation that would have an impact on the costs faced by agents in determining consumption. Whatever the cause, the changes in the level of this coefficient reveal the danger of assuming that this term is fixed over the sample. From the previous subsection, this change in level is also associated with \( q_1, \) becoming significantly different from zero at the 5 percent level. The \( \alpha_1 \) terms are more difficult to interpret; however, the weighting term of \( E_t \Delta y_{kt} \) is more volatile compared to the other \( \alpha_n \) terms. This result is expected because of all the equations in the auxiliary system \( \Delta y_{kt} \) displayed the greatest change over the sample.

Interesting results are also found in the auxiliary system, where a number of coefficients are seen not to converge. Even if coefficients do settle to some level as the sample progresses, a number of these display significant alterations in level at certain points in time. The ability to incorporate this information into the formation of expectations illustrates the appeal of the learning model. Expectations as generated by this auxiliary system will thus alter to reflect the changing coefficient estimates. This concept is important because if the expectation-generating mechanism changes through time, expectations as generated in Chapter 4 are incorrect because we implicitly assumed that such a mechanism was fixed.

Given the possibility of large changes in a number of parameters in the estimated system, two important implications arise from the work considered in this chapter. The first implication is the distinct possibility that expectations as generated and used in Models I and II are wrong. The inability of all system coefficients to converge to some level, which is at least a necessary condition, seems to preclude a learning rule from converging to the REE. By having coefficients that display large movements over the sample, any discussion of an "equilibrium" is meaningless, because it too would change over time. This result suggests that
the generation of expectations in the previous chapter, which relies on some stable parameter set, is incorrect given the amount of change that this chapter has shown to exist in the system.

Because of the possibility of a changing expectation-generating mechanism, it seems more appropriate to formulate expectations in the learning framework discussed in this chapter. By accounting for any change, the agent is allowed to apply any alterations in the system into the determination of consumption. Determining consumption in the learning framework may thus be a more realistic characterization compared to the models presented in the previous chapter. If estimation is confined to fixed coefficients, however, it would seem that Model II may be more appropriate because the restrictions imposed by that model may be less egregious, given changing parameters, compared to the complex restrictions imposed by Model I. Given the possibility of changing parameters, a more simplistic, albeit theoretically incorrect, fixed-coefficient model may be more realistic than another theoretically correct model that is more susceptible to parameter misspecification.

The second implication related to the work of this chapter deals with the inferences obtained from the previous chapter in testing the general, costly adjustment model. Given that expectations generated and used by Models I and II are possibly incorrect and given the large changes observed in the coefficients \( q_{tt} \) and \( \alpha_{2t} \), the strong acceptance of the general model found in the previous chapter is open to question. More appropriate tests of the costly adjustment model seem to be provided with the time-varying/learning model because beyond allowing for a more correct formation of expectations, coefficients in the consumption equation are permitted to change to better reflect changes in the economy.

Testing the special cases \( \lambda_1 \to 0 \) and \( \lambda_1 \to 1 \) proceeded in this chapter by producing 95 percent confidence intervals around the coefficient estimates \( q_{lt}, \alpha_{lt}, \alpha_{2t}, \) and \( \alpha_{3t} \), which suggested the rejection of the two special cases. Thus, it can be concluded that the general model \( 0<\lambda_1<1 \) is appropriate, at least at the 5 percent level of significance, when viewed relative to the whole sample. As discussed in the previous chapter, both costs of adjustment should be taken into account when agents formulate a consumption plan. However, these results also suggest the possibility that the consumption equation may evolve over the sample.
In particular, it seems quite possible for consumption to evolve as a random walk during the earlier part of the sample, only to switch to the general model later. Unfortunately, the t-statistics used in this chapter are not appropriate in determining the validity of this hypothesis. More advanced tools are needed before any definitive statement can be made.
CHAPTER 6: CONCLUSION

Consumption theory has been and always will be, an important area in economic analysis. Over the past 40 years, consumption theory has evolved from simple ad hoc specified models to more complex theories based on microfoundations. Such evolution, however, has been paid for in terms of strong assumptions. These assumptions have received close scrutiny in the literature because the behavior of consumption is closely tied to how much is assumed about the market and the agent. Today’s models, though more theoretically pleasing, have resulted in a large amount of structure imposed upon the system. Many researchers have shown that once this structure is reduced, consumption may evolve as in less sophisticated models.

This paper carries out consumption determination in a permanent income hypothesis (PIH) framework. Created by Milton Friedman, this hypothesis has received much attention by economists since its introduction in 1957. One reason for this popularity is that the PIH seems intuitively plausible. The assumption that rational agents use expected future earnings and current wealth in arriving at an optimal consumption path seems quite believable. Using a time-series representation of the PIH, Hall (1978) shows that the implication of this hypothesis is that consumption follows a random walk. Arrival at this cogent result, however, requires the imposition of strong assumptions upon the agent and the economy.

While the basic supposition of the PIH that the agent uses permanent income in formulating consumption is accepted, the ability of the individual to implement this suggested consumption path is questioned. This paper expresses the belief that capital market imperfections may prohibit the agent from consuming according to the predictions of the PIH. Although we assume that the agent may formulate a consumption path in accord with the PIH, it is only when the agent attempts to implement such consumption predictions that the true imperfect nature of the economy is fully realized. Consumption as predicted by the PIH
may represent a desirable amount because in an uncertain environment some notion of permanent income may represent a best guess of future consumption ability.

In studying consumption within the framework of the PIH, we started with a brief recapitulation of Friedman's original work. By basing consumption on some notion of permanent income the naivety commonly associated with the Keynesian view of consumption determination could be removed. Consumption would now be determined in an optimization framework, with the rational expectations hypothesis (REH) furnishing the tools necessary for the PIH to become fully operational. It was in this environment, using a time-series representation that Hall (1978) made his famous random walk supposition.

Testing the implications of the time-series representation of the PIH has been the focus of much attention on the part of macroeconomists. Chapter 2 provided a brief review of the literature in this area, which provided evidence against the PIH. Contrary to the implications of the PIH as formulated by Hall (1978) and Flavin (1981), we found that more information than simply the lagged level of consumption was important in predicting current consumption. Additional evidence against the PIH was also provided because the predicted smoothness of consumption relative to income is not observed when a difference stationary notion of income, which seems to be consistent with the data, is used.

With the empirical shortcomings of the PIH comes the desire to provide some sort of explanation. Chapter 2 concentrated on the assumption of perfect capital markets used in formulating the PIH. Under this assumption are ideas that the agent can borrow as much as desired, that the lending rate is the same as the borrowing rate, that information is perfect and free, and that there are no transactions costs. In testing for the possibility of market imperfections, Chapter 2 noted two approaches. The first specified some imperfection ex ante deemed relevant by the researcher in affecting consumption. Using this specified imperfection, a model may then be constructed which will allow for the determination of consumption, given the existence of such a phenomenon. The noted disadvantage with this approach was the assumption that the agent could recognize and incorporate that imperfection. In a complex and uncertain economy this action may be a difficult or an impossible task. A second approach
at modeling market imperfections assumed that one segment of the population is liquidity constrained, that is, unable to borrow sufficient amounts for whatever reason. An advantage of this approach compared to the former was that exact specification of what imperfection the agent may be reacting to was not needed, thus reducing the degree of knowledge presupposed upon the agent. Problems with this "liquidity constrained" approach, as noted in Chapter 2, are the implicit assumption that one segment of the population is completely unaffected by imperfections and that only those market failures that result in liquidity-constrained behavior are relevant when discussing consumption determination.

Because of the assumptions required in incorporating market failures into the modeling of consumption, this paper has offered an alternative approach: modeling the costs associated with market failures. We started the study of costly adjustment by introducing the data to be used in the analysis. Chapter 3 began with a reformulation of permanent income that allowed for more information to be incorporated into its derivation, an alternative form that was used throughout the paper. Definitions and analysis of statistical properties of the data then followed. Tests of these data revealed that our notion of consumption is consistent with that observed in Chapter 2, namely, that the PIH is incorrect as it is currently formulated. This conclusion supplied the rationale in providing an alternative consumption model.

With the rejection of the PIH as seen in Chapter 2 and with the use of our data in Chapter 3, this paper offered an alternative approach to modeling consumption. This alternative, which forms the basis of Chapter 4, centered on incorporating the costs that may be generated by market failures as the agent attempts to formulate a consumption path. This costly adjustment approach imposed a lower informational burden compared to explicitly specifying particular imperfections, while providing a more general view of market failures as compared with the liquidity constrained approach. By studying the effect of market failures rather than the multitude of causes, our representative agent need not be omniscient.

In this costly adjustment framework, the agent is assumed to formulate desired consumption in accord with the PIH. Given economic uncertainty consuming according to the level of permanent income may be taken as representing a best guess because market failures
may only be realized when the agent attempts to implement this desired consumption plan. Costs in this framework may arise from two sources: the first arises from the deviation between desired and actual consumption levels and the second is incurred as the agent attempts to change consumption levels too abruptly.

Deviations between desired and actual consumption may generate a cost if the agent takes preliminary steps in implementing desired levels. For example, if the agent spends time and money in performing market research or makes contracts under the supposition of projected levels, the reality of being unable to obtain these desired amounts may leave the agent with additional costs, as the agent attempts to reformulate a consumption path. Altering consumption plans because of market imperfections is thus assumed to be costly. The second cost considered arises from changing consumption levels because market imperfections may introduce rigidities into the system that prohibit instantaneous change. Transaction costs, credit rationing, and imperfect information may all conspire to make abrupt change in consumption costly. Large changes will thus be avoided as the agent attempts to spread alterations in consumption levels over a number of periods.

Chapter 4 introduced two models that allow for these two costs to be incorporated into the determination of consumption. A quadratic loss function was used by Model 1 to incorporate the costs in a very natural and amenable manner. Alternative consumption models were also seen to be nested within this more general framework. The first special case considered was the PIH, which suggested that consumption could be adequately modeled as a random walk. Another special case [a model similar to that discussed in Cushing (1992)] was discussed where the costs associated with altering levels are of primary importance. A general model represented the situation where both costs of adjustment are significant when determining consumption. Derivation of these results was also found when costs of adjustment were incorporated through the utility function. The special case $\lambda_1 \rightarrow 1$ corresponding to Cushing (1992), however, was slightly different.

An important result of Chapter 4 was that given the existence of the two costs, current consumption was found to be a function of more than simply its lagged level. In particular, the
change in consumption is a function of its own lag and the expected change in permanent income. Explicitly, tests of the two special cases may then be performed to study the appropriateness of this general model. Results of these tests using the data of Chapter 2 provided strong evidence to suggest the existence of costly adjustment.

Deriving Model I necessitated strong assumptions upon the agent's understanding of the economic system. By invoking rational expectations, it was assumed that the agent possessed a correct understanding of the system, including the complex cross-equation restrictions implied by the model and the REH. Because of the level of complexity required by Model I, a simpler model, Model II, was introduced. This model assumed that the agent used a simple linear rule in specifying permanent income, which implied that model restrictions were much more transparent compared to the former model. Using a linear rule as an approximation to the infinite expected sums of Model I, Model II may be viewed as imposing a lower level of knowledge upon the agent while reducing the degree of complexity.

Estimation results of Model II, however, were very similar to those obtained under Model I. Both special cases were strongly rejected, suggesting that both costs of adjustment should be used in properly modeling the change in consumption. An important result of Model II is that the assumed naïveté relative to the agent's understanding of the system is not injurious in the costly adjustment framework. Results are the same whether the true or a simplistic approximation of permanent income is used.

The basic conclusion of Chapter 4 is that regardless of which model was used for estimation, evidence suggested that the two costs of adjustment were significant. Thus, both the cost associated with deviations from desired consumption levels and that arising from altering consumption levels have a significant impact upon the consumption decision. In this context, the existence of these costs suggests that the agent encounters capital market imperfections when formulating a consumption path. Precise sources or the precise nature of these imperfections cannot be stated because the agent may not fully understand the market failures or combination of market failures that may confront him. With an agent who is not
omniscient we simply conclude that the existence of these imperfections has manifested itself as the two costs previously discussed.

An important implication of consumption as derived in Chapter 4 deals with the ability of government intervention in altering the optimal level of consumption. Given the acceptance of the costly adjustment model, anything that may alter the agent’s perceptions of permanent income will affect the optimal consumption path. Government policies or actions that are perceived as changing future levels of gross labor income, capital income, and government expenditures will all affect the optimal consumption path. Changes in costs of adjustment through, for example, financial deregulation will also affect consumption through the terms $q_1$ and $q_2$.

What is interesting in these results is that even though rational expectations are assumed, the existence of capital market imperfections is sufficient to ensure that government policies will have an effect on consumption. The ability of government intervention in changing consumption is possible even though the rationality of the forward-looking agent is preserved, for even though the agent still looks towards the future in formulating current consumption, imperfect markets negate the perfect smoothing aspect of the PIH. Once these imperfections are acknowledged, the possibility of policy altering consumption is clearly seen.

Although strong acceptance of the general model was provided regardless of whether Model I or II was used or whether cointegration was utilized or not, one important caveat was found in each model. Likelihood ratio tests dealing with the constancy of coefficients over the sample suggested that, for both Models I and II, coefficients may have changed over the sample. With such a possibility, the strong acceptance of the general costly adjustment model is questionable. Strong acceptance, possibly, may have had more to do with inappropriate expectation formation or shifts in coefficients than with the true validity of the model. Effects of changing coefficients proved to be important in Chapter 5.

Specifying Model II served as an introduction to Chapter 5. Using Model I or Model II to describe the evolution of consumption makes assumptions that many researchers simply accept, namely, the assumption of rational expectations. Although this assumption allows
unknown expectations to be generated in a very tractable manner, it requires that a number of strong assumptions be imposed upon the agent. For example, in formulating Model I, the agent is assumed to know the specification of the auxiliary system and how the infinite expected sums that describe permanent income are found. Additionally, solving out expectations requires that system parameters are fixed and known over the sample. Questions as to how the agent attains this information are avoided by simply assuming that the model has been operating indefinitely.

Chapter 5 offered a framework where the strong informational assumptions of the REH could be relaxed. Instead of assuming that the agent knows the complete system, a learning model was introduced which allowed the agent to derive knowledge about the system as the sample progressed. As utilized in Chapter 5, learning started by specifying a simple linear rule which the agent may use in determining expectations while obtaining coefficient estimates with some form of recursive least squares estimation. Coefficient estimates formed in this manner will only incorporate information up to the point at which they are formed. Convergence of the rule to the rational expectations equilibrium (REE), as discussed by Marcet and Sargent (1989), may occur if these estimated coefficients converge to some level.

Because model coefficients may be subject to continual change over the sample, however, an alternative learning rule was specified. An alternative was desired because if some of the coefficients do not approach some stable level, the specified learning rule [Equation (5.9)] will not converge to the REE. Although this result would be unfortunate if coefficients are fixed through time, it is irrelevant if some of the system coefficients are truly time-varying. In this case, the REE itself would alter through time, making any notion of convergence meaningless. Because of the possibility that coefficients may vary continuously, an alternative learning rule which explicitly incorporated the terms \( E_{t-1} \Delta y_t, E_{t-1} \Delta y_{kt}, E_{t-1} \Delta g_t \) was presented. By directly modeling these terms, any changes that may occur in the expectation-generating mechanism can be directly taken into account when describing \( \Delta c_t \).

Beyond offering an alternative learning model, continuous change was introduced in Chapter 5 under the assumption that coefficients evolve as random walks. This is a more
general approach compared to Marcet and Sargent (1989), because coefficients as specified in that paper may arise as a special case. With random walk coefficients, however, recursive least squares estimation is no longer appropriate, necessitating estimation with the Kalman filter. Following maximum likelihood estimation and the analysis of various diagnostics, examination of the time-varying/learning model could be undertaken. Two important results were found from estimating the change in consumption in this environment. The first point is that a number of coefficients in the auxiliary system fail to converge to some steady level. As discussed, if such coefficients are elements of a learning rule, convergence to the REE cannot occur, at least using the technique of Marcet and Sargent (1989). This problem may not be critical, however, because the standard application of the REH is not appropriate for coefficients that are truly time-varying, suggesting that expectation formation as used in Equation (5.12) may be more proper. The possibility of time-varying coefficients suggest that implementation of the REH as used in Chapter 4 is incorrect because that particular application is dependent upon fixed coefficients.

The second important result involves testing the general model against the two special cases. Results of Chapter 5 question the strong acceptance of the general model found in Chapter 4. Beyond the possibility that the wrong expectations were used in generating these results is the possibility that the coefficients in the consumption equation \(q_0, q_1, \alpha_1, \alpha_2, \text{ and } \alpha_3\) may have changed over the sample. The coefficients \(q_0\) and \(\alpha_2\) reveal behavior that is anything but steady, and \(q_1\) displays a large increase during the late 1970s through the early 1980s. These results suggest that testing the general model in a time-invariant model of Chapter 4 may be hazardous. A more realistic test may be obtained when the time-varying/learning framework is utilized. Tests of the special cases \(\lambda_1 \rightarrow 0\) and \(\lambda_1 \rightarrow 1\) in such an environment showed that the general model could not be rejected at the 5 percent level when the complete sample is considered. An interesting possibility was found, however, in that the consumption equation itself may have shifted over the sample. In particular, it is possible that consumption may have evolved as a random walk before the 1980s. Unfortunately, the tests utilized in this chapter are unable to validate this supposition.
In studying consumption, we have sought to reduce the level of knowledge imposed upon our representative agent. Following the rejection of the permanent income hypothesis, capital market imperfections were introduced into the analysis of consumption by studying the costs that these imperfections may present to the agent. This notion of costs permitted market failures to be modeled without assuming that agents can correctly identify all the imperfections that may confront them. By reducing the amount of knowledge imposed upon the agent, we acknowledge that in a complex economy precise information about such phenomena may be impossible. By being unable to identify all the failures that may affect them, agents may react to the effects that such phenomena may generate.

Reducing the level of knowledge imposed upon the agent showed that consumption does not achieve the cogent form derived by Hall (1978). Rather, consumption was found to be a function of its lagged change and the expected change in permanent income. With the assumption of rational expectations this consumption equation is seen to have very much a Keynesian flavor, because the lagged consumption change along with changes in labor income, capital income, and government expenditure are all found to be important. However, Chapter 4 still made the very strong assumption of rational expectations.

Whereas many papers have been produced which have studied various assumptions about the PIH, relatively few have considered the assumption of rational expectations. Chapter 5 presented an alternative to the strong informational assumptions of the REH. Instead of assuming that the agent knows the system, the more realistic supposition is made that knowledge is only accumulated as the sample progresses. Presenting consumption determination in such a framework acknowledges that the agent may possess less than complete information. By studying the notion of costly adjustment in a time-varying/learning framework, consumption determination under the conditions of imperfect knowledge and information can be modeled in a very flexible manner.

We end this paper by considering some areas for future research. One extension which seems interesting in terms of the work of Chapter 4 is the incorporation of transitory consumption. Chapter 4, following Flavin (1981), explicitly assumed that the transitory
component of consumption was absent, implying that $E_{t-1}y_{pt} = y_{pt-1}$. Given the existence of transitory consumption, as studied in Sargent (1989) and Falk and Lee (1990), $E_{t-1}y_{pt}$ is shown to be equal to $y_{pt} + a_i$, where $a_i$ denotes a disturbance term. Incorporating this more general notion of consumption into the costly adjustment framework derived in Chapter 4 is relatively straightforward and may provide additional insight into how costs of adjustment may react in the presence of a more general definition of consumption.

Another important point is in choosing among a set of learning rules. Choosing among rival rules has yet to be explicitly discussed in the literature. A more in-depth study of learning would seem to benefit from some methodology of choosing from some set of possible rules. This concept seems especially important in the case where rules may not converge to the REE because of parameter instability. In this case adoption of one particular rule may have a significant impact upon estimation and thus be very pertinent to the analysis.

Our final suggestion for further research is a more appropriate method to test for restrictions in a time-varying system. In Chapter 5, we found ourselves unable to test the possibility that consumption may have evolved as a random walk during the early part of the sample. Development of a proper test would allow us to explicitly state whether such a change in structure occurred. Performing such tests would permit a stronger statement to be made about the evolution of consumption.
BIBLIOGRAPHY


Flavin [1981, p. 989, Equation (36)] specifies the following equation

\[(A1) \quad c_t = c_{t-1} + r \left[ \sum_{i=0}^{\infty} \lambda^{i+1} (E_t - E_{t-i})y_{t+i} \right] - (1 + r)u_{t-1} + u_t,\]

where \(c_t\) denotes nondurable consumption, \(y_t\) is personal disposable income, and \(\lambda = \frac{1}{1+r}\).

Suppose that \(y_t\) is difference stationary and can be adequately represented as an AR(3) process,
\[\Delta y_t = p_1 \Delta y_{t-1} + p_2 \Delta y_{t-2} + p_3 \Delta y_{t-3} + \varepsilon_t,\]
which can be rewritten as a VAR(1) of the form
\[
\begin{bmatrix}
y_t \\
y_{t-1} \\
y_{t-2} \\
y_{t-3}
\end{bmatrix} = \begin{bmatrix}
1 + p_1 & p_2 - p_1 & p_3 - p_2 & -p_3 \\
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0
\end{bmatrix}
\begin{bmatrix}
y_{t-1} \\
y_{t-2} \\
y_{t-3} \\
y_{t-4}
\end{bmatrix} + \begin{bmatrix}
\varepsilon_t \\
0 \\
0 \\
0
\end{bmatrix}
\]
or
\[Z_t = AZ_{t-1} + \eta_t.\]

Now consider evaluating the infinite sum in equation (A1) by determining
\[(E_t - E_{t-i})y_{t+i}, \quad \forall j.\] These expectations may be evaluated as
\[(E_t - E_{t-i})y_t = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix} \eta_t,\]
\[(E_t - E_{t-i})y_{t+1} = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix} \Lambda \eta_t,\]
\[(E_t - E_{t-i})y_{t+2} = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix} \Lambda^2 \eta_t,\]
and
thus;
\[r \sum \lambda^{i+1} (E_t - E_{t-i})y_{t+i} = \lambda r [1 \ 0 \ 0 \ 0] (I - \lambda A)^{-1} \eta_t,\]
or;
\[(A2) \quad \Delta c_t = r \Phi \eta_t - (1 + r)u_{t-1} + u_t,\]
where;
\[\Phi = \lambda r [1 \ 0 \ 0 \ 0] (I - \lambda A)^{-1}.\]
Equation (A2) represents the evolution of $c_t$, given that the PIH is correct.

With this characterization of $\Delta c_t$, the following joint system can be estimated:

$$
\Delta y_t = p_1 \Delta y_{t-1} + p_2 \Delta y_{t-2} + p_3 \Delta y_{t-3} + \epsilon_t^y
$$

$$
\Delta c_t = k\Phi \eta_t + \beta_0 \Delta y_t + \beta_1 \Delta y_{t-1} + \beta_2 \Delta y_{t-2} + \epsilon_t^c
$$

It should be noted that $\epsilon_t^e$ is not an MA(1) disturbance term. Following Flavin (1981, p. 992) the component of the error term related to transitory consumption is ignored; that is,

$$(1 + r)u_{t-1} + u_t = 0.$$

Now, under the PIH, only surprises in the income series should result in changes in consumption. This result would suggest that upon estimation $\beta_0$, $\beta_1$, and $\beta_2$ should be jointly zero because these are coefficients on variables that are known by the agent at time $t$.

Or, substituting for $\Delta y_t$:

$$
\Delta y_t = p_1 \Delta y_{t-1} + p_2 \Delta y_{t-2} + p_3 \Delta y_{t-3} + \epsilon_t^y
$$

$$
\Delta c_t = (\beta_0 p_1 + \beta_1) \Delta y_{t-1} + (\beta_0 p_2 + \beta_2) \Delta y_{t-2} + \beta_0 p_3 \Delta y_{t-3} + k\Phi \eta_t + \epsilon_t^c
$$

which may be estimated with the system:

$$
\Delta y_t = p_1 \Delta y_{t-1} + p_2 \Delta y_{t-2} + p_3 \Delta y_{t-3} + \epsilon_t^y
$$

(A3) $$
\Delta c_t = \pi_1 \Delta y_{t-1} + \pi_2 \Delta y_{t-2} + \pi_3 \Delta y_{t-3} + k\Phi \eta_t + \epsilon_t^c.
$$

This system is just identified implying that the $\beta_i$'s may obtained using indirect least squares.

In particular:

$$
\beta_0 = \pi_1 / p_3, \quad \beta_1 = \pi_1 - (\pi_2 / p_3)p_1, \quad \beta_2 = \pi_2 - (\pi_3 / p_3)p_2
$$

Under the null hypothesis, that the PIH is correct, $\beta_0=\beta_1=\beta_2=0$, i.e., there is no excess sensitivity. This supposition can be tested with a likelihood ratio test. Estimation of System (A3) shows that:

$$
\Delta y_t = 0.039 \Delta y_{t-1} + 0.115 \Delta y_{t-2} + 0.183 \Delta y_{t-3}
$$

$$
\Delta c_t = 0.05 \Delta y_{t-1} + 0.021 \Delta y_{t-2} + 0.032 \Delta y_{t-3},
$$

implying that $\beta_0 = 0.254$, $\beta_1 = 0.039$, $\beta_2 = 0.0007$.

Under the null hypothesis it was found that $\nu_{ln} = \sum \epsilon' \epsilon = 0.7612$, this compared to the unrestricted value of 0.696. Performing a likelihood ratio test with 3 degrees of freedom
produces a test statistic of 12.5 with a p-value of 0.007. This result suggests rejection of the hypothesis that $\beta_0 = \beta_1 = \beta_2 = 0$, implying that excess sensitivity does exist when disposable income is assumed to be a difference stationary series.
In this appendix, the restrictions associated with the extended version of Model I are formulated. As in Chapter 4, the change in consumption evolves as

(B1) \[ \Delta c_t = q_1 \Delta c_{t-1} + q_2 E_{t-1} \Delta y_p + \omega_t, \]

or, assuming rational expectations and the notion of the expected change in permanent income as developed in Chapter 4,

(B2) \[ \Delta c_t = q_1 \Delta c_{t-1} + q_2 \left[ (0 1 0 -1 0)A(I-pA)^{-1} + (0 1 0 0)A \right] X_{t-1}, \]

where \(A\) and \(X_{t-1}\) will be defined below. The following auxiliary system is used to generate expectations for the quantity \(E_{t-1} \Delta y_p\)

(B3) \[ Z_t = C_1 Z_{t-1} + C_2 Z_{t-2} + C_3 Z_{t-3} + \eta_t, \]

where \(Z_t = \{ \Delta y_{y_n}, \Delta y_{y_k}, \Delta g_t \}\) and \(C_i\)'s are appropriately dimensioned coefficient matrices.

Estimation proceeds by estimating Equations (B2) and (B3) jointly (i.e., by formulating a new VAR system comprised of \(\Delta c_t\) and \(Z_t\)). This system may be represented as

\[
\begin{bmatrix}
\Delta c_t \\
Z_t
\end{bmatrix} = A_1 \begin{bmatrix}
\Delta c_{t-1} \\
Z_{t-1}
\end{bmatrix} + A_2 Z_{t-2} + A_3 Z_{t-3} + \bar{\eta}_t,
\]

where and the \(A_i\)'s are appropriately dimensioned. As discussed in Chapter 4, however, this new VAR system is highly constrained. Determination of these restrictions is most readily obtained by transforming the above joint system into a VAR(1) by allowing

\[
ZZ_t = \{ \Delta c_t, \Delta y_{y_n}, \Delta y_{y_k}, \Delta g_t \}. \]

Such a transformation results in:

\[
\begin{bmatrix}
ZZ_t \\
ZZ_{t-1} \\
ZZ_{t-2}
\end{bmatrix} = \begin{bmatrix}
a_1 & 0_{4x1} & A_2 & 0_{4x1} & A_3 \\
I_4 & 0_{4x4} & 0_{4x4} & 0_{4x4} & 0_{4x4}
\end{bmatrix} \begin{bmatrix}
ZZ_{t-1} \\
ZZ_{t-2} \\
ZZ_{t-3}
\end{bmatrix} + \xi_t,
\]

The column vectors of zeros arise because we do not allow for \(\Delta c_{t-2}, \Delta c_{t-3}\) in the system. This system may be rewritten as:
(B4) \[ X_t = AX_{t-1} + \varepsilon_t, \]
where \( X_t = [ZZ_t, ZZ_{t-1}, ZZ_{t-2}] \) and \( A \) is the corresponding coefficient matrix. Equation (B4)
can then be used for generating the necessary expectations. For example, \( \sum (1/r)^i y_{hk_i} \)
may be expressed as \[ [0 1 0 0 \ldots 0 1 0 0] A (I_{12} - \rho A)^{-1} X_{t-1}. \] Equation (B4) can thus be
rewritten as

\[
\begin{bmatrix}
1 & 0 & \ldots & 0 \\
0 & 1 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & 1
\end{bmatrix} A X_{t-1} = q_1 [I_{12} 1_{10}] X_{t-1} + q_2 [0 1 0 -1 0_{1\times 10}] A (I_{12} - \rho A)^{-1} X_{t-1}
\]

(B5)

\[
q_2 [0 0 1 0 0_{1\times 10}] A X_{t-1}
\]

The left-hand side of Equation (B5) shows that the terms \( a_{11}, a_{12}, a_{13}, \ldots, a_{110} \) are all subject
to restrictions. Manipulation of Equation (B5) shows that a 10 x 10 simultaneous system must
be solved to attain solutions for \( a_{11}, a_{12}, a_{13}, \ldots, a_{110} \). Assuming that the \( \Delta c_{t-1} \) does not appear
in the equations for \( \Delta y_t, \Delta y_{kt}, \) and \( \Delta g_t \), however, the parameter \( a_{11} \) can be identified easily.

This result can be seen directly from Equation (B5) by examining the first element in each row
vector. Doing so shows that \( a_{11} - \rho a_{11}^2 = q_1 - \rho q_1 a_{11} \), which implies that \( a_{11} = q_1 \). By
determining \( a_{11} \) in this manner, the dimension of the simultaneous system is decreased by a
factor of 1. The system to be solved appears as

\[
\begin{bmatrix}
(I - \rho a_{22}) & -\rho a_{32} & -\rho a_{42} & -\rho & 0 & 0 & 0 & 0 & 0 \\
-\rho a_{23} & (I - \rho a_{33}) & -\rho a_{43} & 0 & -\rho & 0 & 0 & 0 & 0 \\
-\rho a_{24} & -\rho a_{34} & (I - \rho a_{44}) & 0 & 0 & -\rho & 0 & 0 & 0 \\
-\rho a_{25} & -\rho a_{35} & -\rho a_{45} & 1 & 0 & 0 & -\rho & 0 & 0 \\
-\rho a_{26} & -\rho a_{36} & -\rho a_{46} & 0 & 1 & 0 & 0 & -\rho & 0 \\
-\rho a_{27} & -\rho a_{37} & -\rho a_{47} & 0 & 0 & 1 & 0 & 0 & -\rho \\
-\rho a_{28} & -\rho a_{38} & -\rho a_{48} & 0 & 0 & 0 & 1 & 0 & 0 \\
-\rho a_{29} & -\rho a_{39} & -\rho a_{49} & 0 & 0 & 0 & 0 & 1 & 0 \\
-\rho a_{210} & -\rho a_{310} & -\rho a_{410} & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
a_{12} \\
a_{13} \\
a_{14} \\
a_{15} \\
a_{16} \\
a_{17} \\
a_{18} \\
a_{19} \\
a_{110}
\end{bmatrix}
\]
Solving this system allows the parameters \( a_{11}, a_{12}, a_{13}, \ldots, a_{110} \) to be expressed in terms of all the other parameters. It is these expressions, say, \( a'_{11}, a'_{12}, a'_{13}, \ldots, a'_{110} \), that are used in estimating Model I.
APPENDIX C

In this appendix the restrictions associated with the extended version of Model 1 are formulated. As discussed in Chapter 4, the change in consumption evolves as

\[(C1) \quad \Delta c_t = q_1 \Delta c_{t-1} + q_2 \varepsilon_{t-1} \Delta y_{pt} + \omega_t,\]

or, assuming rational expectations and the notion of the expected change in permanent income as developed in Chapter 4,

\[(C2) \quad \Delta c_t = q_1 \Delta c_{t-1} + q_2 \begin{bmatrix} 0 & 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix} \Delta \begin{bmatrix} I - \rho A \end{bmatrix}^{-1} + \begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix} \Delta x_{t-1},\]

where \( A \) and \( X_{t-1} \) are defined below. The following auxiliary system is used to generate expectations for the quantity \( \varepsilon_{t-1} \Delta y_{pt} \):

\[(C3) \quad z_t = c_1 z_{t-1} + c_2 z_{t-2} + c_3 z_{t-3} + \eta_t,\]

where \( z_t = \{ \Delta y_{ht}, \Delta y_{st}, \Delta g_{t}, C_1 \} \), and \( c_i \)'s are appropriately dimensioned coefficient matrices. Estimation proceeds by estimating Equations (C2) and (C3) jointly (i.e., by formulating a new VAR system comprised of \( \Delta c_t \) and \( z_t \)). This system may be represented as:

\[
\begin{bmatrix} \Delta c_t \\ z_t \end{bmatrix} = \begin{bmatrix} A_1 & A_2 \\ I & 0 \end{bmatrix} \begin{bmatrix} \Delta c_t \\ z_{t-1} \end{bmatrix} + A_3 z_{t-2} + A_4 z_{t-3} + \eta_t,
\]

where \( A_i \)'s are appropriately dimensioned. As discussed in Chapter 4, however, this new VAR system is highly constrained. Determination of these restrictions is most readily obtained by transforming the above joint system into a VAR(1) by allowing

\[
\begin{bmatrix} \Delta z_{t-1} \\ \Delta z_{t-2} \end{bmatrix} = \begin{bmatrix} A_1 & A_2 & A_3 & A_4 & \Delta c_t \\ 0 & 0 & 0 & 0 & \Delta y_{ht} \\ 0 & 0 & 0 & 0 & \Delta g_{t} \\ 0 & 0 & 0 & 0 & C_1 \end{bmatrix} + \begin{bmatrix} \Delta z_{t-1} \\ \Delta z_{t-2} \end{bmatrix} + \eta_t.
\]

Such a transformation results in

\[
\begin{bmatrix} \Delta c_t \\ \Delta y_{ht} \\ \Delta g_{t} \\ C_1 \\ z_t \\ z_{t-1} \end{bmatrix} = \begin{bmatrix} A_1 & A_2 & A_3 & A_4 \\ I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & I & 0 \\ 0 & 0 & 0 & I \end{bmatrix} \begin{bmatrix} \Delta c_t \\ \Delta y_{ht} \\ \Delta g_{t} \\ C_1 \\ z_t \\ z_{t-1} \end{bmatrix} + \eta_t.
\]

The column vectors of zeros arise because we do not allow for \( \Delta c_{t-2}, \Delta c_{t-3}, C_{t-2}, C_{t-3} \) in the system. This system may be rewritten as
(C4) \( X_t = AX_{t-1} + \varepsilon_t \)

where \( X_t = \{ZZ_t, ZZ_{t-1}, ZZ_{t-2}\} \) and A is the corresponding coefficient matrix. Equation (C4) can then be used for generating the necessary expectations. For example, \( E_{t-1} \sum (1/\tau)^t y_{n+i} \) may be expressed as \( [0 \ 1 \ 0 \ 0 \ 0 \ 0_{1x10}]A(1_{15} - \rho A)^{-1}X_{t-1} \). Equation (C4) can thus be rewritten as

\[
\begin{bmatrix} 1 & 0_{1x14} \end{bmatrix}AX_{t-1} = q_1 \begin{bmatrix} 1 & 0_{1x14} \end{bmatrix}X_{t-1} + q_2 \begin{bmatrix} 0 & 1 & 0 & -1 & 0 & 0_{1x10} \end{bmatrix}A(1_{15} - \rho A)^{-1}X_{t-1}
\]

(C5)

\[
q_2 \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \end{bmatrix}AX_{t-1}
\]

The left-hand side of Equation (C5) shows that the terms \( a_{11}, a_{12}, a_{13}, \ldots, a_{111} \) are all subject to restrictions. Manipulation of Equation (C5) shows that an \((11 \times 11)\) simultaneous system needs to be solved to attain solutions for \( a_{11}, a_{12}, a_{13}, \ldots, a_{111} \). By assuming the \( \Delta c_{i1} \) does not appear in the equations for \( \Delta y_n, \Delta y_{n+i}, \Delta g_r \), however, the parameter \( a_{11} \) can be identified easily. This result can be seen directly from Equation (C5) by examining the first element in each row vector. Doing so shows that \( a_{11} = \frac{-p}{1-p_{11}} = q_1 - \rho q_1 a_{11} \), which implies that \( a_{11} = q_1 \). By determining \( a_{11} \) in this manner the dimension of the simultaneous system is decreased by a factor of 1. The system to be solved appears as

\[
\begin{bmatrix}
(1-p_{a_{22}}) & -p_{a_{23}} & -p_{a_{24}} & 0 & -p & 0 & 0 & 0 & 0 & 0 \\
-p_{a_{23}} & (1-p_{a_{33}}) & -p_{a_{34}} & 0 & 0 & -p & 0 & 0 & 0 & 0 \\
-p_{a_{24}} & -p_{a_{34}} & (1-p_{a_{22}}) & 0 & 0 & 0 & -p & 0 & 0 & 0 \\
-p_{a_{25}} & -p_{a_{35}} & -p_{a_{45}} & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
-p_{a_{26}} & -p_{a_{36}} & -p_{a_{46}} & 0 & 1 & 0 & 0 & -p & 0 & 0 \\
-p_{a_{27}} & -p_{a_{37}} & -p_{a_{47}} & 0 & 0 & 1 & 0 & 0 & -p & 0 \\
-p_{a_{28}} & -p_{a_{38}} & -p_{a_{48}} & 0 & 0 & 0 & 1 & 0 & 0 & -p \\
-p_{a_{29}} & -p_{a_{39}} & -p_{a_{49}} & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
-p_{a_{30}} & -p_{a_{310}} & -p_{a_{410}} & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
-p_{a_{311}} & -p_{a_{311}} & -p_{a_{411}} & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
\end{bmatrix}
\]
\[
\begin{bmatrix}
q_2(a_{22} + a_{32} - a_{42}) - q_2 p \left( \sum_{i=2}^{4} a_{4i} a_{12} + a_{36} \right) \\
q_2(a_{23} + a_{33} - a_{43}) - q_2 p \left( \sum_{i=2}^{3} a_{3i} a_{13} + a_{37} \right) \\
q_2(a_{24} + a_{34} - a_{44}) - q_2 p \left( \sum_{i=2}^{3} a_{3i} a_{14} + a_{38} \right) \\
q_2(a_{25} + a_{35} - a_{45}) - q_2 p \left( \sum_{i=2}^{3} a_{3i} a_{15} \right) \\
q_2(a_{26} + a_{36} - a_{46}) - q_2 p \left( \sum_{i=2}^{3} a_{3i} a_{16} + a_{39} \right) \\
q_2(a_{27} + a_{37} - a_{47}) - q_2 p \left( \sum_{i=2}^{3} a_{3i} a_{17} + a_{310} \right) \\
q_2(a_{28} + a_{38} - a_{48}) - q_2 p \left( \sum_{i=2}^{3} a_{3i} a_{18} + a_{311} \right) \\
q_2(a_{29} + a_{39} - a_{49}) - q_2 p \left( \sum_{i=2}^{3} a_{3i} a_{19} \right) \\
q_2(a_{310} + a_{310} - a_{410}) - q_2 p \left( \sum_{i=2}^{3} a_{3i} a_{110} \right) \\
q_2(a_{311} + a_{311} - a_{411}) - q_2 p \left( \sum_{i=2}^{3} a_{3i} a_{111} \right)
\end{bmatrix}
\]

Solving this system allows the parameters \(a_{11}, a_{12}, a_{13}, \ldots, a_{111}\) to be expressed in terms of all the other parameters. It is these expressions, say, \(a_{11}^*, a_{12}^*, a_{13}^*, \ldots, a_{111}^*\), that are used in estimating the cointegration version of Model I.
APPENDIX D

We now explicitly consider the idea of convergence in the simple model

(D1) \( y_t = gE_{t-1}y_t + x_t + \varepsilon_t \).

Assuming rational expectations, this equation can be solved as

\[
y_t = \frac{1}{1-g} x_t + \varepsilon_t,
\]

since \( E_{t-1}y_t = \frac{1}{1-g} x_t \),

which we refer to as the REE. OLS estimation of \( y_t \) on \( x_t \) over the complete sample will provide an estimate of this REE, \( \frac{1}{1-g} \).

Suppose now that we specify a learning rule to approximate the unknown expectations. Such a rule may be expressed as

(D2) \( y_t = a_t x_t + \eta_t \).

which may be estimated using recursive least squares as laid forth in system (5.5) of Chapter 5. To determine convergence, substitute the expectation of Equation (D2) into Equation (D1), which results in

(D3) \( y_t = ga_t x_t + x_t + \varepsilon_t \).

Consider what happens if coefficient estimates converge as the sample increases; that is, \( a_t = a_{t-1} \), which implies that as \( t \) increases \( a_t \) becomes time-invariant. This condition allows Equation (D3) to be rewritten as

(D4) \( y_t = (ga + 1)x_t + \varepsilon_t \).

Convergence as discussed by Marcet and Sargent (1989) proceeds by studying the differential equation formed by the time-invariant coefficient in Equation (D4), \( T(a) = (ga + 1) \), and the time-invariant coefficient from Equation (D2), \( a \), and studying whether a
rest point exists. [For more discussion of this procedures, along with proofs, see Marcet and Sargent (1989, pp. 339-347)]. Sargent (1993) suggests the use of a smaller differential equation versus the more complex system that is noted as Equation (5), in Marcet and Sargent (1989).

The differential equation of interest is thus
\[ \frac{d}{dt} a = T(a) - a = ga + 1 - a = (g - 1)a + 1. \]

Standard solution of this differential equation shows that
\[ a = Ae^{(g-1)t} + \frac{1}{1-g}, \]

where A is some constant. To ensure convergence, it is noted that \( g < 1 \), which implies that \( \frac{1}{1-g} > 1 \). Further, as \( t \to \infty \), \( a \to \frac{1}{1-g} \), which is what would be expected if the learning rule converges to the REE. Thus, we have shown that when \( a_t = a_{t-1} \), rule (D2) will converge to the REE, given that \( g < 1 \). One important point in this analysis is that the convergence theorem as presented by Marcet and Sargent (1989) depends critically upon \( a_t \) approaching some stable level. This result offers an explanation as to why estimation with the more general model considered in Chapter 5 may not lead to the attainment of the REE. Convergence in this case is only possible if the variance associated with a given coefficient is close enough to zero.
ACKNOWLEDGMENTS

In completing my Ph.D. several acknowledgments must be made. First, I would like to thank my family who never ceased to ask the question "So, when are you going to graduate?" Special thanks must also be conveyed to my wife Jori for support and encouragement. Last, I would like to thank my committee members.