Author:
Fernandez-Corugedo, Emilio

Title:
Essays in consumption.

Essays in Consumption

By

Emilio Fernandez-Corugedo

"A dissertation submitted to the University of Bristol in accordance with the requirements of the degree of Ph.D. in the Faculty of Social Sciencies"

Department of Economics, February 2000

55,782

Abstract: Consumption research has received considerable attention by economists. Hall (1978) proposed that under certain assumptions a rational agent's change in consumption could not be predicted. Hall's aim was to formulate a simple empirical framework that could test the basic premise that agents attempt to smooth their lifetime consumption. Hall found evidence that consumers indeed attempt to smooth their consumption through time.

Subsequent research in the 1980s and 1990s has found that Hall's specification does not hold for aggregate data: consumption is not unpredictable because it reacts too strongly to current labour income (the excess sensitivity phenomenon) and too weakly to permanent income (the excess smoothness phenomenon). Economists have since attempted to explain why these two results appear to be robust to different periods and across countries; however, no common consensus has been reached to be able to determine which theoretical explanation best explains consumption behaviour for aggregate data.

The aim of this thesis is to understand the extent to which consumers attempt to smooth their consumption, what factors prevent this smoothing and whether behavioural changes occur with time. The approach used is both theoretical and empirical. We examine whether a number of studies that claim to explain the failure of Hall's model or whether a mixture of them, is able to explain consumption behaviour more accurately at the aggregate level. For these purposes we try to integrate as many of the concepts of imperfect information, partial adjustment, excess sensitivity, finite lifetimes, habit formation, and adjustment costs into a general equation for consumption that can be estimated using aggregate time series data. We construct two new data sets for the US and the UK which take on board the procedures suggested by Blinder and Deaton (1985) and Attfield, Demery and Duck (1990) to test all the formulated equations.

Acknowlegdments

I am indebted to my Ph.D. supervisors Professor C.L.F. Attfield and N.W. Duck and to my advisor D. Demery for their tremendous support and guidance these last three years. They have challenged me and kept me interested throughout.

Special thanks also goes to the members of staff and Ph.D. students at the University of Bristol and the Bank of England (too many to name) that have given me tremendous feedback and support during the last three years.

My friends in Bristol, London and Spain have managed to keep me sane (and insane) and have helped both motivate and distract me in the most enjoyable manner. Most importantly my love and thanks goes to my Mum (Christianne), Dad (Emilio), Sister (Clara) and Brothers (Edu and Santiago).

AUTHOR'S DECLARATION

I declare that the work in this dissertation was carried out in accordance with the Regulations of the University of Bristol. The work is original except where indicated by special reference in the text and no part of the dissertation has been submitted for any other degree.

Any views expressed in the dissertation are those of the author and in no way represent those of the University of Bristol.

The dissertation has not been presented to any other University either in the United Kingdom or overseas.

SIGNED: [signature] DATE: 21/12/2000

Contents

[1]When we talk about the failure of REPI we understand that failure to be that one of the Hall (1978) and Flavin's (1981) models of consumption.

[2] Diagnostics are reported in table 7.1.

List of Tables

Chapter 1

Introduction

1.1 Motivation

Modern economists have examined consumption, both at the micro and aggregate level, as the result of a problem whereby agents allocate their lifetime wealth to maximize their lifetime welfare. The problem is no different now to what it was in the 1950s, the only difference being the assumption of rational expectations as the expectational mechanism driving agents' forecasts. This in turn, has led to a more rigorous treatment of uncertainty.

The Rational Expectations Permanent Income Hypothesis. Using the rational expectations framework at a time when large macroeconomic models were being used to forecast economic behaviour, Hall (1978) suggested that under certain conditions, innovations in consumption ought to be unpredictable. According to Hall, changes in consumption represent the new information that becomes available to a consumer. Since this new information is unpredictable by nature and is immediately processed by the consumer, consumption innovations are unpredictable. This prediction turned out to be controversial and led to an unprecedented interest in consumption research which has for the most part refuted the random walk prediction for consumption in levels. The random walk prediction has become known as the rational expectations permanent income hypothesis (REPI).

Merits and Motivations. To dismiss REPI on the evidence obtained from aggregate data is perhaps unfair; one should consider its merits before making such conclusions. The fact that economists continue to seek an explanation for consumption as the result of an intertemporal maximization problem where agents are assumed to be rational still makes Hall's framework an appealing one and thus, it would be unfair to totally reject REPI. Yet in its sophistication and appeal lies its reported failure: strong assumptions have to be made to obtain an expression for consumption in levels from the (first order condition) Euler equation which equates expected marginal utilities through time. One cannot underestimate this, economists are still trying to understand what happens to consumption behaviour when we move away from the assumptions made by Hall (and Flavin (1981)). The work of Zeldes (1989), Caballero (1990), Kimball (1990), Deaton (1991), Carroll (1992, 1997a) validates this point; it is very difficult to solve the first order condition to obtain an expression for consumption. This does not mean that we cannot test the Euler equation; Attanasio (1998) (pp. 20-22) discusses this point at length and concludes: 'Even if it is not possible to obtain a closed form solution for consumption, it is possible to consider equilibrium relationships that can be used to estimate structural parameters. While these [...] are not sufficient to answer many important policy questions, they constitute a basic ingredient of any answer.' The principal weaknesses associated with estimating equilibrium relationships are that instrumental variables are often required (Hansen and Singleton (1982, 1983)) and that the tests used to determine the truth of the behavioural relationship tend to be orthogonality tests which are not very powerful.

A total failure? Theoretical and applied economists have tried to explain why consumption does not follow a random walk in aggregate data. Empirical evidence points to the fact that consumption reacts too strongly to current income and too little to permanent income; those two findings commonly referred to in the literature as the excess sensitivity and excess smoothness puzzles of consumption respectively. Perhaps the gloomiest explanation given for the failure of REPI comes from Attanasio: based on a representative agent framework, aggregation issues make it very difficult to be able to predict or explain the behaviour of consumption accurately. Whilst suggestions of this nature must be acknowledged, if we were to adhere to this view we would be left with nothing to say about most of macroeconomics let alone consumption.

Our aims and work. In this thesis, we continue to take REPI as the theoretical foundation for explaining consumption behaviour at the aggregate level. We do this because we wish to know if agents attempt to smooth their consumption through their lifetimes. Despite the fact that precautionary saving behaviour is neglected when we assume linear marginal utility, we use quadratic utility as the main building block in most of our analysis. We do this because we seek closed form solutions for consumption which can be estimated and tested using time series techniques. This leads us to the two main objectives of our research: to quantify the failure of the random walk hypothesis and explain what factors or behaviour can account for this failure. We relax some of the assumptions made by Hall and Flavin and pay particular attention to consumption dynamics to understand what they can tell us about consumption behaviour, specially consumption smoothing.

1.2 Structure of the Thesis and Overview

The thesis is structured in the following manner; in the next chapter (chapter 2) we review the most important developments in the consumption literature and identify the main theoretical and empirical explanations given for the failure of the random walk hypothesis. The most significant empirical results for the US and UK economies have been obtained using two data sets which end in the mid 1980s. Research on these data sets has demonstrated that consumption suffers from excess sensitivity and smoothness, and it is now acknowledged that all consumption research must now explain these stylized facts.

In chapter 3 we take this point. We construct two new data sets for the UK and the US which start in the 1950s and which include observations on consumption and other relevant variables up to 1996. These data sets are constructed following the recommendations made by Blinder and Deaton (1985) for the US and Attfield, Demery and Duck (1990) for the UK. The two resulting data sets are interesting for a number of reasons:

1. Previous empirical research ends in the 1980s and we are interested in understanding what consumption behaviour in the late 1980s and early 1990s can tell us about the permanent income hypothesis and specially whether excess sensitivity and smoothness are still present;

2. Financial liberalization in the 1980s and 1990s might have reduced constraints on borrowing;

3. There has been an increased participation of consumers in stock markets;

4. There have been changes to the housing market which have enabled consumers to borrow more freely.

We use these data sets to test whether consumption innovations can be approximated as white noise errors. We use the techniques suggested by Campbell (1987), Campbell and Deaton (1989) and Flavin (1993) to test for excess sensitivity and smoothness in consumption. These techniques involve the estimation and imposition of appropriate restrictions to a vector autoregression for savings and the innovation in labour income. To impose these tests, all variables in levels must be $I(1)$ apart from savings (which will be $I(0)$ if disposable income and consumption are co-integrated). We impose a number of tests to test for stationarity, including the variant of the Dickey-Fuller test that examines structural breaks suggested by Banerjee et al. (1992). We find that all variables are $I(1)$ and that disposable income and consumption are co-integrated. We compare the results from the new data sets against previously reported results in the literature.

From these tests, we find that consumption is still too sensitive to current income and too smooth compared to permanent income. Thus we conclude that the random walk prediction does not hold for the data sets that we have used. There is, however, some evidence that consumers attempt to smooth their consumption, thus implying that a total departure from Hall's framework is not necessarily the way forward.

We take three 'concepts' that have been used to explain the failure of REPI as the building block of our fourth chapter: the excess sensitivity hypothesis of Flavin (1993), the partial adjustment to news of Attfield, Demery and Duck (1992) and the notion of limited (aggregate) information of Goodfriend (1992) and Pischke (1995). The aim of chapter 4 is to combine these concepts to develop consumption specifications that can be as general as possible and which can be estimated using time series data. This is done for two reasons: first, to obtain equations with rich dynamics which will enhance our understanding of consumption behaviour and, second to impose appropriate restrictions on these general specifications that will enable us to discriminate between the original models to understand whether excess sensitivity, partial adjustment or incomplete information can explain the failures of the random walk prediction or whether a hybrid combination should be preferred.

Chapter 5 is divided into two sections. In the first section, we undertake the task of determining which of the aforementioned concepts can best explain consumption behaviour from the data sets constructed for chapter 3. We explain what econometric techniques should be implemented to estimate all the equations and we discuss what types of restrictions are appropriate to discriminate between models. We find that imperfect information is an important characteristic of both US and UK consumers and that adjustment to new information is slower than previously thought and may even take forever. In the second section of chapter 5, we use the same data sets as before but consider the shorter period that begins in the first quarter of 1973. We re-estimate all the equations from the first part and we examine the same tests that were reported in chapter 3. We do this because it has been argued that the growth rates of many macroeconomic variables have been substantially lower than before from 1973 onwards. The shorter period does not change the results which we found earlier in any significant way although the results are kinder to the smoothing behaviour of rational agents.

In chapter 6 we deviate from the assumption of infinitely lived agents. Clarida (1991) has suggested that a model based on the same assumptions as Hall, but which has finitely lived agents that have to consider their retirement, is able to explain the phenomena of excess sensitivity and smoothness for a simple labour income process. The principal shortcoming of this paper is that for more complex and persistent labour income processes consumption has signs of excess smoothness. Since we reported in the previous chapter that imperfect information appears to be an important characteristic of consumption behaviour, and since imperfect information is capable of explaining excess smoothness, we introduce this notion into the model of Clarida. We find that excess smoothness is greatly reduced even for more persistent labour income processes.

In chapter 7 we turn to the issues of constraints to borrowing and habit formation. The chapter is divided into two sections. In the first section we return to the case where individuals are infinitely lived but we modify the Hall framework to make it difficult or unpleasant for consumers to have negative assets. We also introduce a degree of habit formation as we assume that consumers cannot or do not constantly monitor their asset levels and tend to be slow to adjust them. The solution to this problem is an expression for consumption that is very similar to Flavin's excess sensitivity hypothesis specification, although the level of consumption is less in our framework than in Flavin's. In the second section we combine this model with the models of imperfect information and partial adjustment to obtain a more general consumption specification that can be compared with those from chapter 4. We find that the equation that introduces imperfect information to the equation developed in the first section of this chapter is able to capture some of the characteristics of UK and US data and it does not reject a number of tests.

In our last chapter, we summarize the objectives and results of the previous chapters and suggest future topics of research on the consumption function and in particular for the rational expectations permanent income hypothesis.

Chapter 2

Review of the Consumption Literature

2.1 Early Literature

2.1.1 The Absolute Income Hypothesis (AIH)

The modern macroeconomic research[1] on the consumption function began after the publication of Keynes' (1936) principle that consumption was a stable, although not necessarily a linear, function of disposable income[2]. The early empirical research in particular focused upon the linear relationship

$$c_t = \alpha + \beta y_t \tag{2.1}$$

where c_t and y_t denote the real values of personal consumption and disposable income respectively at time t. The coefficient β known as the marginal propensity to consume (mpc), was expected to be constant and close to one. The coefficient α measures the autonomous component of consumption and was assumed to be small but positive. Keynes argued that the average

[1]The analytical foundations of the modern theory of aggregate consumer behaviour can be traced back to Ramsey (1928) and Fischer (1930).

[2]'Men are disposed, as a rule and on the average, to increase their consumption as their income increases but not as much as the increase in their income' [Keynes, 1936, pp. 96]

propensity to consume (apc), $\frac{c}{y}$, would exceed the marginal propensity to consume, so that the income elasticity of consumption defined as $\frac{mpc}{apc}$, would be less than unity although it would approach unity as income increased. Hence in the long-run, in the face of income growth, one would expect the income elasticity to be unity.

The actual linear form of equation (2.1) is not without qualifications and must not be interpreted literally. In his own work, Keynes acknowledged that unexpected changes in capital values, substantial changes in the rate of interest as well as changes in the distribution of income could have significant influences upon the mpc; he dismissed however the influence of any other objective factors. Keynes also added that as a rule, the proportion of income saved tends to increase with income but he did not consider that remark a fundamental psychological law. Finally, Keynes recognised that because of habit persistence and slow-adjustment the long-run mpc was likely to be larger than the short-run propensity.

The empirical work that followed Keynes' publication provided mixed results but proved to be discouraging overall. Whilst the income term appeared to account for most of the variability in consumption and the apc was larger than the mpc, the actual stability of the consumption function was questioned. Early empirical studies raised the possibility that equation (2.1) could not explain the true behaviour of consumption: i) the presence of a deterministic trend in α could not be ruled out, ii) the apc did not contain a significant trend, iii) α had a tendency to shift upwards in time, and iv) estimates of β were lower than predicted by the theory. A number of post-war studies also pointed out that the absolute income hypothesis could not explain the commonly observed fact that the apc had remained constant in the US since the 1870s when cross-section data at various points in time indicated that the mpc declined as incomes rose.

In view of this evidence, economists soon attempted to explain the inadequacy of the absolute income hypothesis represented in its simplest format by equation (2.1) to explain the behaviour of consumption. A number of theories were quickly developed. Duesenberry's (1949) Relative Income Hypothesis, Brown's (1952) Habit Persistence Hypothesis, Modigliani's (1980) Life Cycle Hypothesis, Friedman's (1957) Permanent Income Hypothesis and Ball and Drake's (1964) Wealth Hypothesis are the most important early studies that attempted to give an explanation for the failure of equation (2.1).

2.1.2 The Relative Income Hypothesis (RIH)

Duesenberry's work originated as a response to the observation that the cross-section household saving ratio had been declining with income whilst the aggregate personal saving ratio had remained constant through time. The Relative Income Hypothesis was based on two basic premises: a) that at any point in time the propensity to save of an individual can be regarded as a rising function of her percentile position in the income distribution and b) that the aggregate savings ratio is independent of the absolute level of aggregate income, although it may depend on the distribution of income. More specifically, it was presumed that consumers had no reason to follow the same cyclical pattern as income. The explanation for this was based on the fact that consumers do not generally choose to reduce the amounts that they consume; while consumers may be happy to increase their expenditure when income is increasing at the beginning of a cycle, they are rather unwilling to experience a severe reduction in consumption when income starts to fall. Hence, Duesenberry suggested that at any point in time, the past peak income may be considered as a better approximation for the amount of autonomous consumption. Equation (2.1) was duly transformed into

$$c_t = \beta_0 y_t^0 + \beta_1 y_t \tag{2.2}$$

where the constant intercept was replaced with one that depended on past peak income, y_t^0. This equation appeared to be compatible with previous empirical results as it allowed for: 1) a smaller mpc in β_1, 2) a changing intercept in the short-run when incomes fluctuate cyclically, 3) an intercept of zero and a larger mpc ($= \beta_1 + \frac{\beta_0}{(1+g)}$) in the long-run when income grows at a constant rate, g.

A large body of evidence was amassed in support of the Relative Income Hypothesis in the forties and fifties. Duesenberry (1949), Modigliani (1949) and Davis (1952) showed that functions based on RIH performed at least as well as various forms of AIH. Some authors even provided support for equation (2.2) using micro data. The choice of past peak income - an arbitrary empirical specification that had no convincing behavioural explanation for its inclusion - together with the lack of forward-looking behaviour implied on agents by the hypothesis, meant that RIH was soon forgotten.

2.1.3 The Habit Persistence Hypothesis (HPH)

Brown suggested that people's habits played an important part in explaining consumer behaviour. He suggested that the past pattern of consumption is likely to show the influence of habits and that these habits would dictate the dynamic responses of consumers to changes in economic variables. Brown ruled out the discontinuity in adjustment implied by RIH on the grounds that consumers are reluctant to make instantaneous adjustments given the persistent nature of their habits which operate independently of fluctuations in income. Brown argued that whilst consumption may be related to current disposable income, the adjustment process implies a dynamic relationship and the regression equation which captures the short-run behaviour of consumption ought to include past consumption as an additional explanatory variable. Brown used a cost minimising approach to obtain a partial adjustment model

$$c_t = c_{t-1} + \lambda \left(c_t^* - c_{t-1}\right)$$

where c^* is the desired level of consumption and λ is related to the degree of persistence of habits. If c^* is assumed to be proportional to current income the last equation can be expressed as

$$c_t = \beta_0 c_{t-1} + \beta_1 y_t \tag{2.3}$$

From equation (2.3) we can see that the previous stylised findings of the early post-war studies are satisfied; there is a smaller mpc (β_1) and a shift in the intercept in the short-run when income fluctuates cyclically. It also collapses to a proportional relationship between c and y with mpc $= \frac{\frac{\beta}{1-\beta}}{1+g}$, in the long-run when income and consumption both grow at a constant rate g.

As with RIH, HPH performed well empirically but its theoretical foundation was not entirely satisfactory as it relied too much on the myopic backward-looking behaviour of consumers. This argument was exposed after the publication of two important theories that relied heavily on the presumption that individuals were forward-looking and attempted to maximise their lifetime utility: the permanent income hypothesis and the life cycle hypothesis.

2.1.4 The Permanent Income Hypothesis (PIH)

The main departure from the Keynesian consumption function occurred with the development of Friedman's permanent income hypothesis. Friedman's work was - like Modigliani's - inspired by utility maximisation: consumer preferences and the constraints imposed by the resources available to consumers, were the main determinants of consumption behaviour for both the PIH and the LCH. PIH focused on distinguishing between consumption[3] and current expenditure on the one hand and income[4] and current receipts on the other. The idea that underlined this theory was that the consumer was thought to plan its expenditures not on the basis of the income received during the current period but rather on the basis of the income expected during its lifetime. The consumer therefore plans the amount of expenditure to be undertaken on the grounds of a long run view of the resources that will be available to him or her.

Friedman postulated that the amount of income a consumer received each time period, y, could be divided into two components: a permanent component (y^p) and a transitory component (y^t). The permanent component was interpreted '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 economic activity 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.' [pp. 21]. The transitory component was interpreted 'as reflecting all "other" factors, factors that are likely to be treated by the unit affected as "accidental" or "chance" occurrences, though they may, from another point of view, be the predictable effect of specificable forces, for example, cyclical fluctuations in economic activity.' [pp. 21-2] Friedman argued that some of the factors that give rise to the transitory component of income were specific to particular consumers (illnesses, bad harvests for farmers, etc.) but that for any considerable group of consumers the transitory components tend to

[3] '...the value of the services that it is planned to consume during the period in question, which, under conditions of certainty, would also equal the value of the services actually consumed. (...) [Consumption] differs from the value of services it is planned to consume on two counts: first, because of additions to or subtractions from the stock of consumer goods, second, because of divergencies between plans and their realizations.' [1957, pp. 11]

[4] Friedman used the Hicksian measure of income; 'On a theoretical level, income is generally defined as the amount a consumer unit could consume (or believes that it could) while maintaing its wealth constant.' [1957, pp. 10]

average out, so that the mean of the transitory component is expected to be equal to zero[5]. In other words, 'the mean measured of the group would equal the mean permanent component' [pp. 22]

A similar explanation was given for the consumption variable. Consumption expenditures were made up of the sum of a permanent component (c^p) and a transitory component (c^t). The permanent component was defined as the amount that a consumer had planned to consume during a single period to maximise lifetime utility. In a world of certainty the amount of total consumption would be equal to the amount of permanent consumption. The transitory component was again interpreted as reflecting all 'other' factors; 'some of the factors producing transitory components of consumption are specific to particular consumer units, such as unusual sickness, a special favorable opportunity to purchase, and the like; others affect groups of consumer units in the same way, such as an unusually cold spell, a bountiful harvest, and the like. The effects of the former tend to average out; the effects of the latter produce positive or negative mean transitory components for groups of consumer units.' [pp. 22-3]

In its 'most general form' PIH is given by the following three equations:

$$\begin{aligned} c^p &= k(i,\ w,\ u)\, y^p \qquad (2.4) \\ y &= y^p + y^t \\ c &= c^p + c^t \end{aligned}$$

where as before c denotes consumption and y income. Letters without a superscript denote current values, letters with a p superscript refer to permanent values and with a t superscript refer to as transitory values. i is the rate of interest at which the consumer can borrow or lend, w is the ratio of wealth to income and u refers to the taste preferences that consumers have.

The first equation in (2.4) defines the relationship between permanent consumption and permanent income; the ratio between the two variables - the mpc out of permanent income - is independent of the size of permanent income but it does depend on other variables: i, w and u. This means that permanent consumption has a constant marginal propensity to consume with respect to permanent income, but at the same time, that propensity to consume

[5] Friedman himself argues that the mean of the transitory component need not be equal to zero.

is allowed to deviate when any of the ceteris paribus assumptions are breached. The last two equations provide a means of linking actual measured variables $(c,\ y)$ to their 'relevant' permanent components.

The most popular version of PIH[6] was based on the following two relationships

$$\begin{aligned} c_t &= \beta y_t^p + c_t^t \\ y_t^p - y_{t-1}^p &= r\left(y_t - y_t^p\right) \end{aligned}$$

The first relationship states that consumption is made up of a planned component that is proportional to permanent income and a transitory component. The second relationship, which came from Friedman's assumption of adaptive expectations, reflects the assumption that any deviation between current receipts and permanent income is capitalised into wealth and only its annuity is added to past permanent income, hence r is a measure of the real rate of interest. Solving the last relationship, one obtains

$$y_t^p = r\rho\left(y_t + \rho y_{t-1} + \rho^2 y_{t-2} + \cdots\right)$$

where $\rho = (1+r)^{-1}$. Therefore, permanent income may be approximated by a geometrically declining weighted average of current and past actual incomes. To proceed empirically, the researcher would only need to substitute this last result into the first relationship of equation (2.4). Hence, the marginal propensity to consume under PIH is equal to $r\rho\beta$ which is smaller than the mpc predicted by AIH.

2.1.5 The Life Cycle Hypothesis (LCH)

Modigliani's LCH also considered forward looking individuals. The theory was developed to consider the life cycle evolution of income and the consumption needs of households. As with

[6] The specification that Friedman himself employed was based on his assumption that 'the transitory components of income and consumption are uncorrelated with one another and with the corresponding permanent components, or

$$\rho_{y^t y^p} = \rho_{c^t c^p} = \rho_{y^t c^t} = 0$$

where ρ stands for the correlation coefficient between the variables designated by the subscripts.' [pp. 26]

PIH, consumers formulate a consumption plan for future periods in order to attain the maximum amount of lifetime utility. The principal difference between the permanent income hypothesis and the life cycle hypothesis is that the latter recognised 'the finite life of households, (so that the LCH) could focus on those systematic variations in income and in "needs" which occur over the life cycle, as a result of maturing and retiring, and of changes in family size. In addition the LCH was in a position to take into account bequests and the bequest motive' [Modigliani (1986), pp. 300].

In Modigliani's model the consumer proceeds to maximise utility subject to the resources available to him or her (these being the 'sum of current and discounted future earnings over his lifetime and his current net worth' pp. 56). The consumption plan that arises from this maximisation problem is then a function of resources available, the rate of return on capital and the age of the maximising agent.

To solve the maximising problem at the individual level, Modigliani introduced three assumptions: i) a homogeneous utility function, ii) no bequest motive and iii) perfect capital markets. Modigliani then showed that consumption could be approximated by[7]

$$c_t = \alpha_1 y_t + \alpha_2 y_t^e + \alpha_3 A_{t-1} \tag{2.5}$$

where c represents aggregate consumption, y represents current nonproperty income, y^e is 'expected annual nonproperty income', and A represents net worth. The model yielded a number of implications which were summarised in Modigliani's (1986) Nobel Prize Lecture:

1. The saving rate of a country is entirely independent of its per capita income.

2. Differing national saving rates are consistent with an identical individual life cycle behaviour.

3. Between countries with identical individual behaviour, the aggregate saving rate will be

[7]To obtain (2.5) Modigliani had to make further assumptions: All households in the economy have the same utility functions and use the same discount rate; the age distribution, the age distribution of income, and the age distribution of net worth are constant; expected income is proportional to current income; the allocation of consumption is not affected by changes in the degree of uncertainty regarding expectations about future earnings; the planning horizon of the individual household is the whole of the life-span; the rate of time-preference is constant; the actions of the individual conform to his lifetime plans for consumption.

higher the higher the long-run growth rate of the economy. It will be zero for zero growth.

4. The wealth-income ratio is a decreasing function of the growth rate, thus being largest at zero growth.

5. An economy can accumulate a very substantial stock of wealth relative to income even if no wealth is passed on by bequests.

6. The main parameter that controls the wealth-income ratio and the saving rate for given growth is the prevailing length of retirement.

2.1.6 The Wealth Hypothesis (WH)

The wealth hypothesis was derived as a direct criticism to the PIH and LCH and in particular to the role played by the budget constraint and its associated implications. More specifically, Ball and Drake's (1964) theory starts from the premise that 'individuals on average are "short-sighted" in the face of considerable uncertainty about the future and the large subjective margins of error that are likely to be attached to any expectations that they may have.' [pp. 65]. Ball and Drake suggest that the role of the wealth constraint in the LCH and PIH be modified from a means of smoothing consumption to attain maximum lifetime utility, to a function that is able to depict a precautionary motive that leads to an accumulation of assets. This modification of the wealth constraint stems from the observation that 'if an individual adjusts his future consumption to his initial asset holdings and present value of expected income from human wealth, we are deprived of the notion of an excess or undesired holding of assets. [...] variations in the initial stock of assets will simply result in an alteration in future consumption rather than any explicit decision to readjust asset holdings, which simply amounts (to saying that) there is no unique equilibrium value of the asset stock.' [pp. 66]

The WH does not modify the wealth constraint but introduces wealth in the utility function. This modifies the nature of the maximising problem; the intertemporal aspect of the consumer choice problem is overlooked in favour of a static framework where individuals must allocate their resources between consumption and wealth at each time period[8]. Their main equation is

$$c_t = (1-\sigma)\,y_t + \sigma c_{t-1} \tag{2.6}$$

This equation is of course similar to that of the HPH (2.3) and an approximation to the PIH of Friedman[9]. The difference between this equation and those derived from the HPH and PIH, is that equation (2.6) imposes a restriction on the coefficients, namely that they ought to add to unity. The usual assumption is that the sum of those parameters ought to be less than one since it is assumed that $\frac{(1-\sigma)}{\sigma}$ is the long run mpc which is believed to be less than one. The WH therefore suggested that the long-run mpc is close to one, although Ball and Drake played down the importance of this result because they argued that $\frac{(1-\sigma)}{\sigma}$ is not a close approximation to the long-run mpc.

The implications of this model were summarised by Ball and Drake [pp. 69-75]:

1. The savings ratio is independent of the level of income.
2. The savings ratio is a function of the rate of income growth. This implies that the long period stability in the savings ratio can only be explained by a stable trend rate of income growth. Increases in the savings ratio in booms are explained by a high rate of income growth relative to trend.
3. The savings ratio is also a function of the marginal wealth-consumption coefficient and on the rate of at which income has been growing in the past which is reflected in the saving ratio of the previous period.
4. Individuals whose incomes have been stationary for some time will tend to have income

[8]The problem becomes a static one since the utility function depends only on current consumption and the current stock of wealth.

[9]Friedman's PIH states that

$$c_t = kr\rho \sum_{i=0}^{\infty} \rho^i y_{t-i}$$

This can be simplified to

$$c_t = \alpha y_t + \rho c_{t-1}$$

where $\alpha = k\rho r$.

elasticities of consumption close to unity, while those whose incomes have grown relatively rapidly in the recent past will tend to have income elasticities that are further away from unity.

5. Hahn's paradox of income illusion[10] is not encountered in this model as it is in the LCH and PIH ones. This is because in the WH model the asset effect is transitory as the initial stock of assets does not affect consumption in the limit.

2.2 The Modern Literature on Consumption

2.2.1 Three Determining Factors

By the early 1970s research on consumption had become lethargic. The utility maximising theories discussed in the last section came to dominate the profession's thinking about consumption behaviour, not only as a result of their theoretical desirability, but also as a result of their sound econometric performance. The belief was one of optimism; it appeared that the consumption function had been solved.

In the 1970s, however, three important factors sparked an interest in the consumption literature making it one of the most prolific areas in economic research. The first factor was purely empirical. The theories explained above began to have difficulties in predicting the be-

[10]For Hahn's paradox of income illusion to occur we require that, for a given exogenously level of disposable income, the marginal propensity to consume be different for each of the two components (labour and non-labour income) which make up disposable income. To see how Hahn's paradox actually occurs, take the 'standard' permanent income consumption equation

$$C_t = r\left(\frac{Y_t^H}{r} + W_t\right)$$

with Y^Hdenoting income from human wealth, W wealth and r the rate of return from asset holding, and consider the following scenario (taken from Ball and Drake (pp. 72-4)): assume two individuals with the same preferences follow the PIH or LCH. Starting from a position of zero endowments, allow each individual to receive a constant flow of labour income each period and endow both with an initial stock of assets. If we allow both individuals to receive the same amount of disposable income each time period but in such a way that individual A receives most of its disposable income from its labour income and individual B from its return on assets, then according to the tenents of both the LCH and the PIH, individual A will have a higher level of permanent income and therefore consumption than individual B. This is because the marginal propensity to consume out of the labour income component is different to the marginal propensity to consume out of the return on assets. Income illusion does not occur in Ball and Drake's model because the asset effect in their model is transitory and not permanent (pp. 73). The transitory nature of their asset effect arises from their inclusion of assets in the utility function used to portray precautionary motives for asset holding.

haviour of consumption accurately. Equations like (2.6) above that had been regarded as the best fitting and least troublesome of all macroeconomic equations, began to underpredict consumption and suggested that the previous stable relationship between (current) consumption and (current) income no longer existed. In retrospect, these empirical failures were prompted by the impact of the cyclical components in economic variables. These components, usually approximated by fluctuations in variables, gathered momentum as the underlying economic environment became more volatile in the 1970s. Empirical research soon revealed that the performance of these models could be significantly enhanced with the inclusion of extra arguments that would capture that increased volatility in the cyclical components. For instance, Hendry and von Ungern-Sternberg (1981) introduced liquid assets as proxies for wealth, Deaton (1977) advocated the inclusion of an inflation variable to depict possible price illusion on behalf of consumers. Subsequently, most leading macroeconomic models were modified to include these and other variables in their consumption functions.

The second factor is somewhat related to the first one and arises as a result of the development and understanding of more elaborate econometric (time series) techniques. As Deaton (1992) states:

'It is a sobering undertaking to look back at many of the macroeconomic models of the time, and note the (now) obvious time-series problems: spurious correlations between integrated regressors, high coefficients of determination coupled with low Durbin-Watson statistics, and an almost complete lack of diagnostic testing.' [pp. 79]

For David Hendry and his associates, economic theory provides a first approximation to empirical testing but econometric techniques serve to obtain information from raw data about the behaviour of economic variables[11]. Such a trend originated with Davidson et al.'s (1978) revolutionary paper on consumption. That paper initiated the development of a conventional methodology for empirical modelling that led to the formalisation of the now standard procedures of cointegration analysis, dynamic models of error correction, etc. Theoretical economists have criticised the practice of using econometric techniques for the sole purpose of developing empirical formulations that may perform well empirically but lack the support of any sound the-

[11]Hendry's econometric principles are summarised in pages 29-31 of 'The Demand for M1 in the U.S.A., 1960-1988', Review of Economic Studies (1992).

oretical base. Nonetheless, the further understanding of econometrics has helped theorists in an important manner. Cointegration and error correction analysis (Engle and Granger 1987) have allowed economists to establish a clear distinction between long-run and short-run (dynamic) statistical relationships between economic variables. That distinction has enabled economists to gain considerable insights into the relationship of consumption with variables that are thought to influence it both in the long and the short-runs. For instance, various empirical studies have noted that a stationary, or an equilibrium long-run, relationship between *current consumption and current income* is unlikely to hold since the secular, or trending components of these variables tend to exhibit a significant divergence. In other words, to achieve the statistical stationary condition which describes the long-run behaviour of consumption, one needs to assume that consumption depends on other 'secondary variables' besides income in the steady-state. A number of studies have found personal wealth, relative prices, measures of income or age distribution, etc. to perform this secondary role successfully[12]. These findings do, in a way, reconfirm not only the evidence of the 1970s concerning the divergence between consumption and income but suggest also that a number of variables are significant in explaining the behaviour of consumption. According to cointegration analysis, the additional variables will enter the consumption function in a different capacity for their primary role is to explain the long-run divergence between consumption and income rather than to capture the so called short-run shifts in autonomous consumption. (This result has had a relatively important policy implication since it suggests that when consumption depends on another variable besides income, a change in the income process cannot guarantee a corresponding change in consumption unless the other variable is entirely unaffected by the underlying policy).

The third factor is a pure theoretical one. It evolved from the rational expectations revolution that was prompted by Lucas' (1976) critique concerning structural relationships between variables. This critique was severe; in the face of rational expectations, no such things as structural relationships between variables may exist. The consumption function was, according to Lucas, one of those structural relationships that did not exist. Under the theory of rational expectations, expectations are formed on the basis of all the available information relating to the

[12]For instance, Hendry and von Ungern-Sternberg (1981) introduced liquid assets as proxies for wealth, Borooah and Sharpe (1986) introduced the distribution of income. See section 11 in Muellbauer and Lattimore (1994) pp. 276-89 for some current examples.

true or actual governing behaviour of the variable to be predicted. Agents in the economy only perceive a structural relationship between permanent income and consumption, but the consumption functions developed above also asserted that a structural relation between observed income and permanent income existed so that consumption would eventually be determined by observed current income. Lucas argued that there was no reason to expect a stable relation between current and permanent income of that type because changes elsewhere in the economy could alter the optimal way consumers make inferences about permanent income from observed income. Consumption depends on current and expected future incomes. The relationship between past and expected future incomes cannot be properly treated as an invariant feature of the economic environment and it is likely to change whenever changes in policy or other events cause rational agents to change the way in which past incomes affect forecasts of future incomes. What does not change however, is the structural relationship between consumption and permanent income.

2.2.2 Hall's Random Walk Consumption Function (REPI)

Introduction

Hall (1978) attempted to reconcile the Lucas' critique for the consumption function. He argued that the possible structural relationship for consumption did not emanate from the previously thought relationship between (current) consumption and (current) income but that the structural relation that would be invariant to policy interventions and other shifts elsewhere in the economy is the ordering of intertemporal preferences. In other words, what does not change in the face of expectations is the agent's overall aim to maximise lifetime utility.

Before explaining Hall's paper in detail and its implications, it is worth noting the following statement made by Hall (1989b) himself: 'Hall neither tried to repair the traditional consumption function nor tried to estimate the deep parameters of utility. Rather, he formulated a simple empirical test of the idea that consumers maximise the expected value of lifetime utility subject to an unchanging real interest rate.' [pp. 156]

The foundations of Hall's 'theory' came from the principles of utility maximisation associated with the life-cycle/permanent income hypothesis that were then accepted as the most accurate applications of the theory of the consumer to the problem of dividing consumption between the present and the future. Hall acknowledged that the Lucas critique would be applicable to consumption even for the LCH/PIH if expectations were not forward-looking. The overall tenant of these theories could not however, be subjected to the Lucas's critique; the consumer plans his or her expenditures on the basis of his or her long-run of lifetime income expectations instead than on the basis of income in the current period. The concepts and measurement of expectations and wealth, contrary to previously held beliefs that placed them in a second order, came to play a central role in the permanent income hypothesis. The fact that permanent income and expected lifetime income or wealth are not directly observable has been and will continue to be a major handicap in carrying out empirical work that is consistent with the theory. Friedman adopted the adaptive expectations hypothesis that led permanent income to be approximated as a weighted average of current and past values of measured income. Hall's principal aim was to examine the effects of introducing forward-looking rational expectations and uncertainty to consumption behaviour.

The model

Hall considered a conventional life-cycle/permanent income model under uncertainty where households choose a stochastic consumption plan to maximise the expected value of their time-additive utility function subject to an 'evolution of assets' budget constraint. The problem is to maximise

$$V\left(c_t, c_{t+1}, \ldots, c_{t+T}\right) = E_t \sum_{\tau=0}^{T-t} (1+\delta)^{-\tau} u\left(c_{t+\tau}\right) \tag{2.7}$$

subject to,

$$\sum_{\tau=0}^{T-t} (1+r)^{-\tau} \left(c_{t+\tau} - w_{t+\tau}\right) = A_t \tag{2.8}$$

where E_t denotes the mathematical expectations operator conditional on information available at time t, δ is the rate of subjective time preference, r is the rate of interest which is assumed to be constant over time ($r \geq \delta$), c is consumption, A are assets apart from human capital, T is the length of economic life; $u(\cdot)$ is the one period-period utility function that is assumed strictly concave and w are earnings which are stochastic and the only source of uncertainty in this model.

The consumer chooses consumption each time period, c_t, to maximise expected lifetime utility given all available information at that point. It is assumed that the consumer knows the value of w_t when choosing c_t. No specific assumptions are made about the stochastic properties of w except that the conditional expectation of future earnings given today's information exists. In particular, successive w_t's are not assumed to be independent, nor is w_t required to be stationary in any sense.

Hall demonstrates that the result to the problem above can be obtained by solving the following Euler equation

$$E_t u'(c_{t+1}) = \left[\frac{1+\delta}{1+r}\right] u'(c_t) \tag{2.9}$$

The advantage of this specification is that it eliminates 'the term that represents the marginal utility of wealth and therefore the necessity of explicitly modelling the way in which the distribution of future variables influences consumption choices.' [Attanasio, 1998, pp. 20]

Implications of the Euler Equation

The first thing to note about (2.9) is that it is an equilibrium relationship and not a consumption function[13]. The Euler equation states that the expected utility lost from giving up a unit of consumption - the right-hand side - must be equal to the expected utility gained by consuming the proceeds of the extra saving at any future date. More formally, the Euler equation states that the marginal rate of substitution between current and future consumption, $\frac{u'(c_{t+1})}{u'(c_t)}$, must equal their relative price given by the rate of interest and the rate of time preference, $\frac{(1+\delta)}{(1+r)}$. The consumption plan chosen by the consumer therefore depends on the preferences the individual has - the shape of the agent's utility function, the rate of time preference -, the rate of interest and unanticipated events. If expectations are fulfilled, in the sense that the expected value on the right-hand side is equal to the actual realisation, then only the rate of interest and

[13]This specification has been used extensively (after imposing certain assumptions/restrictions) to estimate structural behaviour parameters and to test some of the implications of the permanent income hypothesis. [Attanasio, pp. 20]

preferences do determine the consumption plan.

For the Euler equation to hold, the rational agent must choose consumption optimally at each time period *given all the available information at the time the decision is being made.* Consider now a reduction in consumption at time t of size dc from the value the individual had chosen to satisfy the Euler equation that would finance an increase in consumption at some future date from the value the individual would have chosen otherwise. If the individual is maximising (2.7) subject to (2.8), a marginal change of this type should not increase lifetime utility. If utility were to increase then the previously thought optimal choice c_t would not be so optimal for it would not be yielding the maximum amount of lifetime utility.

Re-write the Euler Equation as follows (arguments here follow Deaton, 1992, pp. 25-9)

$$E_t\left(\frac{(1+r)\,u'(c_{t+1})}{(1+\delta)\,u'(c_t)}\right)=1$$

The assumptions about concave utility imply that the marginal utility functions are decreasing in the level of consumption. If we ignore the rate of time preference and the expectation operator (and assume that taste factors are not an issue so that the marginal utilities are the same in different periods), consumption will be higher at $t+1$ than in t, if at constant consumption, marginal utility in $t+1$ would be higher than in t. This also implies that consumption will be growing most rapidly between periods where the interest rate or reward for waiting is highest. When we introduce the rate of time preference, consumption will be growing when the interest rate is greater than the rate of time-preference and declining when the interest rate is less than the rate of time preference.

Given the level of uncertainty and the expectations operator, we note that these results cannot be derived in general for the case when interest rates and consumption are stochastic. In special cases, for example when utility functions are quadratic (marginal utility is therefore linear) and the real rate of interest is non-stochastic, the results above hold exactly. Otherwise the concavity or convexity of the marginal utility function plays a crucial role as does the covariance between the interest rate and the marginal utility of money[14]. Note also that these

[14]The implications about the convexity and concavity of the marginal utility function are examined in more detail below.

conclusions on consumption are results about the way in which consumption evolves over an anticipated life-cycle path; over such a trajectory, there are incentives to allocate consumption to where it is cheapest, so that interest rates can have an unambiguous effect on consumption. However, interest rates, like earnings, also exert an influence on the level of the path and unanticipated changes in interest rates, like unanticipated changes in earnings, will move the path up and down. In consequence, the theory provides no general result on the effects of changes in interest rates on current consumption.

Friedman's basic premise that agents will consume a portion of their wealth stock without affecting their overall stock of wealth still applies. It can be shown, by substituting the Euler equation when utility is quadratic into the budget constraint and rearranging, that the individual will consume $\frac{1}{(T-t)}$ of his or her expected lifetime resources when the rate of interest is equal to the rate of time preference:

$$c_t = \frac{\left[A_t + \sum_{\tau=0}^{T-t} \left\{(1+r)^{-\tau} E_t\left(w_{t+\tau}\right) - (\tau+1)\left(\frac{\delta-r}{1+r}\right) c^*\right\}\right]}{T-t} \tag{2.10}$$

where c^* is the bliss level of consumption in the quadratic utility function $u(c_t) = -\frac{1}{2}(c^* - c_t)^2$. The expression in square brackets in (2.10) brackets is now regarded as the standard definition of the approximation of permanent income or the amount of expected lifetime wealth which is the sum of non-human and human wealth. This means that consumption in the face of no unpredictable events will be the same in the lifetime of the individual and will only change when unpredictable events regarding the evolution of earnings occur.

The principal implications of the Euler equation were given by Hall (1978) in a number of corollaries:

Corollary 1. 'No information available in period t apart from the level of consumption c_t helps predict future consumption c_{t+1}, in the sense of affecting the expected value of marginal utility. In particular, income or wealth in periods t or earlier is irrelevant, once c_t is known'. This finding, the most important one in the paper is a straight result; the only determinants of the consumption plan are preferences, relative prices and current lifetime resources not current income. Consumption patterns are shaped by tastes and by life-cycle needs, not by the temporal pattern of life-cycle income.

Corollary 2. 'Marginal utility obeys the regression relation $u'(c_{t+1}) = gu'(c_t) + e_{t+1}$, where $g = \frac{(1+\delta)}{(1+r)}$ and e is a true regression error, so that $E_t e_{t+1} = 0$'.

Corollary 3. 'If the utility function is quadratic $u(c_t) = -\frac{1}{2}(c^* - c_t)^2$ (where c^* is the bliss level of consumption), then consumption obeys the exact regression $c_{t+1} = \beta_0 + gc_t - e_{t+1}$, with $\beta_0 = c^* \frac{(r-\delta)}{(1+r)}$. Again, no variable observed in period t or earlier will have a nonzero coefficient if added to this regression'.

Corollary 5. 'Suppose that the change in marginal utility from one period to the next is small, both because the interest rate is close to the rate of time preference and because the stochastic change is small. Then consumption itself obeys a random walk, apart from trend. Specifically $c_{t+1} = \lambda_t c_t + \frac{e_{t+1}}{u''(c_t)} + higher - order\ terms$'. λ is given by

$$\lambda_t = \left[\frac{1+\delta}{1+r}\right]^{\frac{u' c_t}{c_t u''(c_t)}}$$

and represents the rate of consumption growth. The rate of growth of consumption is not necessarily constant but may vary over time and is a function of the current level of consumption via the marginal utility expression and the first differential of marginal utility. Moreover, because of assumptions about the shape of the indifference curve - concave marginal utility - the rate of growth is greater than one so that in the consumption equation we have a random walk.

Solutions to the Euler Equation and the Plausibility of the Assumptions Used by Hall

Hall argues that the above corollaries lead to the result that the simple relationship

$$c_t = \lambda c_{t-1} + \varepsilon_t \tag{2.11}$$

where ε_t is unpredictable at $t-1$, can approximate closely the stochastic behaviour of consumption under the permanent income hypothesis[15].

[15] As Deaton (1992) states, the equation does not say anything about the variance of ε, and there is no reason to believe that the variance is constant. Hence, strictly speaking, equation (2.9) is not a random-walk.

What are the economic implications of this equation? This equation states that the best forecast about the level of consumption in the next period is today's level of consumption. The discrepancies between the two levels of consumption are brought about by unpredictable events at time t, reflected by the disturbance term. Despite the prediction that the change in consumption is unpredictable, this result does satisfy the rational expectations premise. A rational expectation will use all available information relevant to the behaviour of consumption when the expectation is being formed. In period $t-1$, given the information available at that time, the agent will set consumption at c_{t-1} which as we saw above was equal to his or her estimate of his or her permanent income. The right hand side of equation (2.10) gives that estimate of permanent income. To be a rational decision this decision about c_{t-1} would have taken account of all information regarding the evolution of w and r and the needs of the consumer represented by the utility function and δ available at time $t-1$ and earlier. Since in period $t-1$ the agent has consumed an amount equal to his or her permanent income, his or her stock of wealth $(A_t + H_t)$ in period t will be the same as it was at the beginning of $t-1$ if no information about the future has become available in period t and so in period t the consumer's estimate of his or her permanent income will be unchanged and he or she will set consumption, c_t, at the same level as before, c_{t-1}. Only if new information becomes available between periods $t-1$ and t will consumption change in period t. As new information is unpredictable by definition, it must be the case that consumption differs from lagged consumption only by an unpredictable element[16]. Hence, the disturbance term conveys information about the impact of all new information that becomes available to the consumer in period t about his or her lifetime well-being. All the past/predictable information is reflected in the lagged consumption term. Hall demonstrates that it is possible to derive an expression for that unpredictable element:

Non-human assets evolve according to the expression;

$$A_t = (1+r)\left(A_{t-1} - c_{t-1} + w_{t-1}\right)$$

[16] Attfield, Demery and Duck (1991), pp. 208.

and human wealth evolves according to

$$H_t = (1+r)(H_t - w_{t-1}) + \sum_{\tau=0}^{T-t} (E_t w_{t+\tau} - E_{t-1} w_{t+\tau})$$

so that the behaviour of the total wealth stock is given by the following equation:

$$A_t + H_t = (1+r)(A_{t-1} - c_{t-1} + H_{t-1}) + \eta_t$$

where

$$\eta_t = \sum_{\tau=0}^{T-t} (E_t w_{t+\tau} - E_{t-1} w_{t+\tau})$$

The evolution of total wealth then depends, ceteris paribus, on the relationship between two informational variables, η_t and ε_t. By imposing quadratic utility or certainty equivalence, that relationship is given by:

$$\begin{aligned} \varepsilon_t &= \left[1 + \frac{\lambda}{1+r} + \cdots + \left(\frac{\lambda}{1+r}\right)^{T-t}\right] \eta_t \\ &= \alpha \eta_t \end{aligned} \tag{2.12}$$

This is according to Hall 'the modified annuity value of the increment in wealth. The modification takes account of the consumer's plans to make consumption grow at a proportional rate λ over the rest of his life.' (pp. 975-6)

All the Euler equation results discussed above still apply. The martingale consumption equation is simply a stochastic generalisation of the simplest life-cycle model in which consumption is constant over life with (predictable) variations in income offset by appropriate asset transactions.

The economic implications of this solution to the Euler equation and the Euler equation itself, however, must be taken into perspective for they clearly depend on a number of *important* assumptions[17]:

[17] Here we list the four most important ones. Others are:
v) No habits or adjustment costs,

i) Consumption is the only argument in the consumer's utility function,

ii) Capital markets are perfect so that consumers can borrow/lend without any restrictions at a constant rate as long as the present value of their consumption does not exceed the present value of their human and financial wealth (this means that there are no non-linearities in the budget constraint),

iii) The rate of time preference does not exceed the rate of interest,

iv) Certainty equivalence is assumed by Hall.

Together, the first two assumptions imply that rational agents can substitute between current and future expenditures to achieve the maximum level of lifetime utility without any difficulties. The ability to borrow and lend makes the optimal consumption plan independent of current income under no uncertainty. Current income does affect consumption plans in a rational expectations-permanent income framework with certainty equivalence only through its unpredictability, depicted by the error term in the consumption equation. This explains why consumption plans are independent of the level of current income and only depend on the preferences, the rate of interest faced by the consumer and unforeseeable events. The realism of the second assumption has come under pressure recently as it suggests that consumers do not face any type of liquidity constraints.

The third assumption restrains the impatient nature of consumers from surfacing since if that were not the case and if their future income were known with certainty they would consume more than their current income and go substantially into debt. The assumption enables the consumer to be willing to accumulate wealth in the form of savings. If this assumption were not made, and under either no uncertainty about the income process or under certainty equivalence, it would be possible for the agent to undertake too higher levels of consumption earlier in life.

The fourth assumption helps solve the Euler equation. The principal difficulty associated with solving the Euler equation is almost entirely due to the presence of uncertainty and to the resulting expectation operator. The problem disappears if we can pass the operator through the function, but that is only valid if the marginal utility functions are linear, i.e. there is

---

vi) Non-durable goods assumed only,
vii) No measurement errors or transitory shocks to consumption,
viii) The coincidence of the frequency of consumers' decision making with the observation period of the data,
ix) Infinite lifetimes.

certainty equivalence. It is therefore a very powerful assumption made on the face of Jensen's inequality, which places importance upon the actual shape of the indifference curve[18]. As it will be explained below, the actual shape of the indifference curve is crucial in determining the nature of the result: a convex marginal utility function means that not only is the marginal value of consumption higher when consumption is low, but that the rate at which the marginal valuation rises with shortfalls in consumption should be greater when consumption is low than when it is high. A linear marginal utility function therefore rules out a precautionary motive in consumer's behaviour.

Deviations from Quadratic Utility (Certainty Equivalence) In this section, we look at the work of Kimball (1990) which was based on Pratt (1964), Arrow (1965), Leland (1968) and Rothschild and Stiglitz (1970, 1971). The simple 2 period framework is extended by Carroll and Kimball (1996).

Kimball examines the following two period problem (pp. 59),

$$\max_c u(c) + Ev\left(w - c + \tilde{y}\right)$$

where u is the first period utility function, v is the second period utility function, c is first period consumption, $w = w_0 + \bar{y}$ where w_0 is the consumer's initial assets, $\bar{y}$ is the expectation of second period income, and $\tilde{y}$ is the risky component of second period income, such that $y = \bar{y} + \tilde{y}$. The first order condition is

$$u'(c) = Ev'\left(w - c + \tilde{y}\right)$$

Therefore the risky component of second period income will affect consumption in the first period in as far as it affects marginal utility in the second period thus disrupting the first order condition. Define savings as $s = w - c$ and write the first order condition as,

$$u'(c) = Ev'\left(s + \tilde{y}\right)$$

[18] Jensen's inequality tells us that for convex (concave) marginal utility, the expectation of the function is greater (less) than the function of the expectation.

Following Pratt (1964), Kimball [pp. 59] defines the following concepts,

1. If a quantity ψ^* (called the *compensating precautionary premium*) exists that satisfies $v'(w-c) = Ev'\left(w-c+\tilde{y}+\psi^*\right)$ and thus compensates for the effect of the risk $\tilde{y}$ on second-period expected marginal utility then first-period consumption would be unaltered by the addition of the risk and the compensating precautionary premium, and

2. If a quantity ψ (called the *equivalent precautionary premium*) exists that satisfies $v'(w-c-\psi) = Ev'\left(w-c+\tilde{y}\right)$, then the elimination of the risk $\tilde{y}$ at the cost to the consumer of the certain quantity ψ would leave optimal first-period consumption unchanged.

Kimball demonstrates (Lemma, pp. 57-8) that both ψ and ψ^* are not only 'approximately equal "in the small", but almost all important qualitative results about equivalent risk premia are interchangeable with corresponding results about compensating risk premia, including the result about risk premia "in the large" discussed above. Furthermore, because of the close analogy between risk premia and precautionary premia one can be confident that a result about equivalent precautionary premia will imply a corresponding result about compensating precautionary premia.'

Kimball points out that 'since $u'(c)$ is constant for a fixed value of the decision variable c, it can be ignored in the definition of precautionary premia for this model' [pp. 60]. Armed with these concepts, Kimball then demonstrates that the analogy between the theory of risk aversion and the theory of precautionary saving is particularly simple, [pp. 60]: 'The negative of marginal utility, $-v'$, plays substantially the same role for precautionary saving that the utility function itself plays for risk aversion. For example, concavity of v indicates risk aversion, while concavity of $-v'$ (i.e. $v'''(\cdot) > 0$) indicates a positive precautionary saving motive. As another example, the index of absolute prudence in this model, which represents the strength of the precautionary saving motive, is

$$\eta(w,c) = \eta(s) = -\frac{\left(-v'(s)\right)''}{\left(-v'(s)\right)'} = -\frac{v'''(s)}{v''(s)}$$

Thus we can see that deviations from quadratic utility lead to a precautionary saving motive since the third derivative of the utility function is different from zero. These results only apply for a given level of wealth in this framework; Carroll and Kimball (1996), extend this framework to a multiperiod problem to show that precautionary saving of the nature Kimball investigated can occur at *different* levels of wealth and consumption.

A number of studies have tried to determine whether relaxing quadratic preferences can help us explain consumption decisions. Apart from Caballero (1990) who looked at the *consumption function* when preferences are of the CARA class, most of the research that has relaxed quadratic preferences has examined the Euler equation (2.9). Two early studies which first tested the Euler equation with CRRA preferences were those of Hansen and Singleton (1982) and Skinner (1988). The test used in Hansen and Singleton is simple yet powerful (see Deaton, pp. 66-7); if preferences are CRRA, $u(c_t) = \frac{c_t^{1-\rho}}{1-\rho}$, the Euler equation is given by

$$E_t\left(\frac{1+r_{t+1}}{1+\delta}c_{t+1}^{-\rho}\right) = c_t^{-\rho} \tag{2.13}$$

so that if we define the following quantity z_{t+1} as

$$z_{t+1} = \frac{1+r_{t+1}}{1+\delta}c_{t+1}^{-\rho} - c_t^{-\rho}$$

then according to (rational expectations) theory, any variable which is dated at time t or earlier, say w_{tj}, should be orthogonal to z_{t+1}:

$$\frac{1}{T}\sum_{t=1}^{T} w_{tj} z_{t+1} = 0 \tag{2.14}$$

$j = 1, ...J$, where J denotes the number of potential instruments in the consumer's set. Using data on consumption for non-durables and services and using treasury bill rates and the return on New York Stock Exchange stocks, Hansen and Singleton found that the Euler equation can be rejected when (2.14) is estimated using GMM.

Skinner (1988) made a second-order Taylor-Series approximation to the Euler equation (2.13) to examine the importance of *income risk* on precautionary saving decisions. He found mixed evidence: using data from the Panel Study of Income and Dynamics, Skinner tests for the accuracy of the Taylor-Series approximation which warrants the inclusion of a precautionary savings term and finds that the approximation can explain 94% of the data (pp. 249). However, using data from the Consumer Expenditure Survey of 1972-3 to compare saving rates across occupations, Skinner found that those in traditionally 'riskier occupations, such as sales and the self-employed, saved significantly less than average', pp. 252.

Browning and Lusardi (1996) review most of the empirical evidence on the Euler equation and also on precautionary savings at the *micro* level. Browning and Lusardi point out that the evidence on the validity of the Euler equation using the orthogonality condition discussed above is 'deeply ambiguous' (pp. 1835). An explanation for such ambiguity arises from an important problem in all the studies examining precautionary savings: how to construct an observable and exogenous measure of risk that varies across the population (pp. 1835-6)[19]. Browing and Lusardi give an excellent discussion on the strengths and weaknesses of most of the variables chosen to proxy risk. The evidence on the precautionary savings papers reviewed in that paper is mixed, with some studies finding little or no evidence and others claiming that the precautionary motive can explain a large proportion of wealth holdings by households.

Carroll (1997b) has recently advocated that tests performed on the log-linearised Euler equation have to be interpreted with caution: 'the theory implies that the higher-order terms in the approximation cannot be ignored because they are endogenous' [pp. 8][20]. This endogeneity problem suggests that there is no simple way of testing for the occurrence of precautionary motives in consumption.

2.2.3 The Error Correction Equation of Davidson et al.

The error correction approach to modelling originated from Davidson et al.'s (1978) (DHSY hereafter) seminal paper on UK consumer expenditure. The aim of that paper was to develop a framework that could explain: (i) UK data, (ii) the findings of previous models, (iii) exhibit parameter stability over time and (iv) conformed to steady-state postulates of economic theory.

The further understanding of time series analysis[21] rather than economic theory, guided the development of DHSY's work. Their final consumption specification was based on the following premises:

[19]Even Browning and Lusardi disagree on the evidence (pp. 1835).

[20]See the section on buffer stock/precautionary savings below for an explanation of this endogeneity.

[21]Spurious regressions (Granger and Newbold, 1974), General to Specific Modelling (Hendry and Mizon, 1978).

- "We consider it an essential (if minimal) requirement that any new model should be related to existing "explanations" in a constructive research strategy such that previous models are only supplanted if new proposals account for previously understood results, and also explain new phenomena,

- "To avoid directionless "research" and uninterpretable measurements, a theoretical framework is also essential,

- "An econometric model must account for the properties of the data". [pp. 662]

Three previous UK consumption studies that showed remarkable differences in their results, even though they employed similar data sets, were taken by DHSY as their starting point. In line with their premises, DHSY first standardised (to reflect common data sets, methods of seasonal adjustment and other data transformations and functional forms) and then nested the works of Hendry (1974), Ball et al. (1975) and Wall et al. (1975) into a general model. The general framework postulated by DHSY was concerned primarily with the use of (only) directly observable variables without resorting to 'hypothetical constructs and/or [...] unclearly specified but stringent ceteris paribus assumptions [that do] leave many important decisions in formulating an operational model to ad hoc considerations'. [pp. 662]. The emphasis shifted from one of explaining economic theory to one of explaining the dynamic behaviour of economic variables. DHSY conducted statistical tests to determine which of the three models performed better for the UK economy and they found the preferred specification to be that one of Wall et al.:

$$\Delta_4 c_t = (\alpha) + \beta_1 \Delta_4 y_t - \beta_2 \Delta_1 \Delta_4 y_t + \theta D_t^0 \tag{2.15}$$

where $\Delta_4 c_t = c_t - c_{t-4}$ and D_t^0 is a dummy variable. Whilst (2.15) fits the data well, there is a major reason to feel unhappy with this equation since it can be rejected on the grounds of economic theory rather than statistical diagnostic testing. Sound economic theories relate economic variables to situations where any adjustments of the consumer units to positions of disequilibrium are assumed to have been completed. Prior to DHSY's work, no dynamics were generally specified in theoretical economic models because adjustment dynamics can only occur in situations of disequilibrium. Economic models are formulated to produce sensible interpretations for both situations of *static equilibrium*, where the variables are assumed to be unchanging between periods, and for *stable equilibrium*, where all variables are changing at some constant rate. It is straight forward to show that Wall et al.'s specification does not have both a static equilibrium solution and an informative stable solution[22]. Furthermore, from a theoretical point of view, it seems inappropriate that an equation that attempts to explain short-run behaviour assumes that behaviour to be independent of the disequilibria in the levels of the variables. Theoretically speaking, short-run adjustments of the type explained by (2.15) can only arise if the variables in the system are not in long-run equilibrium for otherwise adjustment towards equilibrium would not be necessary. Hence, if one were to start from the premise that the variables are not in long-run equilibrium and therefore necessitate some form of adjustment, that adjustment ought to be a function of the level of disequilibrium between the variables. No term that depicts the amount of disequilibrium exists in (2.15).

A further problem with the above specification is related to the form of the variables. Differenced variables facilitate the study of short-run behaviour without having to specify trend-dominated long-run components. One therefore loses almost all a priori information from economic theory and all long-run information in the data.

All this evidence suggested that the need of either a new economic theory was required to explain why this specification appears to out-perform long-run stable economic models, or that an account ought to be provided of why this specification appears to out-perform models that are based on economic theory (as Hendry and Ball et al.'s were). The latter path was followed since a simple modification of equations in differences could, according to DHSY, resolve the long-run stable equilibrium and loss of data information problems as well as being able to introduce a disequilibrium term into the framework. It is here that the DHSY approach is new.

[22]For static equilibrium all variables must be constant and unchanging, so that $\Delta_4 c_t = \Delta_4 y_t = \Delta\Delta_4 y_t = 0$ and ignoring the long-run effects of the dummy variable D^0, equation (2.15) becomes:

$$0 = \beta_1 \cdot 0 + \beta_2 \cdot 0$$

which is not a long-run solution of any type. To show that (2.15) has no stable equilibrium solution, assume that both consumption and income are growing at a constant rate of 100g% per annum, so that we have $\Delta_4 c = \Delta_4 y = g$ and $\Delta_1\Delta_4 y = 0$. This implies that the 'long-run version' of (2.15) is $\Delta_4 c = \beta\Delta_4 y$ and unless $\beta = 1$ $\Delta_4 c = g \neq \Delta_4 y = g$ we would not have a stable solution.

A non-stochastic steady-state theory for the consumption function of the form

$$\begin{aligned} C_t &= KY_t \\ c_t &= k + y_t \end{aligned} \tag{2.16}$$

is considered. (Lower case letters denote the logarithm of the variable). It is believed that consumption is a constant proportion of income, and it is further assumed that the income elasticity of consumption is unity. Given that the preferred specification for consumption is in differences and has no long-run steady-state solution in the equation, an investigator will wish to reconcile an equation like (2.16) with the short-run equation (2.15) that fits the data reasonably well. DHSY argue that in the absence of a 'well-articulated theory of the dynamic adjustment [of consumption and income it is convenient to postulate a] general rational lag model of the form

$$\alpha(L)\, c_t = k^* + \beta(L)\, y_t + v_t \tag{2.17}$$

where $\alpha(L)$ and $\beta(L)$ are polynomials in the lag operator of high enough order so that v_t is white noise' [pp. 680]. DHSY consider a first order lag in the lag operator to be a close approximation of the dynamic process that represents consumption, thus,

$$c_t = k^* + \beta_1 y_t + \beta_2 y_{t-1} + \alpha_1 c_{t-1} + v_t \tag{2.18}$$

which is close to the specification above that had no steady-state static nor stable solution (when $\alpha_1 = 1$ and $\beta_1 = -\beta_2 = 1$). To ensure that for all values of the estimated parameters the steady-state solution to (2.18) reproduces (2.16), the following restrictions must be imposed:

$$\beta_2 = -\beta_1 + \gamma \textit{ and } \alpha_1 = 1 - \gamma \tag{2.19}$$

and so the following equation results

$$\Delta c_t = k^* + \beta_1 \Delta y_t + \gamma\,(y_{t-1} - c_{t-1}) + v_t \tag{2.20}$$

This is DHSY's significant contribution. Assuming that equation (2.18) passes the restrictions in (2.19) we see that equation (2.20) is guided by long-run theory. There is no loss of long-run information in the data as (2.20) is a reformulated 'levels equation' and the vital 'initial disequilibrium' effect is provided by $\gamma(y_{t-1} - c_{t-1})$ [pp. 681]. Due to this term, this equation is said to have an error correction mechanism (ECM). This equation does represent a marked theoretical improvement over equation (2.15) as it has both static and stable equilibrium solutions. It is easy to show that the static solution in (2.20) is $\gamma(y_{t-1} - c_{t-1})$ since $\Delta_4 c_t = \Delta_4 y_t = \Delta\Delta_4 y_t = 0$ in steady-state. Now since usually $\gamma \neq 0$, then

$$c_t = y_t$$

is the static equilibrium solution. This solution implies that a long-run unitary elasticity of consumption exists. For the stable solution, if consumption and income are assumed to be growing at the same constant rate, g, equation (2.20) solves as

$$\begin{aligned} g &= \beta_1 g + \gamma(y_{t-1} - c_{t-1}) \\ c_{t-1} &= y_{t-1} + \frac{g(\beta_1 - 1)}{\gamma} \end{aligned}$$

so that the long-run consumption function can be expressed as (2.16) but where

$$K = \exp\left[\frac{g(\beta_1 - 1)}{\gamma}\right]$$

Equation (2.16) still exhibits a unitary elasticity as required from the static solution to the problem. Moreover, since $\gamma < 0$, the long-run APC (=MPC in steady-state growth) given by K in equation (2.16) is a 'decreasing function of the growth rate, g, consonant with inter-country evidence' [pp. 681]. The model implies that even a variable or trending observed APC, will not refute a unit-elasticity model. That the model has a unit elasticity becomes one of the tests that the DHSY model has to pass.

The final specification ran by DHSY included an inflation term. The introduction of the inflation term was prompted by the possibility that such a variable could be causing one of two effects: i) Money illusion (Branson and Klevorick (1972)) where individuals see wage inflation as an increase in their real wages and thereby consume more, ii) Price effect (Deaton, 1977), where consumers see the effects of inflation reflected as higher prices in some the goods they often purchase. Such an increase is viewed as a reduction in their purchasing ability and less is consumed.

DHSY's preferred consumption specification includes the inflation term that depicts the price effect. Whilst the preferred equation seems to fit the UK data extremely well, it is however necessary to be aware about the nature of the result for it appears that the inflation term plays an important role in explaining the behaviour of consumption, and without it the predictive powers of an equation like (2.20) become questionable.

Nonetheless, the lasting contribution of DHSY's paper has been one of taking into account the possible differences that exist between long-run behaviour and short-run dynamic adjustments to positions of disequilibrium. Empirical work, at least in the UK, on consumption has focused more on trying to account for previous models and findings in a systematic explicit way (Hendry and von Ungern-Sternberg (1981) amongst others). As a result of this approach more emphasis is being given to short-term dynamics, or adjustment processes and to the long-run properties of the consumption function.

2.2.4 The Error Correction Equation against the Random Walk Hypothesis

Hall's specification and DHSY's consumption equation have come to dominate all research on the consumption function since 1978.

From a theoretical point of view both models are quite different. The random walk model is a forward-looking, theoretically based approach to forecasting and provides an explanation of the path of consumers' expenditure. On the other hand, the ECM approach is a more backward-looking model that pays attention to short-run adjustment dynamics[23]. Whatever specification is relevant empirically will have different implications not just for economic theory but for policy analysis.

1. The principal implication of REPI is that only unanticipated changes will affect the short-run dynamics of consumption. Anticipated changes in variables will have no effect on

[23] This suggests a breach of the Lucas's critique regarding the specification of consumption models.

consumption when they materialise since they ought to have been taken into account once the news about the change became available. Furthermore, changes in the variables and their effect on consumption will depend on whether those variables are permanent or transitory. The error correction mechanism on the other hand, as specified by DHSY, bases changes in consumption on contemporaneous income and price effects.

2. No variables other than past consumption are relevant for predicting future levels of consumption in the REPI case. The ECM however does not start with a fully developed theoretical scheme, and researchers have the tendency to introduce additional regressors at their discretion to develop a model that fits the data quite well. Once that is achieved, the steady-state implications are examined. Hence, any variable is in theory capable of explaining the behaviour of consumption.

3. The ECM allows for the effects of a disequilibrium variable. The deviations reflected by the disequilibrium term could emanate from a number a reasons; genuine errors made by the consumer, a lack of relevant information, a partial or slow adjustment, etc. The REPI does not introduce such a variable because it assumes that consumers are always in equilibrium and need adjust only when new information becomes available. The adjustment is immediate and any deviations from the optimum values are assumed to be white noise[24].

Whilst theoretically speaking Hall's specification is more desirable than the error correction equation, the random walk equation neglects dynamic responses [Hadjimatheou, 1987, pp. 168-170]. It is easy to show that Hall's model imposes serious restrictions compared to the error correction mechanism model;

$$\Delta c_t = \beta_1 \Delta y_t + \beta_2 \Delta\Delta y_t + \gamma (c - y)_{t-1} + v_t$$

which can be written as

$$\frac{C_t}{C_{t-1}} = \left(\frac{Y_t}{Y_{t-1}}\right)^{\beta_1} \left(\frac{Y_t}{Y_{t-1}}\frac{Y_{t-2}}{Y_{t-1}}\right)^{\beta_2} \left(\frac{C}{Y}\right)^{\gamma}_{t-1} (1 + v_t) \cdot$$

[24]This is obviously the strict version of the rational expectations permanent income hypothesis.

If $\beta_1 = 1$, $\gamma = 0$ and income follows the autoregressive process, $Y_t = \lambda Y_{t-1}$, the above equation becomes $C_t = \lambda C_{t-1} + e_t$, where $e_t = \lambda v_t C_{t-1}$ and the error term is expected to be serially correlated. Alternatively, if $Y_t = \lambda Y_{t-1} + \phi Y_{t-1}$, $\phi \sim (0, \sigma^2)$ and $\beta_2 = \gamma = 0$, $\beta_1 = 1$, the equation is reduced to $C_t = \lambda C_{t-1} + w_t$, where $w = C_{t-1}[\phi(1+v_t) + v_t\lambda]$. If at the steady-state $\frac{Y_t}{Y_{t-1}} = \lambda$, so that $(\frac{Y_t}{Y_{t-1}}\frac{Y_{t-2}}{Y_{t-1}}) = 1$, and if $\frac{C}{Y} = K = constant$, the last equation becomes

$$C_t = \lambda^{\beta_1} K^{\beta_2} C_{t-1}$$

so that the white noise error term of Hall is suppressed. These examples show the restrictive nature of Hall's model compared to the general specification (2.20).

2.2.5 Failure of the Rational Expectations/Permanent Income Hypothesis[25]?

The first tests conducted on REPI introduced lagged variables other than lagged consumption to the martingale equation[26]. According to (2.11) no variable other than lagged consumption should be able to predict current consumption so that other lagged variables must have statistically insignificant coefficients when introduced. Hall provided some empirical evidence that appeared to support his theory. He found that the change in consumption was independent of lagged income, although he also found that lagged stock prices did affect current consumption significantly. Hall dismissed this latter result arguing that given the unpredictable nature of stock prices (they follow a random walk themselves) then it is not unreasonable to approximate consumption by a random walk.

However, evidence which refuted REPI soon became available as subsequent studies on consumption have (consistently) produced two well established facts:

1. *Consumption reacts too strongly to changes in actual income.* This is the so-called *'excess sensitivity'* phenomenon. If expectations are rational then Hall's permanent income hypothesis can be refuted since the claim that changes in consumption are unpredictable is not fulfilled. According to the random walk equation, only unpredictable changes in

[25]**When we talk about the failure of REPI we understand that failure to be that one of the Hall (1978) and Flavin's (1981) models of consumption.**

[26]These tests are known as 'orthogonality tests'.

actual income can affect consumption, so consumption should not react not too strongly to actual and past income changes[27].

2. *Consumption reacts too weakly to changes in permanent income.* This is the so-called *'excess smoothness'* result. This result is particularly damaging for the rational expectations permanent income hypothesis since it suggests that permanent income is more volatile than consumption thereby defying the original purpose of the permanent income hypothesis which attempted to explain why consumption appeared to be smoother than actual income.

Excess Sensitivity Tests[28]

A powerful rejection of REPI was provided by the excess sensitivity test of Flavin (1981). Her work was developed with two ideas in mind: to provide a stronger test for consumption than the reduced-form equation (2.11), and to attempt the identification of the consumer's reaction to both *anticipated* and *unanticipated* income shocks. Flavin's model pays specific attention to the role played by current income in providing new information about future income. Such information ought to be used by rational agents under the permanent income hypothesis to upgrade their expectations about permanent income. A shortcoming of this test (in Flavin's original form) is that modelling both the income and the consumption processes is required and the results to this test appear to be somewhat sensitive to the modelling specifications that are used[29].

Formally, Flavin's test used a trended $ARMA$ representation to model the time-series properties of the income process. It is assumed that *agents use that specification when forming expectations about future levels of income.* From the $ARMA$ process for income, it is possible to obtain the actual revision in permanent income warranted by the contemporaneous observation of current income. This revision is given by the forecast error in the $ARMA$ specification and such an error represents unanticipated news associated with the current observation of

[27]The predictability component depends on the nature of the income process so that if income follows a unit root the predictable component carries over to the future.

[28]**Hall and Mishkin's (1982) notion of excess sensitivity has a different interpretation to Flavin's.**

[29]See Deaton (1992) chapter 3.

income[30]. The size of the revision will then depend, amongst other things, on the parameters of the *ARMA* representation of the income process. According to this argument, one can 'specify a structural equation relating the change in consumption to the contemporaneous revision in permanent income (modelled using the income innovation) and the change in current income'. [pp. 976].

The null that Flavin tested in her paper is the truth of the permanent income hypothesis (in the form of equation (2.11)) together with an autoregressive specification for the process governing labour income. That null can be specified in terms of two equations

$$\begin{aligned} \Delta c_t &= \varepsilon_t = \alpha\eta_t \\ &= \alpha\sum_{\tau=0}^{T-t}(1+r)^{-\tau}(E_t - E_{t-1})\left(w_{t+\tau} + y^k_{t+\tau}\right) \end{aligned} \tag{2.21}$$

$$\phi(L)\, y_t = \varepsilon_t \tag{2.22}$$

where $y_t = w_t + y^k_t$. The first equation of the null comes from equations (2.10), (2.11) and (2.12). Flavin allows for the possibility of unanticipated capital gains, so the surprise in the y^k variable which represents non-labour income is allowed to be different from zero. Strictly speaking, the hypothesis that Flavin works with, 'the excess sensitivity' hypothesis, is a substantial generalisation of equation (2.11)[31] and it allows consumption to respond to current and lagged changes in income by more or less than is required by the permanent income theory. The extended model that Flavin works with is the following one [pp. 990]:

$$\begin{aligned} \xi(L)\, y_t &= \mu + \varepsilon_t \\ \Delta c_t &= \gamma + \theta\varepsilon_t + \beta(L)\,\Delta y_t + u_t \end{aligned} \tag{2.23}$$

[30] Flavin also suggests that the error in the *ARMA* representation for income can represent, for econometricians attempting to model consumption, not just the 'true innovation' in income, but also the predictive 'value of all the lagged values of variables observed by the individual, but not explicitly incorporated in the regression.' [pp. 991]. This is an issue related to Campbell's (1987) and West's (1988) superior information.

[31] In her 1993 paper, Flavin argues that 'the consumption data is generated by the excess sensitivity model' [pp. 665]

where $\xi(L) = \sum_{i=0}^{p} \xi_i L^i$, $\xi_0 = 1$ and $\beta(L) = \sum_{i=0}^{p} \beta_i L^i$; $\beta_0 \neq 1$. Flavin rearranges the $AR(p)$ income process equation to express it in terms of the error term ε_t which is then substituted into the consumption equation. Hence, in the unrestricted version of the model the first difference of consumption responds to current and lagged changes in income as well as the innovation in the income process. The β coefficients are measures of excess sensitivity of consumption to current income. They provide the amount of additional response of consumption to the new information contained in current income. Clearly, consumption innovations must be related to the amount of income innovation provided by the error term ε and should not, according to REPI, be affected by other variables. Hence, all the β coefficients that represent the extent to which consumption responds to previously predictable changes in income should be zero.

In her paper, Flavin runs an eight order autoregression ($p = 8$) for the process governing labour income. She imposes the restriction $\beta_0 = \beta_1 = \cdots = \beta_7 = 0$ on the system to obtain a constrained system that can be estimated. Data on non-durable goods from 1949(3) to 1979(1) were used to find that the likelihood ratio statistic for the hypothesis $\beta_0 = \beta_1 = \cdots = \beta_7 = 0$ was 27.02 for $\chi^2(8) = 21.96$. Hence the random walk specification of Hall is rejected by Flavin. [pp. 999]. The estimates for the first three sensitivity parameters were .335, .071 and .049. These results indicated strong excess response of consumption to current income. [pp. 1002].

One of the possible explanations for these results is related to the econometric techniques used by Flavin. Mankiw and Shapiro (1985) and Deaton (1992) pick up this point. Their argument is related to the actual form of modelling the income process when such process appears to be non-stationary. They criticise the method used by Flavin to account for the upward trending behaviour of income. Flavin deals with the non-stationary nature of the income process by fitting exponential time-trends to both consumption and income, and replacing consumption and income in the regressions by their residuals. Mankiw and Shapiro argued that excess sensitivity was induced by the detrending procedure, even if excess sensitivity were not present in the data. Basically, y is a non-stationary variable while Δc is stationary so that running a system like (2.23) cannot provide much information for both sides of the consumption equation are of a different order of integration[32]. The problems about making inferences about

[32] To see this note that in (2.23) the income equation is already in reduced form, and to obtain the reduced form for consumption we only need to substitute the income equation into the consumption equation (see Deaton (1992) pp. 89).

the coefficients on lagged income using standard t and $F-tests$ are essentially the same as the problems that occur in discerning the existence of a unit root in a univariate time series; the use of the standard normal tables at usual significance levels results in over-rejection. Deaton (1992) runs a Monte Carlo experiment[33] to test this point and finds that the $t-statistics$ for excess sensitivity on each of the income variables, and the test for excess sensitivity as a whole (an $F-test$) reject more than the customary 5%[34].

Stock and West (1988) challenged Mankiw and Shapiro's suggestion that excess sensitivity was the result of bad econometric practice. Stock and West used the concepts of cointegration and error correction to provide a means for testing excess sensitivity:

$$c_t = b_0 + b_1 c_{t-1} + b_2 y^d_{t-1} + b_3 y^d_{t-2} + u_t$$

where y^d is the same income measure used by Flavin. Now, if we define savings as $s_t \equiv y^d_t - c_t$ then the consumption equation can be expressed as

$$c_t = b_0 + (b_1 + b_3) c_{t-1} + (b_2 - b_3) \Delta y^d_{t-1} + b_3 s_{t-1} + u_t \tag{2.24a}$$

We see that this model resembles DHSY's study, where the savings variable plays the error correction role if we expect the coefficient of the lagged consumption variable $(b_1 + b_3)$ to be close to one. Sims, Stock and Watson (1990) have shown that in a regression of integrated variables of the same order, standard asymptotic theory can be applied to parameters that can be written as the coefficients on stationary variables. If consumption and disposable income are cointegrated, then the last two variables of the equation are stationary. Hence, it is possible to make inferences about the excess sensitivity parameters b_2 and b_3. Stock and West used Monte Carlo experiments to show that their technique works and the find evidence in favour of excess sensitivity. According to Stock and West, the problem with Flavin's test procedure was that the imposition of a unit coefficient upon the lagged consumption variable altered the asymptotic distributions of the estimates, but once we correct for this excess sensitivity appears

[33] Deaton himself recognises that; 'the Monte Carlo results, although tailored to reflect the actual data, do not generate results that look like Flavin's'. [pp. 94]

[34] The overall $F-test$ rejects 43% of the time, and the $t-test$ for β_0 and β_1 rejects 14% and 21% of the time respectively rather than the correct 5% [pp. 93].

to exist nonetheless.

Excess Smoothness Tests

The test for excess smoothness was developed by Campbell (1987) and Campbell and Deaton (1989) to test the empirical validity of 'Deaton's paradox'. Deaton (1987)[35] relaxed the assumption of stationary income to make inferences about the evolution of permanent income and thereby consumption. Trend-stationary and difference stationary income series were investigated.

Deaton demonstrates that if (labour) income is trend-stationary, and detrended income follows a stationary autoregression, changes in permanent income will be smaller than income innovations. Consumption will not fully respond to news about current income and will therefore be smoother than current income. This was the principal reason Friedman gave for the development of the permanent income hypothesis; permanent income should not react too strongly to changes in current income thereby making permanent income smoother than current income, a fact consistent with the data. Deaton also demonstrates that if the (labour) income process is stationary after first differencing, then permanent income will respond more than one for one to innovations in income. The permanent income hypothesis would then suggest that current income is smoother than permanent income when the (labour) income process is difference stationary. Hence permanent income is incapable of explaining why consumption is smoother than income in actual data. This is the essence of Deaton's Paradox.

Campbell and Deaton (1989) provide an example of this last result through a 'weak' test for excess smoothness in consumption[36]. Following Flavin, it is possible to write the innovation in consumption as a function of the expectational changes of future labour income. Provided that the income process follows an $ARMA(p,q)$

$$\alpha(L)\, y_t = \beta(L)\, \varepsilon_t \tag{2.25}$$

[35] Here we follow Deaton's 1992 book.

[36] We explain a 'stronger test' for excess smoothness in chapter 3.

then the innovation in consumption can be written as follows (see equation (2.10)):

$$\begin{aligned}\Delta c_t &= \frac{r}{1+r}\sum_{\tau=0}^{\infty}(1+r)^{-\tau}\Delta E_t y_{t+\tau} \\ &= \frac{\beta(\rho)}{\alpha(\rho)}\varepsilon_t \end{aligned} \quad (2.26)$$

where $\rho = (1+r)^{-1}$. (Hansen and Sargent (1981) demonstrated that this formula applies both for the stationary and non-stationary cases.) Now, if we consider a second-order autoregressive income process for example

$$(1-\alpha L)(1-(1-\delta)L)y_t = \varepsilon_t$$

then from (2.26) we can write

$$\Delta c_t = \frac{r(1+r)}{(r+\delta)(1+r-\alpha)}\varepsilon_t$$

The terms that play an important role in explaining the volatility of consumption to the type of income process that best fits the data are α and δ. We can see that if a unit root exists, $\alpha = 1$ and the equation above becomes

$$\Delta c_t = \frac{(1+r)}{(r+\delta)}\varepsilon_t \quad (2.27)$$

so that we can see that the ratio $\frac{(1+r)}{(r+\delta)}$ will be greater than one if $0 < \delta < 1$. The size of δ will therefore determine the effect of the income innovation on consumption. The smaller δ is $(0 < \delta)$, the higher the ratio will be and the more volatile consumption will be. The effects of δ on consumption are magnified or dampened by the root α. If the income series does not have a unit root, $\alpha < |1|$, then since $\frac{r}{(1+r-\alpha)} < 1$, then the effects of δ are dampened.

These are the principles that Campbell and Deaton (1989) employed when they tested for the volatility in consumption. Campbell and Deaton (1989) found that the labour income process that best fitted the US data was given by:

$$\Delta y_t = \underset{(3.2)}{8.2} + \underset{(5.5)}{0.442}\Delta y_{t-1} + \varepsilon_t, \; \sigma_\varepsilon = 25.2 \tag{2.28}$$

so that in their case, $\alpha(L) = 1 - 1.442L + 0.442L^2$, thus implying, $\alpha = 1$ and $1 - \delta = 0.442$. Substituting the values of $\delta = 0.558$ and $\alpha = 1$ to equation (2.27) yields

$$\Delta c_t = 1.76\varepsilon_t \tag{2.29}$$

for $r = 10\%$ p.a., equation (2.29) predicts that the standard deviation of changes in consumption ought to be at least 1.76 times than the standard deviation of labour income (i.e. $\sigma_{\Delta c} = 1.76\sigma_\varepsilon$ for $r = 10\%$). From equation (2.28) above and the data, Campbell and Deaton obtain:

$$\sigma_{\Delta y^p} = 1.76 \cdot 25.2 = 44.3$$

whilst the data on consumption stated that for consumption including durables, $\sigma_{\Delta c} = 27.3$, and for consumption excluding durables, $\sigma_{\Delta cn} = 12.4$. This therefore suggests that estimated permanent income is more volatile than labour income. Hence permanent income appears to be noisier than current income (and any measure of consumption). According to the permanent income hypothesis, the standard deviation of consumption and permanent income ought to be the same. If that were the case, then current labour income would be smoother than current consumption, a fact that is not encountered in the data as the standard deviation of consumption excluding durables is less than the standard deviation of current labour income. These results, together with Deaton's Paradox have led some economists to believe that 'the reason consumption is so smooth is because the permanent income theory is false' [Campbell and Deaton, pp. 358].

2.2.6 Some Explanations for the Failure of the Rational Expectations Permanent Income Hypothesis

In view of the results obtained from the tests of excess sensitivity and smoothness of consumption, economists have examined possible reasons for the failure of the permanent income hypothesis. In this section we mention some of the most important explanations.

Private Information

Campbell (1987) developed a model of savings under REPI to account for the possibility that agents may use a different information set to the one used by the researcher when making predictions about the behaviour of permanent income from labour (and capital) income. This information discrepancy can pose problems for excess smoothness tests but not for excess sensitivity tests which examine the significance of past income for making forecasts about the changes in consumption.

Excess smoothness tests (as defined in the previous section) are subject to this information discrepancy because they are based on expectations about future (labour) income. If the econometrician's information set differs from that one used by the agent then the predictions about future labour income and therefore permanent income are likely to differ between the econometrician and the agent too. Hence any predictions about the volatility of permanent income that emanate from the predictions made by the econometrician about future labour income may be flawed for they may not represent the actual behaviour of the representative agent.

Campbell (1987), Campbell and Deaton (1989) and West (1988) developed models that could account for this informational discrepancy. Campbell and Campbell and Deaton examine the (superior) information that is conveyed by savings in order to forecast labour income and hence permanent income. The principle used in those papers is simple yet very powerful: if agents expect lower future incomes and therefore a lower permanent income at time t, then they will reduce the amount of consumption at t. Given that by definition savings are the difference between current income and consumption, a reduction in consumption - explained by lower permanent income due to an expected decrease in future labour income - will induce an increase in savings. Hence, savings do provide information about agents' expectations about future labour income. This relationship is best explained through the following equations[37]:

$$c_t = y_t^p = r\left[A_t + \sum_{i=0}^{\infty} \rho^{1+i} E_t y_{t+i}\right] \tag{2.30}$$

$$s_t \equiv y_t + rA_t - c_t \tag{2.31}$$

Substituting (2.30) into (2.31) and rearranging yields

$$s_t = -\sum_{i=1}^{\infty} (1+r)^{-i} E_t \Delta y_{t+i} \tag{2.32}$$

Hence, savings equal the expected present value of future declines in labour income and therefore savings will rise (fall) if future labour income changes are revised downwards (upwards). This equation is referred to as the 'savings for a rainy day' aspect of the permanent income hypothesis, and can help overcome the information discrepancies mentioned earlier[38]. Take the savings equation

$$s_t = -\sum_{i=1}^{\infty} (1+r)^{-i} E\left(\Delta y_{t+i} | I_t\right) \tag{2.33}$$

where I_t denotes the agent's information set at t. If H_t is the econometrician's information set at t, and it is assumed that

$$H_t \subseteq I_t$$

then *the agent's information set encompasses the one used by the econometrician.* It is also assumed that the *econometrician observes the current saving decision of the consumer*, so that savings are a part of H_t. This is the crucial assumption that is needed to overcome the superior information problem. Given those two assumptions, taking the expectations of (2.33)

[37] Here y is labour income and corresponds to w and A are assets and correspond to y^k in the previous analysis.
[38] We give a short account of the Campbell paper in this chapter. A more detailed description is given in the next chapter.

conditional on the information set H_t:

$$E(s_t|H_t) = -\sum_{i=1}^{\infty}(1+r)^{-i} E(E(\Delta y_{t+i}|I_t)|H_t)$$

which, by the 'law of iterated expectations' and the two assumptions above, is equal to

$$s_t = -\sum_{i=1}^{\infty}(1+r)^{-i} E(\Delta y_{t+i}|H_t) \tag{2.34}$$

so that the econometrician's information set is used instead of the agent's. The crucial assumption here is that savings are observed by the economist so that savings enable the economist to bridge the gap between his information set and the information set of the agent.

From the *savings for a rainy day equation*, it follows that [Campbell, pp. 1253]

$$s_t - \Delta y_t + (1+r)s_{t-1} = -r\varepsilon_t \tag{2.35}$$

$$\varepsilon_t = \left(\frac{1}{1+r}\right)\sum_{i=0}^{\infty}\left(\frac{1}{1+r}\right)^i [E_t y_{t+i} - E_{t-1} y_{t+i}]$$

Equations (2.35) and (2.31) summarise the testable implications of the permanent income hypothesis and are exploited in the papers by Campbell and Campbell and Deaton. Testing for the validity of REPI using (2.35) is more powerful than using equation (2.11) as long as the data do not invalidate the intertemporal budget constraint (2.31) [Campbell, pp. 1254].

The saving equation (2.32) is also important from a statistical and econometric point of view as it allows us to test the permanent income hypothesis through (2.35). Assume that labour income is stationary after taking first differences. Then equations (2.30), (2.31) and (2.11) are also stationary in first differences, but savings[39] is stationary in levels. From the cointegration literature equation (2.31) and the fact that savings are stationary, consumption is a random walk, and total disposable income is stationary after first differencing, then it must be the case that a linear combination of consumption and income exists so that both variables

[39] Intuitively, saving is a discounted present value of changes in expected labour income. These changes must be stationary for otherwise they could be predicted.

are cointegrated. By Engle and Granger's theorem, an 'error correction mechanism' between the cointegrated variables exists which enables 'to put it into VAR form by dropping one of the elements of Δx_t (where $x_t = [y_t,\ A_t,\ c_t]'$) and replacing it with $a'x_t$ (where a' is the cointegrating vector from (2.31), $a = [1,\ 1, -1]$). [...] The resulting model is well-behaved and has the property of cointegration without the restrictions on the VAR coefficients' [Campbell 1987, pp. 1256]. One can accordingly, test REPI as a set of restrictions on a vector autoregression for the change in labour income and savings. Campbell exploits this characteristic of cointegration and proposes a VAR system of labour income innovations and savings (which is obtained by rewriting the ECM into VAR as mentioned above) to test REPI[40].

Campbell, Campbell and Deaton for the US and Attfield et al. (1990) for the UK showed that the data does reject the appropriate restrictions imposed on the VAR thereby implying that the permanent income hypothesis is flawed[41].

West (1988) also tackles the problem of inferior information. He considered a variance bounds test to examine the sensitivity of consumption under the hypothesis that income has a unit root. West investigates whether the findings in Flavin's permanent income hypothesis model vary if consumers are allowed to use additional information than that conveyed by lagged and current labour income for their predictions about their permanent income. If we denote the consumer's and the observer's information sets by I and H respectively, Flavin's model implies

$$c_t - Ec_t|I_{t-1} = \Delta c_t = y_{tI} - Ey_{tI}|I_{t-1}$$

where $y_{tI} = r\,(1+r)^{-1} \sum_{j=0}^{\infty} (1+r)^{-j}\, Ey_{t+j}|I_t$. Thus $var\,(\Delta c_t) = E\,(y_{It} - Ey_{It}|I_{t-1})^2 \equiv \sigma_I^2$. If only current and past income observations are used by the econometrician, then the observer can only hope for

$$\sigma_H^2 = E\,(y_{Ht} - Ey_{Ht}|H_{t-1})^2$$

So that, unless $H_t = I_t$, $var(\Delta c_t) \neq \sigma_H^2$; and it is likely that (Proposition I) that $\sigma_H^2 \geq \sigma_I^2 =$

[40]This test is explained in more detail in the next chapter.

[41]The test demonstrates that there is excess smoothness and that the orthogonality condition (when lagged income is introduced) is breached.

$var(\Delta c_t)$. This implies that the variance that an econometrician estimates is greater than the variance that consumers have. In other words, with lesser information about variables than is required for forecasts which can resemble the agents forecasts, excess smoothness results.

West, using the intertemporal budget constraint (2.31) is able to work out the difference between σ_H^2 and σ_I^2 and in doing so introduces another test for the permanent income hypothesis;

$$\sigma_H^2 = \sigma_I^2 + \left[(1+r)^2 - 1\right] var\left(y_{tI} - y_{tH}\right)$$

or

$$\begin{aligned} \sigma_H^2 &= var\left(\Delta c_t\right) + \left[(1+r)^2 - 1\right] var\left(c_t - rA_t - y_{tH}\right) \\ &= \sigma_{\Delta c}^2 + \sigma_v^2 \end{aligned}$$

Under the permanent income hypothesis, these equations must be true. If consumers do use more information than that conveyed by current and lagged income, then σ_v^2 must be statistically different from zero. If the permanent income hypothesis is true then $\sigma_H^2 - \sigma_{\Delta c}^2 - \sigma_v^2$ must be zero. By equations (2.25) and (2.26) $\sigma_H^2 = \psi^2 \sigma_\varepsilon^2$ and one can calculate σ_H^2 from the estimates of the income process, σ_v^2 can be calculated from the intertemporal budget constraint and $\sigma_{\Delta c}^2$ is obtained from consumption data.

West finds that the null $\sigma_H^2 - \sigma_{\Delta c}^2 = 0$ can be comfortably rejected at the 5% level. Moreover, the null $\sigma_H^2 - \sigma_{\Delta c}^2 - \sigma_v^2 = 0$ can also be rejected, implying that the 'insensitivity of consumption to news about income is unlikely to result purely from the use by the consumer of additional variables to forecast income' [pp. 23]. West introduces two modifications - wealth shocks and transitory consumption - to this model which are not able to explain the aforementioned insensitivity.

The superior information findings are however disputed by Muellbauer and Murphy (1993). They find that lagged saving does not have a significant negative effect on subsequent income for both the US and UK economies. This result questions, at least for the macroeconomy, the notion that private agents may have a superior information set to that one used by the econometrician. Their results can be challenged for they only use a 'moderately sophisticated income-forecasting process' [Muellbauer, 1994, pp. 15].

Variable Interest Rates

Campbell and Mankiw (1989) relax the assumption of a constant interest rate in the permanent income hypothesis. They examine two models of the Euler equation that allow for a varying and uncertain real interest rate. The first model considers a single forward-looking rational agent that consumes his or her permanent income, and the second model allows a proportion of consumers in the economy to be reluctant to substitute consumption intertemporally in response to interest rate movements. Both models are estimated using instrumental variables because the error term in the Euler equation may be correlated with the independent variables in the regression. Estimating the standard random walk equation by instrumental variables where the independent variables are current income innovations and the rate of interest, can be viewed as a restricted version of a more general equation system in which both the dependent and independent variables are regressed directly on the instruments. When there is more than one instrument, the martingale equation places over-identifying restrictions on the systems of equations. Those restrictions are used to test the permanent income hypothesis.

Permanent income consumers The log-linear generalisation of the consumer's Euler equation that accounts for variable interest rates is according to Campbell and Mankiw

$$\Delta c_t = \mu + \sigma r_t + \varepsilon_t$$

The rate of interest is now permitted to be correlated with the error term which is still uncorrelated with lagged variables. Note that r_t is the rate of interest contemporaneous with the innovation in consumption at time t. By definition, σ is the intertemporal elasticity of substitution and should depict the fact that high ex ante real interest rates lead to rapid consumption growth.

Estimation of this equation using instrumental variables yielded disappointing results. Campbell and Mankiw give a number of reasons that explain why an equation like this one is probably misspecificed [pp. 198-200]:

1. The hypothesis that consumption growth is unpredictable - tested with a Wald test for the hypothesis that all coefficients are zero - is rejected at the 5% level or better, which

is inconsistent with Hall's (1978) interpretation of the data. If the PIH were true, and σ were zero, consumption should be a random walk. Furthermore, the over-identifying restrictions of this equation are rejected at the 5% level or better whenever lagged real interest rates are included in the set of instruments.

2. The estimates of σ are highly unstable and small unless the nominal interest rate is used as the instrument in which case it exceeds one.

3. Reversing the Hall regression yields estimates for $\frac{1}{\sigma}$ that are not extremely large as would be predicted by the rational expectations permanent income hypothesis.

Campbell and Mankiw suggest that this misspecification of the model is due to the exclusion of rule-of-thumb consumers.

Rule-of-Thumb Consumers A more general model in which a fraction λ of income goes to individuals who consume their current income, and the remainder goes to consumers that behave according to equation (2.11) was considered. The following model is estimated by instrumental variables

$$\Delta c_t = \mu + \lambda \Delta y_t + \theta r_t + \varepsilon_t$$

where $\theta = (1 - \lambda)\,\sigma$. The two coefficients of interest are now λ (the proportion of rule of thumb consumers) and θ (the effects of interest rates on consumption). Campbell and Mankiw find a number of interesting implications [pp. 200-203]; firstly, rule-of-thumb consumers appear to exist in the economy since the coefficient on current income is substantive and statistically significant. Secondly, they find evidence that the ex ante real interest rate is not in any way determining the amount of consumption growth. The coefficient θ is small and indicates that the intertemporal elasticity of substitution for the permanent income consumers is small as predicted by theory. Finally, the robustness of these results is enhanced by the fact that the over-identifying restrictions are never close to being rejected.

These results suggest that the expected changes in consumption depend on expected changes in income because rule-of-thumb consumers exist in the economy. This explains why the excess sensitivity phenomenon may occur and it suggests that taking a single representative agent to explain the behaviour of aggregate consumption is not entirely correct.

Liquidity Constraints[42]

Deaton (1991) examines the nonlinearities in the intertemporal budget constraint associated with borrowing constraints in a model where agents face uncertain income and are allowed to be 'impatient' (in the sense that the rate of time preference is greater than the rate of interest) but are prudent at the same time (in the sense that they have a convex marginal utility function). In the face of uncertain income, precautionary motives will interact with liquidity constraints since the inability to borrow when times are bad provides an additional incentive to accumulate assets when times are good even for impatient consumers. Deaton shows that the appropriate consumption rule chosen under this framework is dependent on the time-series behaviour of income.

With convex marginal utility, borrowing constraints and impatient consumers Hall's problem[43] is modified. Because those changes make the Euler Equation very difficult to solve unless the problem is redefined, Deaton specifies the framework in terms of the function 'cash in hand' which acts as the state variable. The familiar Euler Equation is changed with the introduction of the borrowing constraint:

$$\lambda(c_t) = \max\left[\lambda(x_t),\ \frac{1+r}{1+\delta}E_t\lambda(c_{t+1})\right]$$

where λ denotes marginal utility and x is cash in hand defined as $x_t = A_t + y_t$ and $\frac{1+r}{1+\delta} = \beta < 1$. Since consumers are not allowed to spend above their cash in hand no borrowing is permitted in the current period[44]. The solution to the problem is difficult to obtain because the marginal utility function λ is nonlinear. To proceed, Deaton changes the problem to look for a stationary

[42] We spend some time with this paper to demonstrate the difficulties associated with solving the Euler equation once we move away from the assumptions made by Hall (1978) and Flavin (1981).

[43] Here we show the modification associated with an i.i.d. income process.

[44] The constraint implies that consumption cannot be higher than cash in hand, so given the convexity of the marginal utility function, that cannot be lower than $\lambda(x_t)$. If $\lambda(x_t) > \beta E_t\lambda(c_{t+1})$, the constraint will bind; otherwise with no liquidity constraints, the two marginal utilities are equated as agents seek to equate marginal utilities across time.

stochastic optimum in which consumption is a function of the state variable x_t, $c_t = f(x_t)$. The marginal utility of money (price of consumption $p(x_t)$) is defined for this purpose as

$$p(x_t) = \lambda [f(x_t)] \text{ or } c_t = \lambda^{-1} [p(x_t)]$$

This enables the Euler equation to be written as

$$p(x_t) = \max \left[\lambda(x), \; \beta \int p\{(1+r)(x - \lambda p(x) + y)\} \, dF(y) \right]$$

After expectations have been taken into account. Note that the only source of uncertainty continues to be labour income. The 'Euler equation' now equates the marginal utility of money today to the maximum value of either the marginal utility of cash in hand in the constrained situation or the discounted expected value of tomorrow's marginal utility of money. The solution to this equation is then used to characterise the equilibrium properties of the marginal utility of money and thus the policy function $f(x)$. The solution is obtained (pp. 1227) with the specification of an updating rule used for a finite number of periods

$$p_n(x_t) = \max \left[\lambda(x), \; \beta \int p_{n-1} \{(1+r)(x - \lambda p_n(x) + y)\} \, dF(y) \right]$$

and with the backward iteration of the functions $p_0(x)$, $p_1(x), \ldots, p_n(x)$ until the function converges. In this problem, $n = 0$ is the last period where everything is spent $p_0(x) \equiv \lambda(x)$ if there are no bequests. The period before that, $n = 1$, $p_1(x)$ is determined by either the borrowing constraint or marginal utility. The problem is solved recursively. This modified problem has the following properties 'the convexity of $\lambda(x)$ implies $p(x)$ is convex. [...] With borrowing constraints, the convexity of $p(x)$ determines the degree of precautionary savings. [Moreover] $p(x)$ is more convex than $\lambda(x)$, so that the inability to borrow in adversity reinforces the precautionary motive' [pp. 1227]. Deaton shows that for i.i.d. income there exist a unique x^* such that:

1. $c = f(x) = x$ when $x \leq x^*$and,

2. $c = f(x) \leq x$, when $x \geq x^*$

What are the implications of this solution? For a given level of assets and a draw of labour income, the agent will spend everything and no assets are accumulated if the total value of assets and income is below the critical level x^* (viewed as an optimal level of assets used by agents to buffer against fluctuations in income). If the amount of cash in hand, x, is greater than the critical value, something will be held over and a new positive level of assets will be carried forward to be added to next period's income.

Some characteristics of the solution are as follows:

1. The distribution of consumption will not be symmetric; the consumer can prevent consumption from being high but it cannot prevent it from being too low.

2. The evolution of marginal utility of money $p(x)$ is a martingale in the standard case, but under borrowing restrictions it follows a renewal process i.e. as long as the consumer carries forward positive assets, we have the martingale result, but as soon as the assets fall to zero, the process loses its memory and starts again.

3. The level of x^* and therefore the amount of smoothing are also determined by the coefficients ρ (which is the coefficient in the isoelastic utility function $u(c) = c^{-\rho}$ and represents prudence) and σ (which is the variance of the income process and therefore represents uncertainty).

When the income process is *serially correlated but stationary* the results change slightly. y now becomes the state variable together with x since both convey information about future consumption decisions. Consumers' behaviour is similar to the previous case; an amount x^* exists whereby levels of x below it lead to all cash in hand to be consumed and levels of x above x^* lead to a proportion of the consumer's cash in hand to be saved for future periods. The actual level of x^* does again depend on the arguments given above, but this time it also depends on the level of income, or the 'state'.

Some characteristics of the solution for serially correlated stationary income:

1. Consumption is smoother than income, but the distribution of consumption is still asymmetric. This time, savings are a much more effective cushion against high consumption than against low consumption.

Table 2.1: Standandard Deviations for Income and Consumption for an AR(1) Income Process

Autocorrelation coefficient, ϕ	-0.4	0	0.3	0.5	0.5	0.9
1. *s.d.*(y)	10.9	10.0	10.5	11.5	14.0	22.5
2. *estimated s.d.* (y)	10.8	10.2	10.0	11.4	13.3	27.5
3. *estimated s.d.*(c)	4.6	5.1	6.7	7.6	10.4	25.9
ratio 3/2	0.43	0.50	0.67	0.67	0.78	0.94

2. The coefficient of the AR process plays an important role as we can see from table 1[45]. Deaton explains these results as follows; 'by assumption, $\delta > r$, so assets are costly to hold. The precautionary demand is a powerful motive to hold assets, but the smoothing of consumption over long autocorrelated swings requires more assets, and more sacrifice of consumption, than is the case when income is i.i.d. or negatively correlated. Positive autocorrelation also restricts the ability to smooth consumption. Once cash on hand falls below [the level x^*], no assets will be held, even if the bad income shock that produced the situation is a signal that further bad income draws are to follow. These bad times have to ridden out without any assets to cushion their impact.' [pp. 1234-5]

When the income process is nonstationary, the analysis is modified to make all the variables in the problem (income, cash in hand and consumption) stationary. Deaton does this by dividing all those variables by the level of income and the solution, obtained with the same techniques as above, is given in terms of these new variables. What are the implications for consumption behaviour when the income process is nonstationary?

1. For the no serial correlation case, a variable $w^* = \frac{x_t^*}{y_t}$ exists such that when $w < w^*$, assets will remain at zero; for $w > w^*$ the evolution of assets is more difficult to follow although it can be shown by simulations that w will eventually decline below w^* in finite time and therefore assets will eventually become zero so that all income is consumed. When income is a random walk and borrowing constraints are present, smoothing is not desirable: if income is above average it is expected to continue that way, so the additional income enables the consumer to get closer to the ideal level of consumption. When a bad draw in income occurs nothing can signal to the agent when the future trough in

[45] See Deaton (1991), Table 1, pp. 1234.

income will occur in order to smooth consumption then. In consequence, the combination of 'the persistence of the random walk and the binding liquidity constraints precludes the accumulation of assets'. [pp. 1238]

2. For the serially correlated case simulations provide a solution to the problem. Deaton finds that as soon as a bad state is announced, savings switch from zero to positive and the consumer begins to accumulate assets. As the slump continues, the savings ratio stops rising and falls below zero if the slump is long enough. Assets go on rising for a while after the savings ratio has started falling, but eventually reach a ceiling above which they cannot go. At this point, the negative savings ratio and asset income help protect consumption against the effects of income which has negative expected growth over the slump. Eventually the slump ends and the boom takes over. As this happens the consumer uses all the accumulated assets to finance a spending boom and sits the boom with consumption equalling income. This result is the exact opposite of that implied by the permanent income hypothesis and does not appear to be supported by the data. Deaton dismisses this point by arguing that 'even in the absence of borrowing restrictions, conditions for aggregation to representative agents are implausible, so that a representative agent formulation is perhaps even more than usually misdirected when there are liquidity constraints'. [pp. 1241] This clearly opens the debate about aggregation in a representative agent framework.

Excess Sensitivity

Flavin (1993) considers an alternative specification to REPI where consumption exhibits excess sensitivity to current income[46]. Excess sensitivity to current consumption, which may originate if individuals are liquidity constrained, is introduced with the assumption that an individual will consume all of its permanent income and a proportion β of its transitory income each period

$$c_t = y_t^P + \beta y_t^T \tag{2.36}$$

[46] In this section we concentrate on the theoretical aspects of Flavin's paper. In chapter 3 we concentrate on the econometric aspects.

where $0 < \beta < 1$ and permanent income is defined by (2.30) and y_t^T denotes transitory income defined as

$$y_t^T = y_t + \left(\frac{r}{1+r}\right) A_t - y_t^p \tag{2.37}$$

Assuming that there are no unanticipated capital gains, then the innovation in consumption is given by

$$\Delta c_t = \beta \Delta y_t + (1-\beta)\left(\frac{r}{1+r}\right)\sum_{\tau=0}^{\infty}\left(\frac{1}{1+r}\right)^{\tau}(E_t - E_{t-1})\, y_{t+\tau} \tag{2.38}$$

Thus, 'even though transitory income and permanent income were defined as the transitory and permanent components of total income, inclusive of asset income, in the statement of the excess sensitivity hypothesis, the terms involving asset income cancel out, with the result that the first difference of consumption is a weighted average of the first difference of labour income and the expectational revision of the annuity value of future labour income' Flavin [pp. 655]. Note that the innovation in consumption due to an innovation in permanent income is less than one to one and can therefore be interpreted as liquidity constraints and/or precautionary savings. To see this, assume that consumers receive news today that their labour income will (permanently) increase tomorrow. If consumers cannot borrow to increase and therefore smooth their consumption they may decide to consume a proportion of their transitory income today until their higher labour income is realised. Moreover, because consumers are reluctant to consume all of their transitory income and therefore all of their disposable income upon hearing the news that they cannot borrow, one can interpret this reluctance as perhaps being explained by precautionary savings.

In this model, savings are now scaled by $(1-\beta)$, viz.

$$s_t = (1-\beta)\left[y_t - \left(\frac{r}{1+r}\right)\sum_{\tau=0}^{\infty}\left(\frac{1}{1+r}\right)^{\tau} E_t y_{t+\tau}\right] \tag{2.39}$$

This model is able to explain both excess sensitivity and smoothness. Sensitivity will occur if

$\beta \neq 0$, and smoothness will occur if

$$var\,(\Delta c_t) = \beta^2 var\,(\Delta y_t) + 2\beta\,(1-\beta)\,cov\,(\Delta y_t, \Delta y_t^p) + (1-\beta)^2\,var\,(\Delta y_t^p) < var\,(\Delta y_t^p)$$

Using data for the US, Flavin found that whilst REPI could be decisively rejected, the restrictions that the excess sensitivity hypothesis imposes on the bivariate system of labour income innovations and savings proposed by Campbell and Campbell and Deaton could not be rejected by the data.

Buffer Stock/Precautionary Savings

Carroll (1992, 1997a) examines the role of precautionary savings in a REPI framework where consumers face important income uncertainty, have a precautionary motive and are impatient. The model predicts similar results[47] to those suggested by Deaton (1991) as consumers engage in 'buffer stock' saving behaviour. Carroll demonstrates that consumers in this framework have a target wealth-to-permanent income ratio - w^* in Deaton's work - such that if wealth is below target the precautionary motive dominates and consumers save, while if wealth is above target impatience will dominate prudence and consumers will dissave. The important result of Carroll's work is that he is able to put forward a formulation for the components of the optimal level of cash in hand w^*[48].

Carroll shows that the Euler consumption equation in a problem similar to Deaton's takes the form:

$$E_t \Delta \ln c_{t+1} \approx \rho^{-1}\,(r-\delta) + \left(\frac{\rho}{2}\right) var_t\,(\Delta \ln c_{t+1}) + \varepsilon_{t+1} \qquad (2.40)$$

if shocks to consumption are lognormally distributed. ρ is the coefficient of risk aversion, r is the rate of interest, δ is the rate of time preference. Consumption growth depends on three factors; the degree of impatience over precaution, a random effect and the conditional variance

[47] Carroll's model differs from that of Deaton in that it does not impose liquidity constraints and in that income is divided into its permanent and transitory parts. Consumers act as if they were liquidity constrained because whilst impatient, they show prudence. Carroll claims that Deaton's results are due to his assumptions about impatient consumers and the convex marginal utility function rather than due to liquidity constraints per se. Carroll also claims that liquidity constraints reinforce the results of his paper.

[48] Deaton did not provide such explanation.

of next year's consumption given information available this year. The first two components are standard to intertemporal consumption behaviour, but the variance term had, until Carroll's work, been somewhat neglected. The variance term is proven to play a significant role in consumers' behaviour. Carroll derives an expression for the average (aggregate) variance of consumption term to obtain further insights

$$E_{.,t}\left[var_{i,t}\left(\Delta \ln c_{i,t+1}\right)\right] \approx \left(\frac{2}{\rho}\right)\left[g - \frac{\sigma^2_{\ln N}}{2} - \rho^{-1}\left(r - \delta\right)\right] \tag{2.41}$$

where g is the growth rate of permanent income and $\sigma^2_{\ln N}$ denotes the variability of permanent income. From the Euler equation, Carroll (1992) demonstrates[49] [pp. 130-2] that the expected variance of consumption growth is negatively related to wealth. Carroll (1997a) (footnote 20) argues that this equation serves as a means of providing inferences about the target level of wealth that consumers will hold to buffer against an uncertain future, although the equation 'should be viewed as a heuristic tool rather than as a rigorous analytical framework. That said, I have found no parameter values for which this kind of reasoning from equation (9) gives the wrong answer.'[pp. 20].

The variance term (and presumably the target level of wealth) in (2.41) increases when the growth rate of permanent income, g and the rate of time preference, δ increase. These results are intuitive; a higher expected growth rate for permanent income will probably reduce the amount of income uncertainty for the individual and given an impatient nature the consumer would consume more. When the consumer discounts the future less (i.e. δ falls) the individual is willing to postpone consumption and wealth increases. From the equation we also notice that the variance term (the target level of wealth) decreases (increases) when the rate of interest and the variability of permanent income increase ($\sigma^2_{\ln N}$). Again the explanation of this result for both variables is straight-forward; as the interest rate increases, the consumer becomes less impatient as consuming more in the present becomes more expensive and so the agent is willing to increase the amount of wealth that he or she holds. When the variability of the permanent income component increases, the prudent component of the individual's behaviour overshadows

[49]Kimball (1990) has shown that for isoelastic utility functions, precautionary saving declines as wealth increases. Since precautionary saving adds to wealth over time, consumption will become less depressed. Hence, the reduction of precautionary saving when wealth increases generates the extra growth in consumption when $E_t var(\Delta \ln c_{t+1})$ is high.

the impatient one and the consumer begins to accumulate more wealth. Finally, the coefficient of relative risk aversion (ρ) has offsetting effects; a higher coefficient represents a stronger precautionary motive (more wealth accumulated due to the $\frac{\rho}{2}$ term) but at the same time, a higher ρ leads to a lower intertemporal elasticity of substitution: thus $\rho^{-1}(r-\delta)$ decreases.

In a recent paper, Carroll (1997b) has suggested that tests for the validity of the Euler equation - as given by equation (2.40) - using instrumental variables must be interpreted with caution because the variance term in that equation is likely to be endogenously determined by all of the variables in (2.41). Thus, for instance, it would be incorrect to use interest rates as one of the instruments to test the Euler equation.

Aggregation with Finite Lives[50]

The existence of finitely lived life-cycle consumers may explain the problems of excess sensitivity and smoothness that appear in the data. Clarida (1991) studied the aggregate stochastic implications of Modigliani's life cycle hypothesis to explain the first and second moment properties of changes in per capita consumption. The principal finding of the paper is that 'smooth per capita consumption in the presence of a permanent shock to per capita labor income is exactly the outcome one should expect from a properly aggregated life cycle model in which saving for retirement, as well for consumption smoothing, is a motive for asset accumulation' [pp. 853-4]. Since savings are required to finance consumption in retirement agents will not react so strongly to permanent changes in their labour income as they need to save for retirement. This means that the MPC of a change in permanent income ought to be less than one (REPI assumes an MPC of one) and it is likely to decline monotonically with age. Clarida's main consumption equation is (in per capita terms)

$$\Delta c_t = \varphi\lambda + \mu\varepsilon_t + \phi\eta_{t-1} \tag{2.42}$$

[50]We examine a version of this model in chapter 6 of this Ph.D. thesis.

where

$$\mu = \sum_{j=1}^{w} \mu(j); \ \mu(j) = \frac{1 + (1+r)^{-1} + \cdots + (1+r)^{-(w-j)}}{1 + (1+r)^{-1} + \cdots + (1+r)^{-(n-j)}} < 1$$

$$\varphi \equiv \frac{n\mu}{w} \geq 1$$

and

$$\eta_{t-1} = \frac{1}{n} \sum_{i=w+1}^{n-1} \varepsilon_{t-n+i} + \frac{1}{n} \sum_{j=1}^{w-1} \varepsilon_{t-n+j} \left(1 - \left(\frac{\mu j + 1}{\mu(1)}\right)\right)$$

n is the number of periods the individual lives, w is the number of periods the individual works for ($n - w$ is therefore the retirement period), j is the age of a consumer at time t, μ is the marginal propensity to consume out of labour income and $\varepsilon_t \left(\equiv \frac{we_t}{n}\right)$ and $\lambda \left(\equiv \frac{wg}{n}\right)$ are functions of the error and the drift term in the following specification for labour income $y_t = g + y_{t-1} + e_t$. Clarida shows that for plausible demographic assumptions, 'the variance of changes in per capita consumption predicted by a properly aggregate life cycle model is substantially less than is implied by the representative agent permanent income hypothesis when shocks to per capita income are permanent' [pp. 854]. The following implications come from (2.42):

1. Aggregate per capita consumption has positive drift even though, by definition of REPI, individual consumption is a random walk without drift.

2. The drift in per capita consumption exceeds the drift in per capita labour income ($\varphi\lambda > 1$ since φ >1 if $r > 0$) whenever the rate of interest is greater than zero.

3. Changes in per capita consumption are correlated with lagged innovations in labour income.

Near Rationality

Cochrane (1989) looks at the utility loss suffered by agents when they follow alternative (not rational) decision rules. Cochrane finds that the utility cost to an agent that decides to set consumption equal to current income rather than to permanent income is less than ten cents to a dollar per quarter. The utility costs are small because the utility costs of deviating from an optimum are an order of magnitude smaller than the deviation itself. An agent will not change its consumption unless a shock forces the agent to be relatively far away from the optimum so that the costs associated with changing consumption behaviour are exceeded by the utility gain from moving consumption to the optimum point.

Partial Adjustment

Attfield, Demery and Duck (1992) examined the possibility that consumers may be slower to adjust to changes to their permanent income than predicted by the permanent income hypothesis. Slow adjustment may be explained by inertia or habit formation.

Two models are examined by Attfield et al.. The first one (referred to as PIH1) is a conventional, forward-looking quadratic cost of adjustment model and the second specification (PIH2) examines the time absorbing costs of planning to enable the agent to reach the optimal level of consumption.

PIH1 In this model, the permanent income problem developed by Flavin (1981) is modified by Attfield et al. through the assumption that consumers wish to minimise the following loss function

$$\min\ L = E_t \sum_{j=0}^{\infty} \rho^j \left[a_0 \left(c^*_{t+j} - c_{t+j} \right)^2 + a_1 \left(c_{t+j} - c_{t+j-1} \right)^2 \right]$$

subject to the constraint

$$\sum_{j=0}^{\infty} \rho^j c_{t+j} = (1+r) A_t + \sum_{j=0}^{\infty} \rho^j y_{t+j}$$

where $\rho = (1+r)^{-1}$. In the paper, $a_0 = \frac{1}{2}$ and $a_1 = \frac{\alpha}{2}$ where, α is a cost of adjustment parameter and c^* is the optimal level of consumption at each time period. The quadratic cost of adjustment makes large consumption changes undesirable. After straight-forward manipulation, Attfield et al. arrive at the following result

$$\Delta c_t = (1+r)(1-\theta)\Delta c_{t-1} + \theta\omega_t \tag{2.43}$$

where $\omega_t = \sum_0^\infty \rho^i \Delta E_t \Delta y_{t+i}$ and θ is a function of the two roots required to solve the first order condition of the problem. The innovation in consumption is no longer a martingale process but an $AR(1)$ one. One must therefore include the lagged dependent variable and examine, under the null hypothesis, whether the last term (the error term) is white noise. This is done by checking whether lagged variables have significant coefficients.

This model explains excess smoothness and excess sensitivity. The innovation in consumption and hence in permanent income is likely to be correlated with lagged income because that innovation is written as a function of ω_t which is itself correlated with lagged income. From (2.43)

$$\frac{var(\Delta c)}{var(\omega)} = \frac{\theta^2}{1-(1+r)^2(1-\theta)^2}$$

and excess smoothness will arise if the ratio is less than one, i.e. if $1 - \frac{2}{[1+(1+r)^2]} < \theta < 1$.

PIH2[51] The costs of adjusting a variable may not be specific to the time the adjustment takes place but may extend to other periods [Attfield et al. pp. 1206]. The planning time required to making sure consumption is at its desired level may produce slow adjustment of consumption to changes in permanent income: according to REPI forward-looking consumers need to ensure that not only is actual consumption at t equal to desired consumption in period t but also that actual consumption in all future periods equals the expected level of desired consumption in all future periods.

Adjustment costs may arise in a rational expectations permanent income problem because given the tendency to discount the future, one would expect the planning effort devoted to the immediate future to exceed the amount of effort put to plan the future. The proportion of permanent income that is unpredictable well before the current consumption decision was made, is accountable for making actual consumption deviate from desired consumption; if a

[51]This model is explained in more detail in subsequent chapters of this thesis.

high proportion of permanent income were predictable well in advance, then actual consumption will be close to its desired level. Thus, in this model, it is assumed that there is 'some time span sufficiently long to ensure that any component of current permanent income which was predictable [well] in advance will have its full effect on current consumption.' [pp. 1212]

Mathematically

$$c_t = E_{t-n}y_t^p + \sum_0^{n-1} \gamma_i \Delta E_{t-i} y_t^p$$

where n defines the time span over which adjustment is less than complete, the γs should follow the pattern $0 \leq \gamma_0 \leq \gamma_1 \leq \cdots \leq \gamma_{n-1} < 1$ Since the change in expectations is unpredictable; $E_t y_t^p - E_{t-1} y_t^p = e_t$ then $E_{t-1} y_t^p - E_{t-2} y_t^p = \zeta_0 e_{t-1}$, where $\zeta_i = [1 + r(1-\gamma_i)]$. By algebraic manipulation, Attfield et al. arrive at

$$\Delta c_t = \sum_0^n \phi_i e_{t-i} \tag{2.44}$$

where $\phi_0 = 1 - \beta_0$, $\phi_j = [\frac{\beta_{j-1}}{\rho - \beta_j}]$ for $0 < j < n$; $\phi_n = \frac{\beta_{n-1}}{\rho}$ and $\beta_i = \frac{[1-\gamma_i]\zeta_{i-1}\beta_{i-1}}{(1-\gamma_{i-1})}$ for i>0 and $\beta_0 = 1 - \gamma_0$. Thus the change in consumption is an $MA(n)$ process, where the error is a function of the change in the expectations about permanent income. This equation can explain excess sensitivity and excess smoothness; since e_t is white noise by definition, we can write

$$var(\Delta c_t) = var(e_t) \sum_0^n \phi_i^2$$

and excess smoothness will arise if $\sum_0^n \phi_i^2 < 1$.

Excess sensitivity arises in this problem because lagged shocks to permanent income, which are likely to be correlated to lagged innovations in labour income, are shown to influence the current change in consumption.

Attfield et al. find that US and UK data favour both the PIH1 and PIH2 models over the random walk specification of Hall. They also find that the PIH2 specification is the preferred one and that the US data is not able to formally reject that specification. For the UK, such rejection is less decisive.

Information-Aggregation

Goodfriend (1992) has shown that the orthogonality restrictions implied by intertemporal optimisation, rational expectations and information processing need not hold under the aggregation of randomly heterogeneous and imperfectly informed representative agents. Economic variables are generated by aggregate and relative components which agents must distinguish to follow optimal decision rules. Agents may be imperfectly informed (as Goodfriend assumes) because data are published with at least one period lag and must distinguish which part of their current income innovation is an aggregate one and which one is a relative one. Then the relative persistence of each component can explain the failures of REPI. These issues are investigated in more detail by Pischke (1995).

Pischke extended Goodfriend's model by assuming that aggregate information may play a small role in household decisions since 'ignoring it is not very costly for most households' [pp. 807][52]. The optimal reaction of agents to innovations in their individual and aggregate economy-wide income is investigated to assess their subsequent consumption decisions. Throughout Pischke's models, it is assumed that agents have individual specific income processes that are different from the time series structure of aggregate income. It is also assumed that the other assumptions of REPI hold.

All the models consider agents that have identical income processes, but each consumer faces a different realization of this process every time period. The simple income process assumes $\Delta y_{it} = \varepsilon_t + (1 - L)\, u_{it}$,where subscripts i denote individual variables while no subscripts refer to aggregate variables. Both errors are uncorrelated by assumption[53]. Pischke examines three scenarios:

1) Complete Aggregate Information. In this case the micro agent has full contemporaneous information on aggregate income so that individual and aggregate income innovations can be distinguished. This results in the permanent income model of Hall and Flavin.

2) Unobservable Aggregate Shocks. The individual cannot distinguish between aggregate and individual income components. A simple income process for the micro agent then looks

[52] This assumption can be based on near-rational considerations.

[53] Other income specifications where the aggregate errors are white noise and the individual errors are random walks can be considered (as well as other modifications).

like this

$$\Delta y_{it} = \eta_{it} - \theta\eta_{it-1}$$

where θ is a function of the individual and aggregate components ε and u[54]. With this income process, Pischke shows that the per capita consumption innovation will be

$$\Delta c_t = \theta\Delta c_{t-1} + A\varepsilon_t$$

where $A = (1 - \frac{\theta}{1+r})$. Consumption does not follow a random walk but an $AR(1)$ process. Excess sensitivity and smoothness are present in this model; if the researcher runs an excess sensitivity test of the type $\Delta c_t = \alpha + \beta\Delta y_{t-1} + e_t$ then the estimated sensitivity coefficient (which should be zero) is

$$\dot{\beta} = \frac{cov\left(\Delta c_t,\ \Delta y_{t-1}\right)}{var\left(\Delta y_{t-1}\right)} = \frac{E\left\{A\left(\frac{\varepsilon_t}{1-\theta L}\right)\varepsilon_{t-1}\right\}}{\sigma_\varepsilon^2} = A\theta$$

and the degree of excess sensitivity depends on the parameter θ. Excess smoothness arises because

$$\frac{\sigma_{\Delta c}}{\sigma_\varepsilon} = \frac{A}{\sqrt{1-\theta^2}} < 1$$

if the rate of interest is small enough and $\theta > 0$. The representative model of Hall would hold if the aggregate and the individual income processes had the same persistence properties.

3) Lagged Information about Aggregate Shocks. (This is Goodfriend's model). In this case, the individual i can only observe both y_{it} and the aggregate shock ε_{t-1} at time t. The consumer has also knowledge of the history of both variables so that it is able to infer u_{it-1}. The income process for the individual now takes the form

$$\Delta y_{it} = v_{it} - u_{it-1}$$

[54]We explain this model in more detail in other chapters of this thesis.

where $v_{it} = \varepsilon_t + u_{it}$ and so the consumer cannot distinguish between the permanent (aggregate) and transitory (individual) components. A consumer will attribute part of the current period innovation to each component given the relative variances of both components. The optimal consumption response to an innovation in income will have two parts; one that accounts for the new innovation in the agent's income and a term that corrects for the error made in predicting both components of income in the previous period. Goodfriend obtains the (aggregate) consumption response to an innovation in income as

$$\Delta c_t = \frac{\omega + r}{1 + r}\varepsilon_t + (1 - \omega)\,\varepsilon_{t-1}$$

where $\omega = \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + \sigma_u^2}$ is a kind of signal extraction parameter. Aggregate consumption does not follow a random walk but an $MA(1)$ process. According to Pischke [pp. 815], the response to an aggregate shock is larger in the no information model than in the lagged information model since $\frac{\omega+r}{1+r} < A = (1 - \frac{\theta}{1+r})$. The lagged information model contains information on contemporaneous shocks, whilst the no information model also contains new information on lagged shocks so that the agent's response to innovations in income differs in both models.

The lagged information model will exhibit both excess sensitivity and smoothness. The excess sensitivity coefficient is given by

$$\hat{\beta} = \frac{E\left\{\left[\frac{\omega+r}{1+r}\varepsilon_t + (1 - \omega)\,\varepsilon_{t-1}\right]\varepsilon_{t-1}\right\}}{\sigma_\varepsilon^2} = 1 - \omega$$

which is different from zero. Excess smoothness arises because

$$\frac{\sigma_{\Delta c}}{\sigma_\varepsilon} = \sqrt{\left(\frac{\omega + r}{1 + r}\right)^2 + (1 - \omega)^2}$$

which Pischke demonstrates is less than one for small values of r.

Pischke looks at the predictions made by the models using empirical estimates for the individual and aggregate parts of the income process. Pischke assumes that the individual income process and the aggregate income process are described by an $MA(2)$ specification in first differences. By definition the consumption processes, differ for the case of no information (consumption follows an $ARIMA(2,1,2)$) and the lagged information model (consumption changes are an $MA(1)$). Pischke obtains estimates for β and the variance ratio for different values for the coefficients in the $MA(2)$ income process and for different measures of the variability of the two components of the income process. He finds that both the no information model and the lagged information model predict parameters which are very close to the time series properties of aggregate consumption but he is not able to say which of the two models best explains the data. Both these models perform better than the full information model.

In two recent studies, Demery and Duck (1999, 2000) have identified the appropriate restrictions that the models of Goodfriend and Pischke imply for the dynamics of aggregate consumption. Using US and UK time series data, Demery and Duck found that although Pischke's model can explain some features of the data better than the permanent income hypothesis, both models can be formally rejected by the data.

Other explanations

Christiano, Eichenbaum and Marshall (1991) have looked at the time horizon in which consumption decisions are made. They specifically address the criticism that consumers have no grounds for planning their actions in an annual, quarterly or monthly basis and they argue that agents make decisions on a *continuous time basis.* The assumption that agent's decisions intervals match the data sampling interval is thereby replaced by the assumption that agents make decisions at time intervals finer than that. Their work is based on the previous finding that temporal aggregation bias can induce serial correlation and spurious Granger-causality findings. Specifically, if the planning interval is shorter than the data interval spurious serial correlation and Granger-causality could be obtained. Christiano et al., find that their continuous time variant of Hall and Flavin's models satisfies the martingale hypothesis. They argue that the empirical findings which suggested that the first difference in consumption is serially correlated and Granger-caused by a variety of other variables, are explained entirely by temporal aggregation bias.

Bernanke (1985) suggested that durable goods and adjustment costs could explain the excess sensitivity phenomenon. Bernanke studies the consumer's optimal spending patterns on durable and nondurables that are jointly determined.

Muellbauer (1988) examined whether lagged dependent variables in consumption represented agents' expectations or adjustment costs, habits or the durability of goods. Muellbauer explains that habits, like convex adjustment costs can account for the excess smoothness finding although he rejects these arguments as the complete explanation for the failure of the REPI specification. Muellbauer (1994) suggests that other types of adjustment costs may in fact account for the failures of the permanent income hypothesis.

Heaton (1993) examined habit formation and time-aggregation issues. He found that monthly consumption changes were negatively correlated, not positively correlated as were the quarterly changes. He suggested a model where the utility function would depend positively on stocks that come from the accumulation of purchases, and negatively on habits. In that model durability would dominate over short periods but habits become important as the observation period increases.

2.3 Conclusion

In this chapter we have reviewed the most important developments in the consumption literature over the last 50 years. In the earlier part of this chapter we reviewed the consumption function developed by Keynes as well as a number of subsequent studies which tried to explain why the Keynesian consumption function failed when estimated on UK and US data. Two of those studies have had a profound effect on the theory of consumption: the permanent income and life cycle hypotheses of Friedman and Modigliani. These two models were embraced by economists for their theoretical properties and their early empirical success.

However, by the late 1970s, it appeared that consumption behaviour could not be modelled empirically. Muellbauer and Lattimore (1995) summarise the then state of consumption research as 'far from satisfactory' [pp. 222]. In 1978, two important studies which have had a profound impact on the consumption literature were published: Davidson et al.'s error correction specification and Hall's consumption function. Davidson and his colleagues put forward a consumption specification that could explain UK data, the findings of previous models, exhibit parameter stability over time and conformed to steady-state postulates of economic theory. Their work was guided by time series techniques which were then beginning to be understood. In this chapter we reviewed their consumption function paying particular attention to the short (dynamic) and long-run characteristics of their model.

Following the rational expectations revolution - most forcefully advocated by Lucas in his famous critique - Hall introduced rational expectations and a number of assumptions into the permanent income model of Friedman. Hall's consumption function advocated that the innovation in consumption represented new information which becomes available to consumers. Because information is unpredictable by nature it cannot be systematically predicted and therefore it should not be possible to predict consumption. Hall tested this claim and found empirical support for his hypothesis[55].

Subsequent research found that Hall's consumption function could be formally rejected in a number of different countries as consumption appeared to react too strongly to current income (the excess sensitivity hypothesis) and too weakly to permanent income (the excess smoothness hypothesis). Thus, it should be possible to predict the innovation in consumption. In this chapter we looked at some of the empirical studies that have tested the truth of the rational expectations permanent income hypothesis and we also reported the principal theories which claim to explain why Hall's consumption function fails at the aggregate level. Two main conclusions can be drawn from this exercise: first, that the main bulk of empirical research on aggregate consumption has been undertaken on two data sets for the US and the UK economies that end in the mid 1980s. Second, there have been many different explanations given for the failure of the random walk hypothesis. Most explanations involve different economic theories/concepts which in turn lead to very different consumption specifications. There is therefore no common concensus about what the consumption function should look like.

In the next few chapters we attempt to deal with these two issues: in the next chapter we address the first issue as we construct two new data sets for the US and the UK that incorporate data up to 1996 to test whether consumption still suffers from excess sensitivity and smoothness. In the remainder of the thesis we address the second issue by constructing consumption specifications which are guided by a number of theories and which try to be as general as possible.

[55]Strictly speaking, Hall only attempted to test the hypothesis that consumers try to smooth marginal utility. Because his marginal utility function was linear this implied that he was testing whether consumers smooth their consumption or not.

Chapter 3

Excess Sensitivity and Smoothness: Evidence from New US and UK Data Sets

3.1 Introduction

Much of the empirical work that has tested the rational expectations permanent income hypothesis (REPI[1]) has used a US data set which was constructed by Blinder and Deaton (1985). Campbell (1987), Campbell and Deaton (1989) and Flavin (1993) are some of the authors that have used this data set that runs from the second quarter of 1953 to the last quarter of 1984. In all these studies REPI fails on the accounts that consumption is both too sensitive to current income and that it is smoother than predicted by the permanent income theory. For the UK, Attfield, Demery and Duck (1990) constructed a data set running from 1955:1 to 1987:2 to mirror Blinder and Deaton's only to report similar findings. In this chapter we extend the data sets[2] used by all these authors to incorporate ten more years of available observations that will establish the strength of the findings reported in all these papers. We pay particular attention to the econometric methods used by Campbell and Deaton and Flavin to examine how these extra observations affect the results reported in those papers.

[1]We refer to the REPI as the models that were developed by Hall (1978) and Flavin (1981).

[2]I am grateful to David Demery for helping me with these data sets.

An examination of the last ten years can provide further insights into the smoothing behaviour of permanent income consumers. In the UK in the 1980s for instance, the consumption to income ratio grew to unprecedented levels, 'peaking in 1987-8 at the highest levels seen in the last 40 years.' (Muellbauer (1994) pp. 1). The early 1990s saw a stunning reversal, with the 'personal-sector saving ratio in 1992, [...] at the highest level since 1980.' (Muellbauer (1994) pp. 1). The latter part of the decade has resulted in a consumption expenditure boom. For the US the story is somewhat different as it did not enjoy such a pronounced recession although it has experienced a consumer expenditure boom in the mid to latter part of the 1990s. Historically, the 1980s and 1990s will probably be characterized by the advent of financial liberalization and the growth in the participation of consumers in the stock market so that it may be possible to take a step back and question the validity of some of the assumptions that were used by Hall when he derived his model. Are consumers constrained in their borrowing? Were they so forward looking as to forecast the rise in the stock market in the last few years? Are rates of return constant?

3.2 An Explanation of the Data Used

Blinder and Deaton (1985) constructed a time series data set from the National Income and Product Accounts (NIPA hereafter) which they believed would help to test a number of consumption specifications. To arrive at these series several adjustments and procedures to the official NIPA data were made.

1. The 1975 tax rebate is removed from the disposable income series on the grounds that this (sizeable) rebate was unanticipated and was not generated by the same stochastic process as the rest of the data. If one were to include this rebate, the estimates of the time series process for income would be distorted.

2. Interest paid by consumers to businesses was subtracted from disposable income to treat these payments symmetrically with business interest payments to consumers. This is done in order to comply with the tax system.

3. 'Personal' non-tax state and local payments are subtracted from taxes and from govern-

ment purchases and added to consumption. The current dollar levels were deflated by a consumption of services deflator. This change raises income and consumption equally.

4. Clothing and shoes which are considered by the NIPA as nondurable goods, were reclassified as durables.

5. Blinder and Deaton achieve the breakdown of disposable income into capital and labour components by attributing propietor's income and personal income taxes to labour and capital according to their overall factor shares and by deducting social insurance contributions from labour income.

All the resulting series are expressed as per capita aggregates. For the UK, similar measures were undertaken by Attfield, Demery and Duck (1990).

The ten extra years of available data do not only provide an interesting period for testing the validity of the permanent income hypothesis but are also specially interesting for the US given the Bureau of Economic Analysis's comprehensive revisions to the annual NIPA not only in terms of scope but also in terms of the number of years subject to revision. The revised and updated versions of these series were obtained following the 10th comprehensive revision to the National Income and Product Accounts. Three major revisions were proposed: '(1) Definitional and classificational changes that update the accounts to portray more accurately the evolving U.S. economy, (2) statistical changes that update the accounts to reflect the introduction of new and improved methodologies and the incorporation of newly and available and revised source data, and (3) presentational changes that update the NIPA tables to reflect definitional, classificational, and statistical changes to make the tables more informative.' (Survey of Current Business, January/February 1996, pp. 1) The most important innovations are;

1. 'improved chain-type measures of real output and prices that eliminate the overstatement of real GDP growth for periods after the base year and the understatement of real GDP growth for periods before the base year' (Survey of Current Business, 1996, pp. 1),

2. 'a new treatment of government investment that provides a more complete picture of investment through the consistent treatment of fixed assets whether purchased by the public

or private sector and that improves the international comparability of U.S. estimates of saving and investment' (Survey of Current Business, 1996, pp. 1),

3. 'a new methodology for calculating depreciation that improves the empirical basis of these estimates by replacing straight-line depreciation patterns with estimates based on studies of prices of used equipment and structures in resale markets' (Survey of Current Business, 1996, pp. 1-2).

The data obtained from the revised NIPA estimates for this study runs from 1959:3 to 1996:1 whilst the data obtained from the ONS runs from 1955:1 to 1996:2. The same adjustments made by Blinder and Deaton for the US and Attfield, Demery and Duck for the UK were made to the data. This renders four real per capita series on total consumption, non-durable and services consumption, disposable income and labour income for both the US and the UK.

3.3 Consumption, Permanent Income and Innovations - Univariate Framework Results for Excess Smoothness

In the next two sections, we follow the procedures used by Campbell and Deaton to examine the excess smoothness debate. This section follows Campbell and Deaton's simple approach of comparing the variance of the innovation in consumption with a measure of the 'estimated' variance of permanent income innovations to test for excess smoothness.

3.3.1 Overview and Unit Root Tests

According to REPI, the innovation in consumption is driven by the expectational change in the discounted sum of current and future labour income

$$\Delta c_{t+1} = \Delta y^p_{t+1} = r \sum_{i=1}^{\infty} (1+r)^{-i} (E_{t+1} - E_t) y_{t+i} \tag{3.1}$$

This result was first noted by Flavin (1981). If the process for labour income can be modelled correctly, a researcher will be able to calculate the innovations in permanent income given by the right hand side of the expression above and compare these to the innovations in consumption (these can be calculated directly from the data). Hansen and Sargent (1981), Quah (1990) have shown that the innovation in permanent income can be sensitive to the process assumed for labour income. If it is assumed that the income process can be represented by a *trend stationary income process* $y_t = B(L)\epsilon_t$[3] then we will have

$$r\sum_{i=1}^{\infty}(1+r)^{-i}(E_{t+1}-E_t)\,y_{t+i} = \left(\frac{r}{1+r}\right)B\left(\frac{1}{1+r}\right)\epsilon_{t+1}$$

where $B(L)$ denotes the polynomial in the lag operator. On the other hand, if the *labour income process is a difference stationary process* $\Delta y_t = A(L)\varepsilon_t$ then we will have

$$r\sum_{i=1}^{\infty}(1+r)^{-i}(E_{t+1}-E_t)\,y_{t+i} = A\left(\frac{1}{1+r}\right)\varepsilon_{t+1}$$

where $A(L)$ denotes the polynomial in the lag operator. Thus, comparing the variance of the innovation in consumption with $var(\epsilon_t)$ and $\left(\frac{r}{1+r}\right)B\left(\frac{1}{1+r}\right)$ for a trend stationary income process, or $var(\varepsilon_t)$ and $A\left(\frac{1}{1+r}\right)$ for a difference stationary income process allows a simple test for the truth of REPI. Previous research on unit roots for US data has reported that difference stationary processes explain the behaviour of labour income better than trend stationary processes. When A and $var(\varepsilon_t)$ have been estimated on US aggregate time-series data, the implied variance of permanent income has been significantly larger than the variance of consumption changes suggesting that consumption is smoother than predicted by REPI. As Quah points out, 'this result is remarkably robust across alternative specifications for A' (footnote 10, pp. 457). This result is intuitive; if labour income is non-stationary, innovations to this process will be persistent and will therefore imply a revision to permanent income (and thus consumption) of a similar amount. Since it can be established from the data that the volatility of consumption is less than the volatility of labour income, this being the primary reason Friedman developed the permanent income hypothesis, we find the result that consumption is smoother than permanent income. (This fact has been referred to in the literature as the 'Deaton Paradox').

[3] 'We can, without loss, take the trend to be identically zero since here we are interested only in the second-moment properties of consumption and income' Quah (1990), pp. 455.

We carry out this type of test with our two data sets to examine whether the above definition of permanent income can explain consumption. This provides a preliminary test for Flavin's definition of permanent income. We begin by testing for the presence of unit roots in all the variables that will be used throughout the analysis in this chapter through the tests developed by Dickey and Fuller (DF thereafter), Phillips and Perron (PP) and the variant of the DF test that allows for structural breaks developed by Banerjee, Lumsdaine and Stock (BLS) (1992). The results for the unit root tests without structural breaks are reported in tables 1 and 2 (DF) and 3 and 4 (PP) for the US and in tables 5 and 6 (DF) and 7 and 8 (PP) for the UK. We test two types of nulls for the DF and PP Unit Root tests: i) that the series X_t has a unit root against the alternative that X_t is stationary against a linear trend and ii) that the series X_t has a unit root against the alternative that X_t is stationary around a fixed mean. For the first type of tests we run a regression

$$\Delta X_t = \alpha_0 + \alpha_1 X_{t-1} + \alpha_2 t + \sum_{j=1}^{p} \gamma_j \Delta X_{t-j} + \varepsilon_t \tag{3.2}$$

and we test the Null hypothesis that a) $\alpha_1 = 0$ (through a $t - test$) and b) $\alpha_1 = \alpha_2 (= \alpha_0) = 0$ (through an $F - test$[4]). The lagged terms in ΔX_t are included in the DF test to ensure that there is no correlation in the errors. For the second type of tests we run the regression

$$\Delta X_t = \alpha_0 + \alpha_1 X_{t-1} + \sum_{j=1}^{p} \gamma_j \Delta X_{t-j} + \varepsilon_t \tag{3.3}$$

and test the hypothesis that a) $\alpha_1 = 0$ and b) $\alpha_1 (= \alpha_0) = 0$. The Phillips and Perron statistics were devised as an alternative to the inclusion of lagged terms in the equations above since they adjust the test statistics to take account of serial correlation and potential heteroskedasticity in the disturbances. In all tables, the values reported in parentheses denote the critical values at the 10% level of significance.

To test for the possibility that structural breaks are present in our data, we follow the procedures suggested by BLS. BLS estimate the 'asymptotic distributions for recursive, rolling

[4]Recall that both types of tests do not have the conventional critical values.

and sequential tests for unit roots and/or changing coefficients in time series regressions. The recursive and rolling tests are based on changing subsamples of the data. The sequential statistics are computed using the full data set and a sequence of regressors indexed by a 'break' date.' (Journal of Business and Economic Statistics, 1992, pp. 271). The *recursive minimum ADF t-statistic* and *rolling minimum ADF t-statistic* are tests based on the ADF test so that the observations on X_t are assumed to be generated by equation (3.2). The *recursive minimum ADF t-statistic* is computed using subsamples $t = 1, ..., k$, for $k = k_0, ..., T$ where k_0 is a start-up value and T is the size of the full sample. Equation (3.2) is estimated for each subsample and the minimum (maximum) value of the $t(k/T)$ across all the subsamples is chosen and compared to the critical value of $t_{DF}^{\min}$ ($t_{DF}^{\max}$) provided in table 1 of BLS. Rolling statistics are computed using subsamples that are constant fraction δ_0 of the full sample, rolling through the sample. Again, the minimum (maximum) values of the $t(k/T)$ across all the subsamples is chosen and compared to the critical value of $t_{DF}^{\min}$ ($t_{DF}^{\max}$) provided in table 1 of BLS. The *sequential minimum ADF t-statistic* is based on the observation that the variable X_t is generated by the following equation,

$$\Delta X_t = \alpha_0 + \alpha_1 X_{t-1} + \alpha_2 t + \sum_{j=1}^{p} \gamma_j \Delta X_{t-j} + \alpha_3 D + \varepsilon_t$$

which allows for a single shift or break in a deterministic trend at an unknown date. The deterministic regressor D allows for the possibility of a shift or jump in the trend at period k. In the shift in trend model,

$$\begin{aligned} D &= t \text{ for } (t > k) \\ D &= 0 \text{ for } (t \leqslant k) \end{aligned}$$

and the shift in mean model

$$\begin{aligned} D &= 1 \text{ for } (t > k) \\ D &= 0 \text{ for } (t \leqslant k) \end{aligned}$$

The sequential statistics are computed using the full sample, sequentially incrementing the date

Table 3.1: Dickey-Fuller Tests for the US; Trend

Test (C.V.10%)	$\alpha_1 = 0$ (-3.13)	$\alpha_0 = \alpha_1 = \alpha_2 = 0$ (4.03)	$\alpha_1 = \alpha_2 = 0$ (5.34)
Variable			
Δyd	-3.932	5.157	7.731
yd	-2.205	10.14	2.897
Δy	-3.558	4.225	6.332
y	-1.955	8.049	2.288
Δc	-4.510	6.828	10.22
c	-3.782	8.488	7.167
Δcn	-3.567	4.274	6.410
cn	-2.641	6.592	4.088

Table 3.2: Dickey-Fuller Tests for the US; No Trend

Test (C.V.10%)	$\alpha_1 = 0$ (-2.57)	$\alpha_0 = \alpha_1 = 0$ (3.78)
Variable		
Δyd	-3.824	7.318
yd	-1.156	12.68
Δy	-3.481	6.064
y	-1.154	10.292
Δc	-4.524	10.25
c	-0.465	5.194
Δcn	-3.426	5.870
cn	-1.28	6.389

of the hypothetical break (or shift). Allowing k (the unknown date of the hypothetical break or shift) to be increased sequentially, minimum values of $t(k/T)$ for the shift in trend and shift in mean models are compared to the critical values provided in table 2 of BLS. A further test of interest involves testing the null $H_0 : \alpha_3 = \alpha_1 = 0$ in both the trend-shift and mean-shift models. Table 2 in BLS reports the critical values of the F-statistics used to test these hypotheses. Table 3 in BLS reports the size and power of the recursive, rolling and sequential tests. The unit root tests for structural breaks are reported in tables 9 and 10 for the US and 11 and 12 for the UK. We remind the reader at this point that tables 1 to 3 in BLS were obtained through Monte Carlo simulations for 100, 200 and 500 observations. We do not have those exact number of observations and so comparison to the critical values is not totally accurate.

Table 3.3: Phillips and Perron Tests for the US; Trend

Test (C.V.10%)	$\alpha_1 = 0(z)$ (-18.2)	$\alpha_1 = 0(T)$ (-3.13)	$\alpha_0 = \alpha_1 = \alpha_2 = 0$ (4.03)	$\alpha_1 = \alpha_2 = 0$ (5.34)
Variable				
Δyd	-151.57	-12.455	51.71	77.56
yd	-10.59	-2.24	17.43	2.6134
Δy	-147.9	-12.157	49.27	73.9
y	-6.193	-1.772	14.195	1.658
Δc	-106.11	-9.205	28.233	42.34
c	-8.149	-2.045	27.144	2.102
Δcn	-95.826	-8.443	23.756	35.63
cn	-4.191	-1.238	40.851	1.038

Table 3.4: Phillips and Perron Tests for the US; No Trend

Test (C.V.10%)	$\alpha_1 = 0(z)$ (-11.2)	$\alpha_1 = 0(T)$ (-2.57)	$\alpha_0 = \alpha_1 = 0$ (3.78)
Variable			
Δyd	-151.38	-12.48	77.868
yd	-0.329	-0.682	24.073
Δy	-147.73	-12.183	74.21
y	-0.378	-0.689	20.027
Δc	-106.1	-9.239	42.666
c	-0.034	-0.092	38.499
Δcn	-95.31	-8.438	35.596
cn	-0.249	-0.841	61.072

Table 3.5: Dickey-Fuller Tests for the UK; Trend

Test (C.V.10%)	$\alpha_1 = 0$ (-3.13)	$\alpha_0 = \alpha_1 = \alpha_2 = 0$ (4.03)	$\alpha_1 = \alpha_2 = 0$ (5.34)
Variable			
Δyd	-3.627	4.405	6.581
yd	-1.627	5.793	2.016
Δy	-3.947	5.199	7.798
y	-1.891	7.974	2.143
Δc	-3.772	4.756	7.116
c	-2.237	4.339	2.823
Δcn	−3.741	4.674	7.004
cn	-2.280	4.214	2.876

Table 3.6: Dickey-Fuller Tests for the UK; No Trend

Test (C.V.10%)	$\alpha_1 = 0$ (-2.57)	$\alpha_0 = \alpha_1 = 0$ (3.78)
Variable		
Δyd	-3.477	6.074
yd	0.901	6.978
Δy	-3.893	7.578
y	0.420	9.723
Δc	-3.702	6.871
c	0.300	3.624
Δcn	-3.677	6.762
cn	0.244	3.372

Unit Root Tests: Evidence

To establish the number of times that a variable needs to be differenced before it becomes stationary, we first examine the null that all the differenced variables possess a unit root (i.e. $\alpha_1 = 0$) against the alternative that these differenced variables are stationary[5]. We then report the results for the case of all the variables in levels. The results of the DF and PP tests on both equations, suggest that US data appears to be integrated of order one and that the drift component is significant.

For the UK, both the DF and PP tests report that all data are integrated of order one. As for the US data, we find that the drift components appear to be significant in UK data.

[5]The test that all variables had a unit root when they were differenced twice was rejected in all cases and is not reported here.

Table 3.7: Phillips and Perron Tests for the UK; Trend

Test (C.V.10%)	$\alpha_1 = 0(z)$ (-18.2)	$\alpha_1 = 0(T)$ (-3.13)	$\alpha_0 = \alpha_1 = \alpha_2 = 0$ (4.03)	$\alpha_1 = \alpha_2 = 0$ (5.34)
Variable				
Δyd	-190.99	-14.939	74.393	111.59
yd	-7.471	-1.927	11.036	2.608
Δy	-172.05	-13.285	58.854	88.28
y	-6.444	-1.868	8.15	2.109
Δc	-169.71	-13.098	57.188	85.782
c	-4.110	-1.554	12.787	2.367
Δcn	-146.85	-11.556	44.499	66.748
cn	-3.3488	-1.449	16.749	2.313

Table 3.8: Phillips and Perron Tests for the UK; No Trend

Test (C.V.10%)	$\alpha_1 = 0(z)$ (-11.2)	$\alpha_1 = 0(T)$ (-2.57)	$\alpha_0 = \alpha_1 = 0$ (3.78)
Variable			
Δyd	-189.52	-14.834	110.03
yd	0.586	0.905	14.491
Δy	-171.42	-13.253	87.818
y	0.326	0.447	10.193
Δc	-167.24	-12.958	83.955
c	0.654	1.154	17.376
Δcn	-143.87	-11.407	65.044
cn	0.605	1.247	23.387

Table 3.9: Recursive and Rolling Augmented Dickey-Fuller Tests of Unit Roots; US data

		Recursive		Rolling	
	t_{DF}	t_{DF}^{min}	t_{DF}^{max}	t_{DF}^{min}	t_{DF}^{max}
yd	-2.65	-2.865	-0.372	-3.3	-0.548
y	-2.12	-2.123	-0.043	-3.159	-0.770
c	-3.32	-3.327	-0.462	-2.567	-0.400
cn	-2.36	-2.598	-0.404	-2.506	-0.218

Table 3.10: Sequential Augmented Dickey-Fuller Tests of Unit Roots; US data

	Trend Shift		Mean Shift	
	t_{DF}^{min}	F_t^{max}	t_{DF}^{min}	F_t^{max}
yd	-3.000	5.142	-2.961	4.932
y	-2.538	20.642	-2.064	13.05
c	-3.016	13.336	-2.993	13.504
cn	-2.523	9.891	-2.481	9.354

Structural Break Tests

The results from the recursive and rolling tests for all the series for both the US and UK suggest that the presence of a unit root cannot be rejected in any case. The sequential tests indicate the presence of a unit root for all the series and for the US, the F-tests suggest that there are no structural breaks in the data as the joint hypothesis $H_0 : \alpha_3 = \alpha_1 = 0$ cannot be rejected. For the UK, the critical values of the $F - test$ are close to the 2.5% significance level. We conclude with the evidence from all the tests that all the series in the UK have a unit root without structural breaks. Thus it is necessary to calculate the value of the lag operator $A(L)$.

Table 3.11: Recursive and Rolling Augmented Dickey-Fuller Tests of Unit Roots; UK data

		Recursive		Rolling	
	t_{DF}	t_{DF}^{min}	t_{DF}^{max}	t_{DF}^{min}	t_{DF}^{max}
yd	-1.857	-2.371	-0.3766	-3.147	-0.0915
y	-2.214	-2.642	-0.2698	-2.473	0.15233
c	-1.885	-2.2211	1.2	-3.1738	0.71416
cn	-1.9062	-2.3359	1.22	-2.8502	0.66735

Table 3.12: Sequential Augmented Dickey-Fuller Tests of Unit Roots; UK data

	Trend Shift		Mean Shift	
	t_{DF}^{min}	F_t^{max}	t_{DF}^{min}	F_t^{max}
yd	-3.346	18.251	-3.357	18.38
y	-3.54	20.429	-3.719	22.551
c	-3.355	19.848	-3.561	22.365
cn	-3.459	21.805	-3.711	24.288

3.3.2 Levels data

To compare our results with Campbell and Deaton's, the following equation was estimated on their claim that an $ARI(1,1)$ process tends to explain the behaviour of labour income well

$$\Delta y_t = \underset{(0.009953)}{0.05588} - \underset{(0.08366)}{0.018477}\Delta y_{t-1} + \varepsilon_t \qquad \begin{array}{l} \sigma_\varepsilon = 0.1064923 \\ R^2 = 0.0003 \end{array} \tag{3.4}$$

(standard errors are reported in parentheses). According to REPI, the innovation in consumption ought to be

$$\Delta c_t = \frac{(1+r)}{(1+0.018477)+r}\varepsilon_t \tag{3.5}$$

The multiplier on the right-hand side is 0.982 when r is zero and 0.983 when the rate of interest is 10%. Thus (3.5) predicts that the standard deviation of consumption should be 0.98 times the standard error of labour income. The standard error of labour income was calculated in (3.4) to be 0.106, whilst the standard deviation of consumption changes calculated from the data set is 0.081 therefore smoother than the standard deviation predicted by REPI ($0.106 \times 0.98 = 0.104$). The standard deviation of nondurable consumption (thought to be a closer approximation to the domain of the permanent income hypothesis) is 0.0489 suggesting again that permanent income innovations are smoother than actual innovations in consumption. Even scaling[6] the standard

[6] 'The c_{nt} (nondurable consumption) measure has the disadvantage that it is only a component of consumption; to use it, one must postulate that total consumption is unobservable and a constant multiple of c_{nt}.' Campbell (1987), pp. 1260. Thus we write $c_t = \lambda c_{nt}$, λ being this scaling factor. c_t then becomes our nondurable measure of consumption. Campbell adds the following warning:

'Blinder and Deaton report that the share of nondurables and services in measured total consumption expenditure has displayed a secular decline over the sample period. This casts some doubt on the practice of using nondurables and services consumption as a proxy for the total; nevertheless I follow this tradition and estimate deviation of non-durable consumption to render it comparable to the predicted innovation does not change the nature of the results. When we scale this standard deviation by the ratio of the mean of consumption to the mean of consumption of non-durables we get $0.0489 \times 1.18 = 0.058$.

For the UK, the $ARI(1,1)$ for labour income is given by the equation,

$$\Delta y_t = \underset{(1.104)}{4.704} - \underset{(0.0787)}{0.04295}\Delta y_{t-1} + \varepsilon_t \qquad \begin{array}{l} \sigma_\varepsilon = 13.40 \\ R^2 = 0.018 \end{array} \tag{3.6}$$

and the change in consumption is given by

$$\Delta c_t = \frac{(1+r)}{(1+0.042947)+r}\varepsilon_t \tag{3.7}$$

The multiplier on the right-hand side is 0.996 for a rate of interest of zero and 0.962 for $r = 10\%$ and it is therefore somewhat sensitive to the rate of interest used. Again, (3.7) predicts that the standard deviation of consumption should be 0.96 times the standard error of labour income which was calculated as being 13.40 in (3.6). As in the US case, when we compare the predicted standard deviation from REPI ($13.40 \times 0.96 = 12.899$) to the actual standard deviation of consumption (12.507) we find evidence that consumption in the UK appears to be less volatile than predicted by the permanent income hypothesis.

The principal shortcoming of this type of test is that it relies on the labour income specification to be fundamentally correct. In this case, the equations for labour income for both the UK and the US perform poorly; the coefficients do not appear to be significant, there is a very low R^2 and for the US there is evidence of structural breakdown in our equation from Chow tests. Campbell and Deaton have argued that a logarithmic version of the permanent income hypothesis maybe more desirable, their justification being that 'most of the relevant time series are more easily transformed to stationarity in logarithmic form, and most atheoretical specifications of consumption functions have found that form tends to fit the data better.' (pp. 358).

---

a constant scale factor.' Campbell (1987), footnote 15, pp. 1260.

Table 3.13: Dickey-Fuller Tests for the Change in Labour Income

Variable	$\alpha_1 = 0$	$\alpha_0 = \alpha_1 = 0$	$\alpha_1 = 0$	$\alpha_0 = \alpha_1 = \alpha_2 = 0$	$\alpha_1 = \alpha_2 = 0$
$\Delta \log y(US)$	-3.043	4.633	-3.554	4.211	16.314
$\log y(US)$	-2.736	12.63	-1.827	8.910	4.455
$\Delta \log y(UK)$	-4.377	9.581	-4.362	6.346	9.516
$\log y(UK)$	-0.541	9.799	-2.682	9.078	3.613

3.3.3 Logarithmic Data

A first step required to conduct tests with logarithmic data is to establish the order of integration of the variables required for the test, in this case labour income. DF unit root tests conducted for labour income are reported in table 13[7] and show that the variable is $I(1)$ for both the US and the UK.

In order to compare the robustness of the results reported in the papers of Campbell and Deaton and Attfield, Demery and Duck to the extended data sets, we transform the REPI model into logarithmic form and adhere to the suggestion that the increased volatility observed in the data can be modelled better in terms of logarithmic expressions. The following equations represent the logarithmic approximations derived by Campbell and Deaton to test the validity of the permanent income hypothesis; for the change in consumption we have

$$\frac{\Delta c_{t+1}}{y_t} \approx \frac{r}{1-\mu} \sum_{i=1}^{\infty} \rho^i \left(E_{t+1} - E_t\right) \Delta \log y_{t+i} \tag{3.8}$$

where μ is the mean growth of labour income, r continues to represent the rate of interest and ρ is equal to $(1+\mu)/(1+r)$. A restriction that $r > \mu$ is imposed so that the sum is finitely bounded. The 'savings for a rainy day' equation is given by[8]

$$\frac{s_t}{y_t} \approx -\sum_{i=1}^{\infty} \rho^i E_t \Delta \log y_{t+i} - \kappa \tag{3.9}$$

under the conditions that r and μ/r are small so that the savings ratio is small[9] (and κ is a constant defined by Campbell and Deaton). An equation of interest, used extensively to test

[7] The first two columns refer to regression (3.3) whereas the remaining columns refer to regression (3.2).
[8] In the next section we explain the economic implications and rationale of this expression.
[9] Otherwise the approximations do not hold.

the validity of REPI, is obtained from the linear combination of equations (3.8) and (3.9)

$$\frac{s_t}{y_t} - \Delta \log y_t - \frac{s_{t-1}}{\rho y_{t-1}} \approx -\sum_{i=1}^{\infty} \rho^i \left(E_{t+1} - E_t\right) \Delta \log y_{t+i} \approx -\frac{\Delta c_t}{y_{t-1}} \tag{3.10}$$

Campbell and Deaton claim that the equivalence of the first and last expressions of this equation can be checked directly from the data. Equation (3.10) can be used to detect excess smoothness in consumption. Based on an $AR(1)$ for the rate of growth of labour income, we note that the term in the centre can be expressed as a function of ε_t

$$\sum_{i=0}^{\infty} \rho^i \left(E_t - E_{t-1}\right) \Delta \log y_{t+i} = \sum_{i=0}^{\infty} (\pi\rho)^i \varepsilon_t = \frac{\varepsilon_t}{1 - \pi\left(1 + \mu - r\right)} \tag{3.11}$$

where π is the autoregressive coefficient on the $AR(1)$ equation for the growth rate of labour income. Hence once we obtain a value for π, it is possible to calculate the (theoretically) predicted standard deviation of changes in consumption, these changes being expressed as ratios of the level of labour income, the right hand side of (3.10). The predicted standard deviation of the change in consumption is equal to the standard error of the innovation in the autoregressive equation multiplied by the coefficient on ε_t in equation (3.11). This can then be compared with either the actual standard deviation in the innovation of consumption or the standard deviation of the approximately equivalent expression on the left hand side of (3.10) which we term $\Delta\eta_t$. A finding that the predicted exceeds the actual standard deviation implies the presence of excess smoothness. Tables 14 (for the US) and 15 (UK) below enable this comparison. We have the following equations for the growth rate of labour income in the US and the UK respectively,

$$\Delta \log y_t = \underset{(0.345)}{1.858} + \underset{(0.083)}{0.097} \Delta \log y_{t-1} + \varepsilon_t \qquad \begin{array}{l} \sigma_\varepsilon = 3.603 \\ R^2 = 0.0094 \end{array} \tag{3.12}$$

$$\Delta \log y_t = \underset{(0.488)}{2.095} - \underset{(0.078)}{0.0559} \Delta \log y_{t-1} + \varepsilon_t \qquad \begin{array}{l} \sigma_\varepsilon = 5.924 \\ R^2 = 0.0032 \end{array} \tag{3.13}$$

and from equation (3.11) we have

$$\sum_{i=1}^{\infty} \rho^i (E_{t+1} - E_t) \Delta \log y_{t+i} = \sum_{i=1}^{\infty} (-0.097\rho)^i \varepsilon_t = \frac{\varepsilon_t}{1 - 0.097(1 + \mu - r)} \tag{3.14}$$

$$\sum_{i=1}^{\infty} \rho^i (E_{t+1} - E_t) \Delta \log y_{t+i} = \frac{\varepsilon_t}{1 + 0.0559(1 + \mu - r)} \tag{3.15}$$

The sample average rate of growth of labour income in the US, μ, was 2.06% per annum. Hence, the coefficient that multiplies ε_t is 1.106 for a rate of interest of 10% per annum. Given that the standard error of the innovation in labour income from (3.12) is 3.603, we have that the predicted standard deviation for consumption changes is $3.603 \times 1.106 = 3.98$. For the UK, the growth rate of labour income was 2.035% per annum, so that the coefficient that multiplies ε_t is 0.947 for the same rate of interest of 10%. The predicted standard deviation for consumption changes in the UK is $5.92 \times 0.947 = 5.61$. The actual standard deviations for the innovation in consumption are reported in tables 14 and 15.

Table 3.14: Means, Actual and Predicted Standard Deviations for the US

US	Mean	s.d.	Scaling factor, λ
$\Delta \log c_t$	2.200	2.609	
$\frac{\Delta c_t}{y_{t-1}}$	2.481	2.924	
$\Delta \eta_t$	-2.45	3.329	
$\Delta \log c_{nt}$	1.991	1.903	
$\lambda \frac{\Delta c_t}{y_{t-1}}$	2.261	2.159	1.180 (mean)
$\lambda \Delta \eta_t$	-2.195	2.620	1.180
$\lambda \frac{\Delta c_t}{y_{t-1}}$	2.558	2.443	1.335 (OLS)
$\lambda \Delta \eta_t$	-1.618	2.688	1.335

As we can see from table 14, the actual standard deviations of all US measures are smaller than the calculated standard deviation of 3.98. This suggests the existence of excess smoothness in US data.

For the UK, the predicted standard deviation for consumption changes produces interesting results: the predicted innovation is always higher than the actual innovation in the consumption ratio, but it is always less than the equivalent term $\Delta \eta_t$ in (3.10). It appears that there is excess smoothness in consumption for both the US and the UK, but as we can see from tables 14 and 15

Table 3.15: Means, Actual and Predicted Standard Deviations for the UK

UK	Mean	s.d.	Scaling factor, λ
$\Delta \log c_t$	2.092	4.704	
$\frac{\Delta c_t}{y_{t-1}}$	2.463	5.457	
$\Delta \eta_t$	-2.182	7.633	
$\Delta \log c_{nt}$	1.915	3.469	
$\lambda \frac{\Delta c_t}{y_{t-1}}$	2.256	4.038	1.086 (mean)
$\lambda \Delta \eta_t$	-1.953	6.588	1.086
$\lambda \frac{\Delta c_t}{y_{t-1}}$	2.780	4.976	1.338 (OLS)
$\lambda \Delta \eta_t$	-0.919	6.792	1.338

it is less clear that the approximations in equation (3.10) hold, specially for the UK. As we will see in the next sections, we must acknowledge these results when attempting to ascertain the degree of smoothness in consumption from the superior information test devised by Campbell.

We must also note at this point that the equations for the change in the log of labour income for both the UK and the US do not fare better than their level equivalents. For the US in particular, there is again evidence of a structural break in the estimated equation. For the purposes of this section however, we conclude that the simple tests point towards evidence of excess smoothness.

3.4 Superior Information and VARs: Testing for Excess Smoothness

3.4.1 Theory

As we noted when we moved from the linear to the log-linear specification, the tests used in the previous section are subject to an important criticism; *the process for labour income must be modelled correctly.* Those tests assumed that income was correctly specified as a univariate stochastic process by consumers and researchers alike. Closer examination reveals that all the equations explaining the change in labour income, whether in logs or levels, do not report the behaviour of labour income accurately and so it is unlikely that *rational* consumers would choose to use those (exact) specifications.

Throughout the last section it was assumed that the expectations conditional on consumers' *information in both equation (3.1) and in the middle term of equation (3.10) were identical to the expectations conditional on the econometrician's information set that only included current and lagged values of labour income.* Even if the econometrician introduced more variables other than current and past values of labour income it is very likely that the information set used by the consumer will be larger than the econometrician's information set and so it seems unlikely that a researcher would be able to accurately predict the innovations in labour income and hence consumption. West (1988) has demonstrated that if consumers do in fact use a larger information set to forecast future values of labour income than the univariate equation for labour income, *the overall effect is to smooth permanent income.*

Superior information on the agents' behalf poses significant problems to some of the tests attempting to ascertain the merits of REPI. Indeed, in light of this evidence, we may suspect that the calculations made in the previous section may be suggestive of this fact, that consumers are using a larger informational set than that one used by the econometrician. Campbell and Deaton (1989) follow Campbell (1987)'s approach of using the consumer's own behaviour to reveal his or her expectations and therefore finesse the informational problem. The consumer's behaviour is represented by current and past values of consumption and savings. To understand how the superior information problem is overcome, start with the savings equation (3.9)

$$\frac{s_t}{y_t} \approx -\sum_{i=1}^{\infty} \rho^i E_t \left[\Delta \log y_{t+i} \mid I_t\right] - \kappa \tag{3.16}$$

where I_t denotes the consumers' information set. We assume that the econometrician has a smaller information set H_t that consists of all the current and lagged values of the saving ratio and the growth rate of labour income. (In the previous section the information set of the econometrician included only observations on labour income). Campbell and Deaton impose the law of iterated expectations to some of the REPI equations to express them in terms of the smaller information set H_t. Two equations are used

$$\frac{s_t}{y_t} \approx -\sum_{i=1}^{\infty} \rho^i E_t \left[\Delta \log y_{t+i} \mid H_t\right] - \kappa \tag{3.17}$$

and

$$\frac{s_t}{y_t} - \Delta \log y_t - \frac{s_{t-1}}{\rho y_{t-1}} \approx -\sum_{i=0}^{\infty} \rho^i \left[E_t\left(\Delta \log y_{t+i} \mid H_t\right) - E_{t-1}\left(\Delta \log y_{t+i} \mid H_t\right)\right] \tag{3.18}$$

The implications behind these two equations are both simple and powerful. Because s_t/y_t is in H_t, the left hand side of (3.16) which is conditional on the information set I_t must be equal to the left hand side of (3.17) which in turn is conditional to the smaller information set H_t. Moreover, because $\Delta \log y_t$ and $s_{t-1}/\rho y_{t-1}$ are also in H_t, the econometrician is able to observe the left hand side of (3.18) which in turn is equal to the left hand side of (3.10) which is conditional on the information set of the consumer I_t. The crucial assumption that is made throughout this argument is that *the consumer's saving behaviour is observed by the econometrician.* The economic interpretation of these arguments is interesting; if the permanent income hypothesis is true, the agent's saving behaviour reveals the agent's expectations about the behaviour of future labour income. If the present value of expected future changes in labour income is negative (positive), REPI suggests that individuals will save (dissave) in the current period in order to satisfy a basic premise: *agents wish to smooth their lifetime consumption.* REPI can therefore be tested under the smaller information set available to the econometrician with equations (3.17) and (3.18).

The approach taken by Campbell and Deaton is to estimate a VAR system containing the growth rate of labour income and the savings ratio. The intuition behind the VAR is simple; as we have just seen, lagged values of saving ought to explain the behaviour of labour income, whilst allowing consumption, and therefore savings, to depend on past values of income enables an econometrician to test the excess sensitivity hypothesis. The VAR can then be used to obtain forecasts of future income growth used to form the revisions of the right hand side of (3.18). These revisions are compared to the innovations in savings on the left hand side. Excess smoothness will arise in this case if the standard deviation of the left-hand side of (3.18) is less than the standard deviation of the right hand side calculated from the VAR[10].

[10]The use of a VAR framework for labour income and savings was originally advocated by Campbell (1987). The following arguments come from Campbell, pp. 1255-7. The importance of the VAR framework for testing the premise of REPI come from the cointegration literature. If we have a vector with the following variables $x_t = [A_t, y_t, c_t]'$and those variables are cointegrated:

1) a set of restrictions on a VAR on Δx_t cannot be imposed since REPI has implications for consumption and

Consider a basic VAR model stacked into a first order system

$$\begin{bmatrix} \Delta \log y_t - \mu \\ \vdots \\ \Delta \log y_{t-p+1} - \mu \\ s_t/y_t - \sigma \\ \vdots \\ s_{t-p+1}/y_{t-p+1} - \sigma \end{bmatrix} = \begin{bmatrix} a_1 & \cdots & a_p & b_1 & \cdots & b_p \\ 1 & \ddots & & & & \\ & & 1 & & & \\ c_1 & \cdots & c_p & d_1 & \cdots & d_p \\ & & & 1 & \ddots & \\ & & & & & 1 \end{bmatrix} \begin{bmatrix} \Delta \log y_{t-1} - \mu \\ \vdots \\ \Delta \log y_{t-p} - \mu \\ s_{t-1}/y_{t-1} - \sigma \\ \vdots \\ s_{t-p}/y_{t-p} - \sigma \end{bmatrix} + \begin{bmatrix} u_{1t} \\ 0 \\ \vdots \\ u_{2t} \\ 0 \\ \vdots \end{bmatrix} \tag{3.19}$$

where μ is the mean growth of labour income and σ is the mean saving ratio. For the purpose of the analysis, we write this VAR structure more succinctly in matrix notation as

$$z_t = Az_{t-1} + u_t \tag{3.20}$$

where A is the companion matrix for the VAR. For all i,

$$E\left[z_{t+i} \mid H_t\right] = A^i z_t \tag{3.21}$$

and

$$E\left[z_{t+i} \mid H_t\right] - E\left[z_{t+i} \mid H_{t-1}\right] = A^i z_t - A^{i+1} z_{t-1} = A^i u_t \tag{3.22}$$

Defining the selection vectors e_1' and e_2' such that $e_1' z_t = \Delta \log y_t$ and $e_2' z_t = s_t/y_t$ we can write the left hand side of equation (3.18) as

$$\frac{s_t}{y_t} - \Delta \log y_t - \frac{s_{t-1}}{\rho y_{t-1}} = \left(e_2' - e_1'\right) z_t - \rho^{-1} e_2' z_{t-1}$$

its change,

2) no finite VAR representation exists for Δx_t.

A solution to the second point is to use an error-correction model for the vector x_t. It is possible to re-write an error-correction model in VAR form by dropping one of the elements of Δx_t in the error correction system and replacing it with $\alpha' x_t$ where α denotes the cointegrating vector. This is where the VAR system that involves savings and the innovation in labour income comes from. In the Appendix we report the cointegrating results and show that savings are stationary.

and substitute (3.20) to obtain,

$$\frac{s_t}{y_t} - \Delta \log y_t - \frac{s_{t-1}}{\rho y_{t-1}} = \left[\left(e_2' - e_1'\right) A - \rho^{-1} e_2'\right] z_{t-1} + \left(e_2' - e_1'\right) u_t \tag{3.23}$$

In addition we can express the right hand side of (3.18) in terms of the VAR

$$\sum_{i=0}^{\infty} \rho^i \left[E_t \left(\Delta \log y_{t+i} \mid H_t\right) - E_{t-1} \left(\Delta \log y_{t+i} \mid H_t\right)\right] = \sum_{i=0}^{\infty} e_1' \left(\rho A\right)^i u_t \tag{3.24}$$

where we have made use of the result in (3.22). From equations (3.23) and (3.24) it is possible to compare the actual with the theoretically predicted innovation standard deviation of $(s_t/y_t - \Delta \log y_t)$ which is approximately equivalent to the theoretical innovation of $\Delta c_t/y_t$. The predicted innovation standard deviation of the right hand side of (3.18) is equal to

$$\sqrt{e_1' \left(I - \rho A\right)^{-1} \Omega \left(I - \rho A\right)^{-1} e_1} \tag{3.25}$$

where Ω is the variance covariance matrix obtained from the unrestricted $VAR(p)$ equation. The actual innovation in $(s_t/y_t - \Delta \log y_t)$ is simply $u_{2t} - u_{1t}$ and so the actual innovation standard deviation is equal to the square root of

$$\left(e_2' - e_1'\right) \Omega \left(e_2 - e_1\right) \tag{3.26}$$

Excess smoothness is observed in the data if the predicted innovation warranted by (3.25) is larger than the actual innovation warranted by (3.26).

Before reporting the degree of smoothness in consumption for the US and UK, it is interesting to note that REPI imposes testable restrictions to the coefficients of the companion matrix A. According to REPI, the left hand side of (3.23) ought to be equal to the left hand side of (3.24). This requires that two sets of constraints be satisfied, namely,

$$\left(e_2' - e_1'\right) A - \rho^{-1} e_2' = 0 \tag{3.27}$$

$$-\sum_{i=0}^{\infty} e_1' (\rho A)^i = e_2' - e_1' \tag{3.28}$$

The first of the restrictions claims that no lagged elements of labour income growth or the savings ratio should affect the ratio of consumption changes to income. The test for 'excess sensitivity' can be considered a special case of this restriction. The second restriction represents a test for excess smoothness; it suggests that the innovations in consumption ought to be equal to the change in the present value of labour income. Thus, if this restriction is satisfied then we cannot report that consumption is smoother than implied by REPI. Note further that if $| \rho A |< 1$, we can write $\sum_{i=0}^{\infty} e_1' (\rho A)^i = (I - \rho A)^{-1}$ in (3.28). Provided that the inverse of $(I - \rho A)$ exists, the first constraint (3.27) can be rearranged to produce

$$-e_1' (I - \rho A)^{-1} = e_2' - e_1'$$

which is identical to the no 'excess smoothness' constraint. The 'excess sensitivity' constraint imposes the following restrictions on the coefficients of the companion matrix A; $a_1 = c_1, ..., a_p = c_p, d_1 = b_1 + \rho^{-1}, d_2 = b_2, ..., d_p = b_p$. In the results that follow we test these restrictions with both Wald and LR tests.

The condition that the inverse term, $(I - \rho A)^{-1}$, exists together with the 'excess sensitivity' constraint, provides another interesting insight into the permanent income hypothesis. Using a $VAR(1)$ to illustrate this point, the restricted companion matrix is

$$A = \begin{pmatrix} a_1 & b_1 \\ a_1 & b_1 + \rho^{-1} \end{pmatrix}$$

so that

$$(I - \rho A) = \begin{pmatrix} 1 - \rho a_1 & -\rho b_1 \\ -\rho a_1 & -\rho b_1 \end{pmatrix}$$

Consequently, the determinant of the matrix can only be non-zero if ρ and b_1 are not zero. This imposes the condition that there must be Granger-causality from savings to the log of income. *Such Granger-causality is guaranteed if consumers have superior information and so it is worth* *testing for it since detection of its absence suggests that the bivariate system of labour income and savings is not able to overcome the superior information problem.*

3.4.2 Results

The first step that we must undertake to test whether REPI explains the behaviour of US and UK consumers is to estimate unrestricted VARs for the innovation in labour income and the level of savings for both countries[11]. The diagnostic selection criteria used to select the lag length of the VARs in this study is the Akaike Information Criteria although we also examine the Schwartz Bayesian Criterion together with the significance of lag lengths reported by t and likelihood ratio test statistics. For both the US and the UK, and for the purposes of comparison, we also provide VARs with the lag lengths reported in previous studies.

For the US, the data favoured a $VAR(2)$ system and there was also evidence that some of the coefficients in a $VAR(5)$ were significant. We therefore report results for a $VAR(2)$ and a $VAR(5)$ together with the $VAR(1)$ specification that some authors (Campbell, Campbell and Deaton and Flavin) had used previously. For the UK, there was evidence that a $VAR(3)$ system would encapsulate the characteristics of the data best. We also report results for a $VAR(1)$ and a $VAR(4)$ to compare the results with Attfield, Demery and Duck (1990). The next set of tables examine the degree of excess smoothness in consumption for $VAR(1)$, $VAR(2)$ and $VAR(5)$ systems for the US, and $VAR(1)$, $VAR(3)$ and $VAR(4)$ systems for the UK. These tables include Wald tests for the restrictions (3.27) and (3.28) implied by REPI on the companion matrix, the predicted innovation standard deviation and the actual innovation standard deviation. In the final column the ratio of the predicted to the actual innovation standard deviation is reported; a ratio of less than one implying excess smoothness. The standard error associated with this ratio is computed as[12] $\sqrt{D'\theta D}$ where D is the vector of derivatives of the standard deviation ratio with respect to the VAR parameters, and θ is the variance-covariance matrix of the parameters. D was computed using numerical derivatives in Gauss[13]. Tables in the appendix show the different values taken by all the coefficients under the different estimated VAR specifications for both the

[11] As we shall see in the results that follow, selection of the appropriate lag order can be an important part of the story since it can affect results obtained from imposing the appropriate restrictions with Wald and LR tests.

[12] See Campbell and Deaton, pp. 366, footnote under table 3.

[13] I am grateful to W.C. Lau for helping me with the Gauss code for the calculation of D.

Table 3.16: Tests for Excess Smoothness for the US

	WaldTest (*p − value*)	*Predicted* *Innovation s.d.*	*Actual* *Innovation s.d.*	*Ratio* (*s.e.*)
Total Consumption				
VAR(1)	11.368 (0.034)	3.483	3.193	0.917 (0.098)
VAR(2)	18.675 (0.0009)	4.352	3.123	0.718 (0.105)
VAR(5)	30.017 (0.008)	4.13	2.982	0.722 (0.158)
Non-durable($\lambda = 1.18$)				
VAR(1)	8.828 (0.012)	3.416	2.540	0.744 (0.153)
VAR(2)	19.574 (0.0006)	4.287	2.449	0.571 (0.131)
VAR(5)	25.572 (0.0041)	4.183	2.367	0.566 (0.148)
Non-durable($\lambda = 1.33$)				
VAR(1)	9.180 (0.012)	3.243	2.603	0.803 (0.151)
VAR(2)	17.323 (0.00167)	4.096	2.529	0.617 (0.134)
VAR(5)	24.804 (0.0057)	4.019	2.435	0.606 (0.150)

US and the UK. The appendix also shows the results of the Granger Causality test mentioned above. As it can be seen, the weakest implications of the REPI are satisfied: in all instances the null that there is no Granger Causality from savings to income is easily rejected for both the UK and the US and the coefficient on saving in the $VAR(1)$ equation explaining labour income change (or the sum of coefficients in the higher order cases[14]) is negative as one would expect[15].

The results from table 16 are consistent with previously reported findings for the US economy. The Wald test statistics continue to reject the coefficient restrictions imposed on the companion matrix by REPI and it appears that the failure of the restrictions can still be ex-

[14] Only the VAR(4) case for UK total consumption has a sum that is positive.

[15] In tables 16 and 17, the measure of savings in the first column of each table is defined as the difference between disposable income and total consumption, including purchases of durables. The second column uses the difference between disposable income and the consumption of nondurables inflated by a factor that represents the ratio of the mean of total consumption to the mean of consumption excluding durables. The third column is the same as the second column but with a scaling factor which is the reciprocal of the marginal propensity to consume estimated from a simple bivariate regression of consumption of nondurables and services on income. (See Campbell and Deaton's Tables II and III, pp. 364 and 366 respectively for more).

Table 3.17: Tests for Excess Smoothness for the UK

	WaldTest (*p* − *value*)	*Predicted Innovation s.d.*	*Actual Innovation s.d.*	*Ratio* (*s.e.*)
Total Consumption				
VAR(1)	6.845 (0.033)	5.931	7.459	1.258 (0.166)
VAR(3)	33.460 (0.0001)	6.459	6.949	1.076 (0.258)
VAR(4)	32.802 (0.00007)	6.675	6.942	1.040 (0.261)
Non-durable($\lambda = 1.08$)				
VAR(1)	5.255 (0.072)	5.277	6.483	1.228 (0.127)
VAR(3)	35.843 (0.0000)	5.355	5.972	1.115 (0.175)
VAR(4)	35.118 (0.0003)	5.768	5.973	1.036 (0.182)
Non-durable($\lambda = 1.33$)				
VAR(1)	7.956 (0.019)	5.121	6.631	1.295 (0.145)
VAR(3)	42.399 (0.0000)	5.184	6.057	1.168 (0.185)
VAR(4)	41.685 (0.0000)	5.467	6.055	1.107 (0.194)

plained in all cases, irrespective of the consumption measure used, by the finding of excess smoothness as the ratio of the actual to the predicted standard deviations reported in the last column is significantly less than one.

For the UK (table 17), the results are somewhat different to those for the US. As in the US, only in once instance (for the $VAR(1)$ case of nondurable consumption scaled by the ratio of the means) does the Wald test fail to reject the restrictions imposed by REPI. In terms of the finding of excess smoothness, the results differ substantially to those of the US, the ratio being always greater than one. It appears that the restrictions on the companion matrix fail not because consumption is too smooth but because the predictions of the permanent income model are less volatile than the actual innovations in consumption[16].

[16]It is important to note at this point that the results are somewhat dependent of the approximations used for the derivation of equation (3.18).

3.5 Testing for Excess Smoothness: Flavin's Approach

3.5.1 Theory

As we have seen above, Campbell and Deaton used projection arguments to infer the variance of revisions in permanent income from a bivariate autoregression of income and savings, *under the null hypothesis of the truth of the REPI.* Flavin generalises this analysis to show that Campbell and Deaton's 'algorithm provides a consistent estimate of the variance of revisions in permanent income if the consumption data is generated by an alternative to the PIH- the excess sensitivity hypothesis.' Flavin (1993), pp. 653.

The alternative hypothesis that Flavin works with is that consumption is *too sensitive to current income*

$$c_t = \beta y_t^T + y_t^p$$

where $0 < \beta < 1$ is the marginal propensity to consume out of transitory income and where transitory income is defined as the residual $y_t^T \equiv y_t + \left(\frac{r}{1+r}\right) A_t - y_t^p$. Permanent income is defined in the standard Flavin (1981) form. The change in consumption is

$$\Delta c_t = \beta \Delta y_t + (1-\beta)\left(\frac{r}{1+r}\right)\sum_{\tau=0}^{\infty}\left(\frac{1}{1+r}\right)^{\tau}(E_t - E_{t-1})\, y_{t+\tau} \tag{3.29}$$

and the 'savings for a rainy day' equation is now scaled by $(1-\beta)$, viz.

$$s_t = (1-\beta)\left[y_t - \left(\frac{r}{1+r}\right)\sum_{\tau=0}^{\infty}\left(\frac{1}{1+r}\right)^{\tau} E_t y_{t+\tau}\right] \tag{3.30}$$

We use the VAR approach to derive the restrictions on the companion matrix *under the null hypothesis of excess sensitivity.* We transform Flavin's framework into logs to compare her results with those of Campbell and Deaton and to test whether the null hypothesis of excess sensitivity holds for the revised data sets. The logarithmic form of equation (3.9) is

$$\frac{s_t}{y_t} \approx -(1-\beta)\sum_{i=1}^{\infty}\rho^i E_t \Delta \log y_{t+i} - \varpi$$

where ϖ is a constant different from Campbell and Deaton's. The equivalent form of (3.10) is given as

$$\frac{s_t}{y_t} - (1-\beta)\Delta\log y_t - \frac{s_{t-1}}{\rho y_{t-1}} \approx -(1-\beta)\sum_{i=1}^{\infty}\rho^i\left(E_{t+1}-E_t\right)\Delta\log y_{t+i} \approx -\frac{\Delta c_t}{y_{t-1}} \tag{3.31}$$

Note that we are now scaling by the factor $(1-\beta)$. This framework is still able to overcome the superior information problem[17]. The equivalent restrictions under the null of excess sensitivity for the companion matrix A in equation (3.20) are

$$\left(e_2^{'} - (1-\beta)\, e_1^{'}\right) A - \rho^{-1} e_2^{'} = 0 \tag{3.32}$$

$$-(1-\beta)\sum_{i=0}^{\infty} e_1^{'}\,(\rho A)^i = e_2^{'} - (1-\beta)\, e_1^{'} \tag{3.33}$$

These constraints now impose the following restrictions on the coefficients of the companion matrix A; $a_1 = (1-\beta)\, c_1, ..., a_p = (1-\beta)\, c_p, d_1 = (1-\beta)\, b_1 + \rho^{-1}, d_2 = (1-\beta)\, b_2, ..., d_p = (1-\beta)\, b_p$.

3.5.2 Orthogonality and Smoothness Tests

In this section we summarize Flavin's (1993) discussion about the parameter restrictions implied by the orthogonality and smoothness conditions on the bivariate autoregression of labour income and savings (see pp. 658-9 for more details). We explain the why the orthogonality condition implies the smoothness condition but the reverse is not true unless the bivariate autoregression of labour income and savings is of order one.

Consider the general bivariate autoregression of labour income and savings

$$\begin{bmatrix} y_t \\ s_t \end{bmatrix} = \begin{bmatrix} a(L) & b(L) \\ c(L) & d(L) \end{bmatrix} \begin{bmatrix} y_t \\ s_t \end{bmatrix} + \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{st} \end{bmatrix} \tag{3.34}$$

where ε_{yt} and ε_{st} denote the innovations in labour income and savings respectively and $a(L)$,

[17] See the appendix in the next chapter for an explanation.

$b(L)$, $c(L)$ and $d(L)$ are polynomials in the lag operator. Represent this expression as

$$z_t = A(L)z_t + u_t \tag{3.35}$$

where $z_t = [y_t, s_t]$. If the PIH holds, such that $\Delta c_t = \Delta y_t^p = \varepsilon_{ypt}$, then the accounting identity

$$\Delta c_t \equiv s_t + (1+r)\, s_{t-1} + \Delta y_t$$

can be written as

$$s_t + (1+r)\, s_{t-1} + \Delta y_t = \varepsilon_{ypt} \tag{3.36}$$

Equation (3.36) places testable restrictions on (3.34). If the PIH hypothesis holds, augmenting (3.34) by the consumption equation imposes the following structure on the reduced-form (VAR) representation of y_t, s_t, Δc_t

$$\begin{bmatrix} y_t \\ s_t \\ \Delta c_t \end{bmatrix} = \begin{bmatrix} a(L) & b(L) \\ a(L) - L & b(L) + (1+r)L \\ 0 & 0 \end{bmatrix} \begin{bmatrix} y_t \\ s_t \end{bmatrix} + \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{yt} - \varepsilon_{ypt} \\ \varepsilon_{ypt} \end{bmatrix} \tag{3.37}$$

where

$$\varepsilon_{ypt} = \left(\frac{r}{1+r}\right) \sum_{i=0}^{\infty} (1+r)^{-i} \left(E_t - E_{t-1}\right) y_{t+i}$$

and $(1+r)^{-i} = \delta$. Note the crucial implication of (3.37): given 'the exact linear dependence among the three equations in system (3.37), a test of the restrictions on any two of the three equations will yield numerically identical values of the likelihood-ratio statistic. Further, since the form of the parameter restrictions on the consumption equation is the orthogonality of Δc_t with respect to lagged y_t and lagged s_t, no restrictions are imposed on the parameters of the companion equation (either the y_t or s_t equation) in any paired comparison in which Δc_t is one of the two variables in the system. Thus testing the cross-equation parameter restrictions on the bivariate autoregression (equation (3.34)), will yield exactly the same value of the likelihood-ratio statistics as testing the orthogonality condition in a regression of Δc_t on lagged y_t and lagged s_t [...]' (pp. 658)

To see how the orthogonality restriction is more general than the smoothness restriction consider the expression for the innovation in permanent income expressed in terms of the VAR of labour income and savings

$$\widetilde{\varepsilon}_{ypt} = r\delta\,[1,0]\,[1 - A(\delta)]^{-1} \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{yt} - \varepsilon_{ypt} \end{bmatrix} \tag{3.38}$$

where $\widetilde{\varepsilon}_{ypt}$ is the econometrician's inference on ε_{ypt} which comes from the estimation of the VAR for labour income and savings and $A(\delta)$ is the 2×2 matrix in equation (3.35) evaluated at $L = \delta = (1+r)^{-1}$. If the parameters of $A(L)$ satisfy the PIH restrictions (as stated in (3.37)), then (3.38) reduces to

$$\widetilde{\varepsilon}_{ypt} = r\delta\,[1,0]\,[1 - A(\delta)]^{-1} \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{yt} - \varepsilon_{ypt} \end{bmatrix} = [1,-1] \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{yt} - \varepsilon_{ypt} \end{bmatrix} = \varepsilon_{ypt}$$

which is of course Campbell and Campbell and Deaton's original contribution to the literature: if the PIH holds then the econometrician is able to estimate the innovation of permanent income from a VAR of the innovation in labour income and savings (this is equivalent to restriction (3.28) above which is the smoothness restriction on the VAR). Note that the smoothness restrictions are (in terms of the individual parameters of $A(L)$)

$$\begin{aligned} c(\delta) &= a(\delta) - \delta \\ d(\delta) &= b(\delta) + (1+r)\delta \end{aligned} \tag{3.39}$$

Further, given (3.36) it must be true that

$$\varepsilon_{ct} = [1,-1] \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{yt} - \varepsilon_{ypt} \end{bmatrix}$$

thus 'if the parameter restrictions imposed by the PIH in $A(L)$ hold, $\varepsilon_{ct} = \varepsilon_{ypt} = \widetilde{\varepsilon}_{ypt}$.' It follows trivially from this that if the PIH holds then the variance of consumption should be equal to the variance of the innovation in permanent income and further that the econometrician will be able to accurately infer the variance of the innovation in permanent income from the VAR of labour income and savings.

It remains to note that the smoothness restrictions (3.39) is a special case of the orthogonality restriction imposed in (3.37) since the restrictions

$$\begin{aligned} c(L) &= a(L) - L \\ d(L) &= b(L) + (1+r)L \end{aligned} \tag{3.40}$$

are the smoothness restrictions if and only if L is evalued at δ.

3.5.3 Results

We test the orthogonality and sensitivity restrictions using LR tests. Again, we estimate the unrestricted VAR system, impose the restrictions (3.32) and (3.33) and test them. From the LR tests estimates of the excess sensitivity parameter, β, are generated when the restricted model is estimated. It is interesting that for a paper published in 1993, Flavin decided to use the Blinder and Deaton data set that ends in 1984[18].

Table 18 reaffirms the findings of excess smoothness obtained in the previous section for the US. These findings are also consistent with Flavin's research although, *unlike Flavin we do reject the excess sensitivity restrictions for all the consumption measures to conclude that the null of excess sensitivity does not hold.* Moreover, from the standard errors reported in the last column there is little evidence in favour of the significance of the excess sensitivity parameter.

The equivalent LR tests for UK data shown in table 19 suggest that there is excess smoothness. The excess sensitivity results are more puzzling however, the LR test cannot reject the null of excess sensitivity, suggesting that this alternative hypothesis explains the behaviour of consumption in the UK better than REPI. However, closer examination reveals that the excess sensitivity parameters are always negative - implying a negative marginal propensity to consume out of transitory income - and are mostly statistically insignificant. The findings of the rejection of the null of excess sensitivity along with a negative and insignificant marginal

[18]The savings measures used in the next two tables are the same as those used in tables 3.16 and 3.17.

Table 3.18: LR Tests for Orthogonality and Sensitivity for the US

	Orthogonality	Sensitivity	*value of* β (*s.e.*)
Total Consumption			
VAR(1)	$\chi^2(2)$ =11.542	$\chi^2(1)$ =5.413	−4.436 (9.51)
VAR(2)	$\chi^2(4)$ =18.456	$\chi^2(3)$ =15.608	0.881 (0.557)
VAR(5)	$\chi^2(10)$ =29.489	$\chi^2(9)$ =22.310	1.306 (0.662)
Non-durable($\lambda = 1.18$)			
VAR(1)	$\chi^2(2)$ =8.521	$\chi^2(1)$ =7.747	0.253 (0.268)
VAR(2)	$\chi^2(4)$ =18.351	$\chi^2(3)$ =9.923	0.580 (0.205)
VAR(5)	$\chi^2(10)$ =24.676	$\chi^2(9)$ =20.526	0.395 (0.181)
Non-durable($\lambda = 1.33$)			
VAR(1)	$\chi^2(2)$ =9.730	$\chi^2(1)$ =8.926	0.152 (0.289)
VAR(2)	$\chi^2(4)$ =16.623	$\chi^2(3)$ =10.947	0.493 (0.223)
VAR(5)	$\chi^2(10)$ =24.201	$\chi^2(9)$ =21.637	0.3185 (0.204)

propensity to consume out of transitory income is a puzzle for further research beyond the present work.

3.6 Are the Above Results Sensitive to the Logarithmic Approximations?

The results in the previous two sections may be sensitive to the logarithmic approximations used by Campbell and Deaton and Flavin[19]. In this section we examine the permanent income model in levels and leave aside Campbell and Deaton's claim that a logarithmic approximation may explain the data better. By reverting to a model in levels it maybe possible to determine whether the previously reported failures of REPI and the excess sensitivity hypothesis were the result of the logarithmic approximations.

[19]Recall, from tables 14 and 15, that the left hand-side of equation (3.10) was not approximately equal to the ratio of the change in consumption over lagged labour income. That approximation was crucial to the tests used in the last two sections and may be accountable for some of the results encountered thus far.

Table 3.19: LR Tests for Orthogonality and Sensitivity for the UK

	Orthogonality	Sensitivity	*value of* β (*s.e.*)
Total Consumption			
VAR(1)	$\chi^2(2)$ =6.845	$\chi^2(1)$ =0.452	−1.152 (0.708)
VAR(3)	$\chi^2(4)$ =31.993	$\chi^2(3)$ =7.347	−2.569 (1.336)
VAR(4)	$\chi^2(8)$ =31.702	$\chi^2(7)$ =8.876	−2.287 (1.151)
Non-durable($\lambda = 1.08$)			
VAR(1)	$\chi^2(2)$ =5.32	$\chi^2(1)$ =0.205	−1.434 (1.191)
VAR(3)	$\chi^2(6)$ = 35.852	$\chi^2(5)$ =5.208	−3.098 1.939
VAR(4)	$\chi^2(8)$ =35.71	$\chi^2(7)$ =7.593	−2.646 (1.497)
Non-durable($\lambda = 1.33$)			
VAR(1)	$\chi^2(2)$ =7.374	$\chi^2(1)$ =0.665	−1.552 (1.028)
VAR(3)	$\chi^2(6)$ =40.777	$\chi^2(5)$ =6.913	−2.768 (1.346)
VAR(4)	$\chi^2(8)$ =40.688	$\chi^2(7)$ =8.795	−2.448 (1.129)

3.6.1 Likelihood Ratio Tests for the Permanent Income and Excess Sensitivity Hypotheses

From the definition of savings as the difference between disposable income and consumption, $s_t = y_t + \frac{r}{1+r}A_t - c_t$, and equation (3.1) we find a statement about savings

$$s_t = -\sum_{i=0}^{\infty} \left(\frac{1}{1+r}\right)^i E_t \Delta y_{t+i+1} \tag{3.41}$$

If (3.41) is true, the restrictions derived by Campbell and Deaton to be imposed on the coefficients of the companion matrix A will still be a valid way to test the truth of the permanent income hypothesis. Take a VAR in levels

$$\begin{bmatrix} \Delta y_t \\ \vdots \\ \Delta y_{t-p+1} \\ s_t \\ \vdots \\ s_{t-p+1} \end{bmatrix} = \begin{bmatrix} a_1 & \cdots & a_p & b_1 & \cdots & b_p \\ 1 & \ddots & & & & \\ & & 1 & & & \\ c_1 & \cdots & c_p & d_1 & \cdots & d_p \\ & & & 1 & \ddots & \\ & & & & & 1 \end{bmatrix} \begin{bmatrix} \Delta y_{t-1} \\ \vdots \\ \Delta y_{t-p} \\ s_{t-1} \\ \vdots \\ s_{t-p} \end{bmatrix} + \begin{bmatrix} u_{1t} \\ 0 \\ \vdots \\ u_{2t} \\ 0 \\ \vdots \end{bmatrix}$$

and write it in matrix form as

$$z_t = Az_{t-1} + u_t$$

we can express the savings equation (3.41) in terms of the VAR as

$$-s_t = -e_2' z_t = \sum_{i=1}^{\infty} \rho^i E_t \Delta y_{t+k} = \sum_{i=1}^{\infty} e_1' \rho^i A^i z_t$$

where e_1 and e_2 are the same selection vectors as those defined above, $\rho = (1+r)^{-1}$and we have made use of the results in (3.21). Since this equation for savings must be true for all values of the vector z, the following restrictions can be imposed to the companion matrix A to test REPI

$$-e_2' = \sum_{i=1}^{\infty} e_1' \rho^i A^i \tag{3.42}$$

and these are the same restrictions as those imposed by equation (3.28). The equivalent excess sensitivity hypothesis restriction is given by:

$$-e_2' = (1-\beta) \sum_{i=1}^{\infty} e_1' \rho^i A^i \tag{3.43}$$

Hence, provided (3.41) and (3.30) are correct definitions of the behaviour of savings under REPI or the excess sensitivity hypothesis respectively, we can overcome the informational problem and test both hypotheses.

Table 3.20: LR Tests for Orthogonality and Sensitivity for the US

	Orthogonality	Sensitivity	*value of β* (*s.e.*)
Total Consumption			
VAR(1)	$\chi^2(2)$ =117.9234	$\chi^2(1)$ =2.218	−0.62078 (0.37332)
VAR(2)	$\chi^2(4)$ =12.3884	$\chi^2(3)$ =11.0436	−0.25852 (0.25944)
VAR(5)	$\chi^2(10)$ =23.5714	$\chi^2(9)$ =22.1958	−0.33438 (0.31231)
Non-durable($\lambda = 1.33$)			
VAR(1)	$\chi^2(2)$ =106.398	$\chi^2(1)$ =2.225	−0.48393 (0.25846)
VAR(2)	$\chi^2(4)$ =54.6484	$\chi^2(3)$ =16.6592	−0.43222 (0.35448)
VAR(5)	$\chi^2(10)$ =30.8536	$\chi^2(9)$ =23.5026	−0.83257 (0.54582)

The same techniques used previously to test for the truth of the permanent income and the excess sensitivity hypotheses, are used in this section[20,21]. First, the lag-order for the unrestricted *VAR* system is calculated using the same criteria that were used previously. Again, the data favoured the $VAR(2)$ specification for the US and the $VAR(3)$ specification for the UK. Other *VAR* specifications are reported to compare with previously published results by other authors. The second step involved testing for Granger causality from savings to the change in labour income. Again, the weakest implications of the REPI are satisfied since the null of no Granger causality from savings to income is easily rejected for both the US and the UK and the coefficient on savings in the $VAR(1)$ equation explaining labour income changes (or the sum of coefficients in the higher order cases) is negative. Finally, a likelihood ratio test is used to test restrictions (3.42) and (3.43) on the companion matrix A. The results of this last step are reported in tables 20 and 21.

The findings reported in these tables are similar to those found for the logarithmic versions, namely the failure of the permanent income and excess sensitivity hypotheses. Whilst the test statistics are kinder to the excess sensitivity hypothesis, that hypothesis is ruled out on the

[20]The definitions of savings used in this section are similar to the ones employed by Campbell when he tested for permanent income using data in levels and therefore differ somewhat from the definitions of savings used above which were similar to those employed by Campbell and Deaton.

[21]The results of all the calculations and Granger tests on the *VARs* are given in the appendix.

Table 3.21: LR Tests for Orthogonality and Sensitivity for the UK

	Orthogonality	Sensitivity	*value of* β (*s.e.*)
Total Consumption			
VAR(1)	$\chi^2(2)$ =8.584	$\chi^2(1)$ =1.706	−1.067 (0.593)
VAR(3)	$\chi^2(6)$ =28.644	$\chi^2(5)$ =9.496	−2.181 (1.058)
VAR(4)	$\chi^2(8)$ =30.146	$\chi^2(7)$ =12.154	−2.138 (1.086)
Non-durable($\lambda = 1.33$)			
VAR(1)	$\chi^2(2)$ =9.486	$\chi^2(1)$ =2.348	−1.021 (0.576)
VAR(3)	$\chi^2(6)$ =45.468	$\chi^2(5)$ =10.384	−2.172 (1.129)
VAR(4)	$\chi^2(8)$ =31.172	$\chi^2(7)$ =12.112	−2.023 (1.054)

account that the excess sensitivity coefficients continue to be negative and statistically insignificant in most cases. Since these results are not that different from those results reported by the logarithmic versions, we tentatively conclude that the logarithmic approximations alone cannot be blamed for the failure of the permanent income and excess sensitivity hypotheses. Thus, there is strong evidence that these two hypotheses are not able to explain the characteristics of US and UK consumption behaviour accurately.

3.7 Are the Results Sensitive to the Data Sets Used? A Comparison of the Blinder and Deaton and Attfield, Demery and Duck Data Sets with the Revised Data Sets

The failures of the permanent income and excess sensitivity hypotheses encountered so far in this chapter may be the result of the data revisions made by the ONS and specially the NIPA. To examine whether the revisions made to these data sets can be considered the driving force behind the results reported in the previous sections we compare the revised data sets to the original Blinder and Deaton and Attfield et al.'s data sets. To do this, the same tests are imposed on all the data sets for equivalent periods, 1959:3 to 1984:4 for the US and 1955:1 to 1987:2 for the UK. We report the results of these tests for both the level and logarithmic

Table 3.22: Tests for Excess Smoothness; Blinder and Deaton Data

	WaldTest *(p − value)*	*Predicted* *Innovation s.d.*	*Actual* *Innovation s.d.*	*Ratio* *(s.e.)*
Total Consumption				
VAR(1)	12.029 (0.002)	4.807	3.382	0.703 (0.105)
VAR(2)	11.564 (0.021)	4.986	3.388	0.679 (0.127)
VAR(5)	30.747 (0.0006)	4.699	3.063	0.652 (0.154)
Non-durable($\lambda = 1.28$)				
VAR(1)	12.496 (0.0019)	3.919	2.317	0.591 (0.205)
VAR(2)	15.417 (0.0039)	4.713	2.259	0.479 (0.209)
VAR(5)	32.600 (0.0003)	5.708	2.046	0.358 (0.152)
Non-durable($\lambda = 1.52$)				
VAR(1)	5.613 (0.063)	3.143	2.439	0.776 (0.165)
VAR(2)	7.253 (0.123)	3.679	2.391	0.649 (0.175)
VAR(5)	17.866 (0.057)	4.835	2.229	0.461 (0.154)

approximations.

3.7.1 The US Data

Logarithmic Approximations

As we can see from tables 22-5, these tests show that the revisions made to the NIPA cannot be considered the driving force of the results that were reported in the previous sections so that the failures of the permanent income and the excess sensitivity hypotheses are likely to be explained by the extra observations in the 1980s and 1990s. Apart from the last measurement of non-durable consumption (the Blinder and Deaton data set accepts the permanent income hypothesis whereas the revised data set does not), there are no significant differences in the results obtained from the two data sets. Both tend to reject the permanent income hypothesis and *fail to reject* the excess sensitivity theory of Flavin. These findings are also consistent with Campbell and Deaton's and Flavin's for the US that used the original (and larger) Blinder and Deaton data set running from 1953 to 1984 and suggest that testing for REPI can be quite

Table 3.23: Tests for Excess Smoothness; Revised Data

	WaldTest (*p − value*)	*Predicted Innovation s.d.*	*Actual Innovation s.d.*	*Ratio* (*s.e.*)
Total Consumption				
VAR(1)	8.910 (0.001)	4.552	3.243	0.712 (0.135)
VAR(2)	14.942 (0.005)	5.158	3.157	0.612 (0.140)
VAR(5)	28.266 (0.002)	4.054	2.922	0.721 (0.154)
Non-durable($\lambda = 1.16$)				
VAR(1)	9.493 (0.009)	4.472	2.567	0.574 (0.203)
VAR(2)	17.179 (0.002)	5.721	2.468	0.431 (0.197)
VAR(5)	26.601 (0.003)	5.446	2.311	0.424 (0.181)
Non-durable($\lambda = 1.37$)				
VAR(1)	5.896 (0.052)	3.884	2.695	0.694 (0.198)
VAR(2)	11.9712 (0.018)	4.997	2.608	0.522 (0.193)
VAR(5)	20.835 (0.022)	4.969	2.448	0.493 (0.172)

sensitive to the data period and the definition of consumption chosen[22].

Levels

Tables 26-27 compare the Blinder and Deaton to the revised data sets using Campbell's permanent income test for data in levels. We do not report excess sensitivity tests as Flavin did not report such tests for data in levels[23].

The results in tables 26 and 27 are similar to the ones reported for the logarithmic approximations; the $VAR(1)$ and $VAR(5)$ cases used by Campbell continue to reject REPI regardless of whether the Blinder and Deaton or the revised data sets are used.

[22]Note that the Blinder and Deaton data set does fail to reject the permanent income hypothesis for one definition of non-durable consumption and that the excess sensitivity theory is rejected for total consumption, a finding consistent with Flavin's original study.

[23]The original Blinder and Deaton data set from 1953:2 to 1984:4 in levels does in fact reject the excess sensitivity hypothesis. The shorter Blinder and Deaton data set and the revised data sets continue to reject Flavin's hypothesis in levels.

Table 3.24: LR Tests for Orthogonality and Sensitivity; Blinder and Deaton Data

	Orthogonality	Sensitivity	*value of* β (*s.e.*)
Total Consumption			
VAR(1)	$\chi^2(2)$ =11.696	$\chi^2(1)$ =7.308	0.521 (0.193)
VAR(2)	$\chi^2(4)$ =11.478	$\chi^2(3)$ =7.172	0.514 (0.186)
VAR(5)	$\chi^2(10)$ =29.38	$\chi^2(9)$ =18.394	0.75 (0.172)
Non-durable($\lambda = 1.28$)			
VAR(1)	$\chi^2(2)$ =13.571	$\chi^2(1)$ =0.221	0.473 (0.121)
VAR(2)	$\chi^2(4)$ =13.892	$\chi^2(3)$ =0.875	0.488 (0.119)
VAR(5)	$\chi^2(10)$ =28.830	$\chi^2(9)$ =6.555	0.504 (0.102)
Non-durable($\lambda = 1.52$)			
VAR(1)	$\chi^2(2)$ =5.561	$\chi^2(1)$ =0.878	0.316 (0.142)
VAR(2)	$\chi^2(4)$ =6.784	$\chi^2(3)$ =1.252	0.332 (0.139)
VAR(5)	$\chi^2(10)$ =16.910	$\chi^2(9)$ =6.370	0.363 (0.113)

3.7.2 The UK data

In this section we compare the results of Attfield, Demery and Duck for their data period from 1955:1 to 1987:2 against the new data set for the same period. We only report the results for their $VAR(4)$ specification for both logs and levels.

As we can see from table 28, the extended data set cannot be accountable for the failures of the permanent income hypothesis.

3.8 Conclusions: What Have We Learned About the Behaviour of Agents with Ten More Years of Available Information?

At this point, we compare the results that we have obtained from the two extended data sets with those reported by other authors with older data sets. The evidence provided by the further ten years of available observations lead to three main conclusions:

1. *Granger causality* is still present in our data. This suggests that the problem of inferior

Table 3.25: LR Tests for Orthogonality and Sensitivity; Revised Data

	Orthogonality	Sensitivity	*value of β* *(s.e.)*
Total Consumption			
VAR(1)	$\chi^2(2)$ =8.719	$\chi^2(1)$ =6.432	0.352 (0.197)
VAR(2)	$\chi^2(4)$ =14.420	$\chi^2(3)$ =10.571	0.453 (0.203)
VAR(5)	$\chi^2(10)$ =27.197	$\chi^2(9)$ =24.741	0.409 (0.230)
Non-durable($\lambda = 1.16$)			
VAR(1)	$\chi^2(2)$ =8.952	$\chi^2(1)$ =0.853	0.505 (0.167)
VAR(2)	$\chi^2(4)$ =15.452	$\chi^2(3)$ =3.829	0.568 (0.166)
VAR(5)	$\chi^2(10)$ =24.449	$\chi^2(9)$ =14.096	0.471 (0.143)
Non-durable($\lambda = 1.37$)			
VAR(1)	$\chi^2(2)$ =8.137	$\chi^2(1)$ =1.459	0.382 (0.180)
VAR(2)	$\chi^2(4)$ =11.129	$\chi^2(3)$ =4.837	0.447 (0.188)
VAR(5)	$\chi^2(10)$ =19.802	$\chi^2(9)$ =13.061	0.352 (0.158)

Table 3.26: LR Tests of Restriction (3.35) for Blinder and Deaton Data

	Smoothness Test
Total Consumption	
VAR(1)	$\chi^2(2)$ =7.872
VAR(5)	$\chi^2(10)$ =24.081
Non-durable($\lambda = 1.33$)	
VAR(1)	$\chi^2(2)$ =6.223
VAR(5)	$\chi^2(10)$ =26.606

Table 3.27: LR Tests for Restriction (3.35) for the Revised Data Set

	Smoothness Test
Total Consumption	
VAR(1)	$\chi^2(2)$ =6.3018
VAR(5)	$\chi^2(10)$ =31.0292
Non-durable($\lambda = 1.52$)	
VAR(1)	$\chi^2(2)$ =4.5382
VAR(5)	$\chi^2(10)$ =32.9932

Table 3.28: LR Tests on Consumption for a VAR(4)

	ADD data set	Revised Data set
VAR(4) in levels	$\chi^2(8)$ =47.1	$\chi^2(8)$ =35.26
VAR(4) in logs	$\chi^2(8)$ =51.5	$\chi^2(8)$ =42.93

information should not be present in our VAR tests.

2. *The rational expectations permanent income model* (in particular, Flavin's (1981) version of Hall's (1978) specification) and

3. *The excess sensitivity hypothesis do not hold for US and UK data.* The Wald and Likelihood ratio tests that introduce the equivalent restrictions for both hypotheses on the coefficients of a VAR system for the change in labour income and savings report failures of these hypotheses. Whilst the failure of REPI is not new, the failure of the excess sensitivity hypothesis is. This questions the theoretical arguments used by Flavin (1993) to explain the excess sensitivity phenomenon.

Both failures on the extended data sets may not be so surprising if we take into account the behaviour of UK and US consumers in the latter part of the 1980s and early 1990s. Far from the smoothing nature expected by both the permanent income and excess sensitivity hypotheses, we observe a degree of variability in most of the time series; an increase in consumption in the second half of the 1980s, followed by a subsequent decrease in the early part of the 1990s. This behaviour is even more pronounced for the UK economy. Nonetheless, the aggregate data for both the US and the UK display some of the characteristics that one would expect if agents were to smooth their behaviour: firstly, we find that savings do Granger-cause labour income innovations and in particular, savings help to forecast declines in labour income in all the estimated VARs (the sum of the coefficients of saving in the VAR equation explaining labour income change are negative). This result is robust to the level and logarithmic approximations for both the US and the UK.

What is surprising is the failure of the less restrictive excess sensitivity hypothesis in both levels and logs for all the consumption specifications. The revisions made to the data sets used for both the US and the UK, cannot explain this finding - the shorter and revised period 1959:3 to 1984:4 for the US fails to reject the hypothesis - which leads us to believe that such failures must be the result of the behaviour of all the economic variables in the latter part of the 1980s and early 1990s. It is not straightforward to find an explanation for the finding that the excess sensitivity hypothesis fails. In her original paper, Flavin (1985) suggested that a finding of a significantly positive marginal propensity to consume out of transitory income would not invalidate the basic postulates of the (smoothing) life cycle model but could suggest binding liquidity constraints. For Flavin, consumers are not myopic when they plan their intertemporal consumption pattern but are prevented from realising those plans because they lack liquidity. If we take this view to be the correct interpretation of the excess sensitivity hypothesis, then it could be possible to link this failure to a number of factors. Muellbauer (1994) provides a number of reasons for the change in behaviour of US and UK consumers that can explain why the permanent income hypothesis continues to fail. Two of those reasons could perhaps explain the failure of excess sensitivity. One is the role played by (illiquid) assets and the other is financial deregulation. Both, it is claimed, can lead to an ease in the amount of credit constraint that consumers face. Muellbauer points out that the advent of financial liberalization led credit to become easily available as banks fought to gain new borrowers. Financial liberalization also made illiquid assets such as pensions, tax benefits from saving schemes, houses, etc. more spendable hence decreasing the impact of liquidity constraints[24].

Some authors have questioned the interpretation of Flavin's work as representing the behaviour of forward-looking, liquidity constrained agents. Hadjimatheou (1987) argues that forward-looking 'risk-averse consumers must be aware of the probability of being faced with liquidity constraints in the future and they should therefore be expected to allow fully for it in their plans' [pp. 78].

Apart from the tempting conclusion that the assumptions inherent in the derivation of the Hall and Flavin's papers maybe to blame for the failure of both hypotheses, it is also possible that a more 'damaging', yet more interesting conclusion could account for the results that were reported above. The failure of both of the smoothing life cycle models may imply that the smoothing behaviour of economic agents, if existent, may have changed or changes through time. Given that Flavin's excess sensitivity hypothesis cannot be rejected for the original Blinder and Deaton and the equivalent period revised data set, but is rejected for the data set

[24]These are perhaps the factors that contribute to the excess sensitivity parameter to be insignificant.

that includes the observations in the 1980s and 1990s, one could conjecture that the smoothing behaviour of agents has changed.

3.9 Appendix 1: Cointegrating Results for Disposable Income and Consumption

3.9.1 US data

Total Consumption (t-statistics in brackets)

$$y_t^d = \underset{(8.13)}{0.627} + \underset{(174.94)}{1.028}\, c_t \;\; R^2 = 0.995$$

Test of no cointegration

i. Intercept and no Trend (95% CV is -2.882):

Augmented DF Test with 0 lags -3.6, 1 lag -2.94, 4 lags -2.98

ii. Intercept and Trend (95% CV is -3.44):

Augmented DF Test with 0 lags -3.7, 1 lag -2.95, 4 lags -2.96

Non-Durable Consumption

$$y_t^d = \underset{(-11.31)}{-0.691} + \underset{(241.94)}{1.335}\, cn_t \;\; R^2 = 0.997$$

Test of no cointegration

i. Intercept and no Trend (95% CV is -2.882):

Augmented DF Test with 0 lags -4.7, 1 lag -4.14, 4 lags -4.17

ii. Intercept and Trend (95% CV is -3.44):

Augmented DF Test with 0 lags -4.7, 1 lag -4.13, 4 lags -4.16

3.9.2 UK data

Total Consumption

$$y_t^d = \underset{(3.86)}{-34.037} + \underset{(140.76)}{1.136}\, c_t \;\; R^2 = 0.991$$

Test of no cointegration

i. Intercept and no Trend (95% CV is -2.87):

Augmented DF Test with 0 lags -4.3, 1 lag -2.82, 4 lags -2.40

ii. Intercept and Trend (95% CV is -3.43):

Augmented DF Test with 0 lags -4.3, 1 lag -2.83, 4 lags -2.41

Non-Durable Consumption

$$y_t^d = \underset{(-15.78)}{-135.81} + \underset{(155.63)}{1.338}\, cn_t \;\; R^2 = 0.993$$

Test of no cointegration

Intercept and no Trend (95% CV is -2.87):

Augmented DF Test with 0 lags -4.4, 1 lag -2.73, 4 lags -2.25

Intercept and Trend (95% CV is -3.43):

Augmented DF Test with 0 lags -4.4, 1 lag -2.75, 4 lags -2.26

3.10 Appendix 2: Coefficient Results and Granger Causality Tests, Logarithmic Data

3.10.1 US data

VAR(1) : Total Consumption (Standard Errors in brackets).

Labour Income			
a_1, b_1	0.123 (0.083)	−0.086 (0.041)	
Serial Corr	$\chi^2(4) = 14.3$	Normality	$\chi^2(2) = 31.2$
Savings			
c_1, d_1	−0.098 (0.085)	0.879 (0.042)	
Serial Corr	$\chi^2(4) = 3.61$	Normality	$\chi^2(2) = 4.96$

VAR(1): Non-durable Consumption, ($\lambda = 1.18$)

Labour Income			
a_1, b_1	0.103 (0.079)	−0.124 (0.034)	
Serial Corr	$\chi^2(4) = 4.99$	Normality	$\chi^2(2) = 15.6$
Savings			
c_1, d_1	−0.069 (0.079)	0.903 (0.034)	
Serial Corr	$\chi^2(4) = 1.3$	Normality	$\chi^2(2) = 0.17$

VAR(1): Non-durable Consumption, ($\lambda = 1.33$)

Labour Income			
a_1, b_1	0.112 (0.079)	−0.125 (0.033)	
Serial Corr	$\chi^2(4) = 4.94$	Normality	$\chi^2(2) = 15.4$
Savings			
c_1, d_1	−0.072 (0.089)	0.887 (0.037)	
Serial Corr	$\chi^2(4) = 1.65$	Normality	$\chi^2(2) = 0.27$

VAR(2) : Total Consumption

Labour Income				
a_1, a_2, b_1, b_2	0.326 (0.101)	0.017 (0.08)	−0.403 (0.988)	0.348 (0.100)
Serial Corr	$\chi^2(4) = 3.19$	Normality	$\chi^2(2) = 32.5$	
Savings				
c_1, c_2, d_1, d_2	0.132 (0.107)	−0.086 (0.087)	0.723 (0.105)	0.176 (0.106)
Serial Corr	$\chi^2(4) = 2.2$	Normality	$\chi^2(2) = 4.7$	

VAR(2): Non-durable Consumption, ($\lambda = 1.18$)

Labour Income				
a_1, a_2, b_1, b_2	0.334 (0.119)	0.026 (0.080)	−0.434 (0.119)	0.329 (0.123)
Serial Corr	$\chi^2(4) = 4.77$	Normality	$\chi^2(2) = 17.6$	
Savings				
c_1, c_2, d_1, d_2	0.01 (0.122)	−0.053 (0.082)	0.806 (0.122)	0.103 (0.127)
Serial Corr	$\chi^2(4) = 1.96$	Normality	$\chi^2(2) = 0.26$	

VAR(2): Non-durable Consumption, ($\lambda = 1.33$)

Labour Income				
a_1, a_2, b_1, b_2	0.349 (0.123)	0.029 (0.080)	−0.403 (0.109)	0.295 (0.114)
Serial Corr	$\chi^2(4) = 5.4$	Normality	$\chi^2(2) = 17.5$	
Savings				
c_1, c_2, d_1, d_2	0.337 (0.141)	−0.058 (0.924)	0.773 (0.126)	0.123 (0.131)
Serial Corr	$\chi^2(4) = 2.1$	Normality	$\chi^2(2) = 0.42$	

VAR(5) : Total Consumption

Labour Income					
a_1, a_2, a_3, a_4, a_5	0.289 (0.109)	−0.132 (0.115)	0.059 (0.111)	0.149 (0.109)	−0.172 (0.847)
b_1, b_2, b_3, b_4, b_5	−0.355 (0.143)	0.295 (0.148)	−0.121 (0.147)	0.051 (0.147)	0.099 (0.109)
Serial Corr	$\chi^2(4) = 2.81$		Normality	$\chi^2(2) = 43.0$	
Savings					
c_1, c_2, c_3, c_4, c_5	0.001 (0.116)	−0.101 (0.118)	−0.141 (0.118)	0.265 (0.115)	−0.075 (0.090)
d_1, d_2, d_3, d_4, d_5	0.755 (0.111)	0.205 (0.157)	0.071 (0.157)	−0.259 (0.156)	0.117 (0.116)
Serial Corr	$\chi^2(4) = 7.52$		Normality	$\chi^2(2) = 3.19$	

VAR(5): Non-durable Consumption, ($\lambda = 1.18$)

Labour Income					
a_1, a_2, a_3, a_4, a_5	0.268 (0.128)	−0.0236 (0.129)	0.049 (0.129)	0.205 (0.125)	−0.177 (0.081)
b_1, b_2, b_3, b_4, b_5	−0.378 (0.124)	0.285 (0.191)	−0.138 (0.188)	−0.052 (0.186)	0.162 (0.127)
Serial Corr	$\chi^2(4) = 4.77$		Normality	$\chi^2(2) = 29.9$	
Savings					
c_1, c_2, c_3, c_4, c_5	−0.509 (0.131)	−0.122 (0.133)	−0.102 (0.133)	0.358 (0.128)	−0.149 (0.083)
d_1, d_2, d_3, d_4, d_5	0.877 (0.127)	0.103 (0.196)	−0.075 (0.193)	−0.28 (0.191)	0.276 (0.130)
Serial Corr	$\chi^2(4) = 6.59$		Normality	$\chi^2(2) = 0.52$	

VAR(5): Non-durable Consumption, ($\lambda = 1.33$)

Labour Income					
a_1, a_2, a_3, a_4, a_5	0.271 (0.132)	−0.018 (0.134)	0.069 (0.134)	0.209 (0.130)	−0.175 (0.081)
b_1, b_2, b_3, b_4, b_5	−0.344 (0.115)	0.243 (0.176)	−0.134 (0.173)	−0.033 (0.171)	0.147 (0.117)
Serial Corr	$\chi^2(4) = 4.84$		Normality	$\chi^2(2) = 29.1$	
Savings					
c_1, c_2, c_3, c_4, c_5	−0.056 (0.152)	−0.139 (0.154)	−0.139 (0.154)	0.408 (0.149)	−0.162 (0.093)
d_1, d_2, d_3, d_4, d_5	0.851 (0.133)	0.117 (0.202)	−0.087 (0.198)	−0.276 (0.196)	0.275 (0.135)
Serial Corr	$\chi^2(4) = 7.35$		Normality	$\chi^2(2) = 0.94$	

Granger Causality tests for the US (p-values in brackets)

	Total Consumption	Non-durable, $\lambda = 1.18$	Non-durable, $\lambda = 1.33$
t-test (VAR-1)	−2.074 (0.039)	−3.642 (0.0003)	−3.796 (0.0002)
Wald test (VAR-2)	18.675 (0.001)	21.890 (0.00002)	22.878 (0.00001)
Wald test (VAR-5)	17.888 (0.003)	26.784 (0.0006)	28.099 (0.00003)

3.10.2 UK data

VAR(1): Total Consumption

Labour Income			
a_1, b_1	−0.056 (0.0762)	−0.096 (0.034)	
Serial Corr	$\chi^2(4) = 3.1$	Normality	$\chi^2(2) = 10.4$
Savings			
c_1, d_1	−0.216 (0.097)	0.823 (0.043)	
Serial Corr	$\chi^2(4) = 25.1$	Normality	$\chi^2(2) = 3.19$

VAR(1): Non-durable Consumption, ($\lambda = 1.08$)

Labour Income			
a_1, b_1	−0.052 (0.077)	−0.039 (0.023)	
Serial Corr	$\chi^2(4) = 4.14$	Normality	$\chi^2(2) = 11.5$
Savings			
c_1, d_1	−0.188 (0.028)	0.929 (0.028)	
Serial Corr	$\chi^2(4) = 33.2$	Normality	$\chi^2(2) = 10.9$

VAR(1): Non-durable Consumption, ($\lambda = 1.33$)

Labour Income			
a_1, b_1	−0.049 (0.077)	−0.044 (0.022)	
Serial Corr	$\chi^2(4) = 4.31$	Normality	$\chi^2(2) = 10.5$
Savings			
c_1, d_1	−0.231 (0.106)	0.921 (0.029)	
Serial Corr	$\chi^2(4) = 29.8.$	Normality	$\chi^2(2) = 5.57$

VAR(3): Total Consumption

Labour Income			
a_1, a_2, a_3	−0.041 (0.086)	0.087 (0.085)	−0.095 (0.078)
b_1, b_2, b_3	−0.117 (0.072)	−0.092229 (0.078)	0.112 (0.073)
Serial Corr	$\chi^2(4) = 2.0$	Normality	$\chi^2(2) = 9.96$
Savings			
c_1, c_2, c_3	0.028 (0.100)	0.080 (0.099)	−0.245 (0.090)
d_1, d_2, d_3	0.436 (0.084)	0.235 (0.091)	0.217 (0.085)
Serial Corr	$\chi^2(4) = 8.19$	Normality	$\chi^2(2) = 9.74$

VAR(3): Non-durable Consumption, ($\lambda = 1.08$)

Labour Income			
a_1, a_2, a_3	0.009 (0.093)	0.131 (0.092)	−0.094 (0.079)
b_1, b_2, b_3	−0.137 (0.084)	−0.061 (0.892)	0.161 (0.085)
Serial Corr	$\chi^2(4) = 2.76$	Normality	$\chi^2(2) = 8.84$
Savings			
c_1, c_2, c_3	0.1403 (0.101)	0.056 (0.099)	−0.202 (0.086)
d_1, d_2, d_3	0.415 (0.092)	0.329 (0.969)	0.211 (0.092)
Serial Corr	$\chi^2(4) = 6.1$	Normality	$\chi^2(2) = 31.4$

VAR(3): Non-durable Consumption, ($\lambda = 1.33$)

Labour Income			
a_1, a_2, a_3	0.043 (0.098)	0.170 (0.097)	−0.098 (0.078)
b_1, b_2, b_3	−0.159 (0.077)	−0.058 (0.082)	0.179 (0.077)
Serial Corr	$\chi^2(4) = 2.38$	Normality	$\chi^2(2) = 8.12$
Savings			
c_1, c_2, c_3	0.205 (0.125)	0.109 (0.122)	−0.250 (0.099)
d_1, d_2, d_3	0.402 (0.097)	0.319 (0.104)	0.229 (0.098)
Serial Corr	$\chi^2(4) = 6.06$	Normality	$\chi^2(2) = 28.7$

VAR(4): Total Consumption

Labour Income				
a_1, a_2, a_3, a_4	−0.046 (0.087)	0.102 (0.087)	−0.534 (0.086)	0.082 (0.079)
b_1, b_2, b_3, b_4	−0.114 (0.074)	−0.013 (0.815)	0.074 (0.081)	0.082 (0.075)
Serial Corr	$\chi^2(4) = 7.06$	Normality	$\chi^2(2) = 7.03$	
Savings				
c_1, c_2, c_3, c_4	0.029 (0.100)	0.097 (0.100)	−0.221 (0.992)	0.119 (0.092)
d_1, d_2, d_3, d_4	0.453 (0.086)	0.201 (0.094)	0.191 (0.948)	0.066 (0.087)
Serial Corr	$\chi^2(4) = 3.25$	Normality	$\chi^2(2) = 8.78$	

VAR(4): Non-durable Consumption, ($\lambda = 1.08$)

Labour Income				
a_1, a_2, a_3, a_4	−0.001 (0.094)	0.151 (0.094)	−0.293 (0.093)	0.087 (0.080)
b_1, b_2, b_3, b_4	−0.131 (0.086)	−0.109 (0.095)	0.102 (0.095)	0.106 (0.087)
Serial Corr	$\chi^2(4) = 6.81$	Normality	$\chi^2(2) = 6.19$	
Savings				
c_1, c_2, c_3, c_4	0.130 (0.101)	0.087 (0.102)	−0.137 (0.100)	0.058 (0.087)
d_1, d_2, d_3, d_4	0.416 (0.093)	0.270 (0.103)	0.144 (0.103)	0.133 (0.095)
Serial Corr	$\chi^2(4) = 2.09$	Normality	$\chi^2(2) = 28.7$	

VAR(4): Non-durable Consumption, ($\lambda = 1.33$)

Labour Income				
a_1, a_2, a_3, a_4	0.026 (0.091)	0.192 (0.099)	−0.025 (0.098)	0.079 (0.080)
b_1, b_2, b_3, b_4	−0.152 (0.078)	−0.103 (0.088)	0.124 (0.877)	0.096 (0.079)
Serial Corr	$\chi^2(4) = 6.87$	Normality	$\chi^2(2) = 5.84$	
Savings				
c_1, c_2, c_3, c_4	0.189 (0.125)	0.142 (0.126)	−0.169 (0.124)	0.079 (0.101)
d_1, d_2, d_3, d_4	0.410 (0.098)	0.263 (0.111)	0.166 (0.111)	0.121 (0.112)
Serial Corr	$\chi^2(4) = 2.55$	Normality	$\chi^2(2) = 25.6$	

Granger Causality for the UK

	Total Consumption	Non-durable, $\lambda = 1.08$	Non-durable, $\lambda = 1.33$
t-test (VAR-1)	−2.846 (0.005)	−1.656 (0.099)	−2.027 (0.019)
Wald test (VAR-3)	11.642 (0.009)	7.044 (0.071)	10.538 (0.014)
Wald test (VAR-4)	11.343 (0.023)	7.541 (0.109)	10.757 (0.029)

3.11 Appendix 3: Coefficient Results and Granger Causality Tests, Level Data

3.11.1 US data

VAR(1) : Total Consumption (Standard Errors in brackets).

Labour Income			
a_1, b_1	−0.085 (0.080)	−0.136 (0.043)	
Serial Corr	$\chi^2(4) = 9.35$	Normality	$\chi^2(2) = 65.9$
Savings			
c_1, d_1	−0.256 (0.087)	0.842 (0.046)	
Serial Corr	$\chi^2(4) = 3.73$	Normality	$\chi^2(2) = 14.5$

VAR(1): Non-durable Consumption, ($\lambda = 1.33$)

Labour Income			
a_1, b_1	−0.032 (0.083)	−0.223 (0.062)	
Serial Corr	$\chi^2(4) = 7.69$	Normality	$\chi^2(2) = 78.3$
Savings			
c_1, d_1	−0.159 (0.085)	0.734 (0.063)	
Serial Corr	$\chi^2(4) = 2.46$	Normality	$\chi^2(2) = 21.1$

VAR(2) : Total Consumption

Labour Income				
a_1, a_2, b_1, b_2	0.171 (0.113)	0.0003 (0.082)	−0.444 (0.102)	0.343 (0.107)
Serial Corr	$\chi^2(4) = 3.67$	Normality	$\chi^2(2) = 87.1$	
Savings				
c_1, c_2, d_1, d_2	−0.104 (0.126)	−0.088 (0.091)	0.648 (0.112)	0.217 (0.116)
Serial Corr	$\chi^2(4) = 3.94$	Normality	$\chi^2(2) = 15.0$	

VAR(2): Non-durable Consumption, ($\lambda = 1.33$)

Labour Income				
a_1, a_2, b_1, b_2	0.217 (0.129)	0.053 (0.094)	−0.498 (0.128)	0.309 (0.129)
Serial Corr	$\chi^2(4) = 4.42$	Normality	$\chi^2(2) = 81.7$	
Savings				
c_1, c_2, d_1, d_2	−0.124 (0.133)	0.036 (0.092)	0.697 (0.1316)	0.038 (0.136)
Serial Corr	$\chi^2(4) = 2.73$	Normality	$\chi^2(2) = 20.4$	

VAR(5) : Total Consumption

Labour Income					
a_1, a_2, a_3, a_4, a_5	0.141 (0.117)	−0.067 (0.122)	0.102 (0.124)	0.046 (0.122)	−0.164 (0.081)
b_1, b_2, b_3, b_4, b_5	−0.405 (0.106)	0.346 (0.155)	−0.192 (0.160)	0.074 (0.158)	0.079 (0.158)
Serial Corr	$\chi^2(4) = 3.84$		Normality	$\chi^2(2) = 102.1$	
Savings					
c_1, c_2, c_3, c_4, c_5	−0.138 (0.132)	−0.134 (0.133)	−0.141 (0.133)	0.238 (0.136)	−0.120 (0.090)
d_1, d_2, d_3, d_4, d_5	0.688 (0.123)	0.264 (0.174)	0.028 (0.177)	−0.308 (0.173)	0.178 (0.131)
Serial Corr	$\chi^2(4) = 7.63$		Normality	$\chi^2(2) = 9.67$	

VAR(5): Non-durable Consumption, ($\lambda = 1.33$)

Labour Income					
a_1, a_2, a_3, a_4, a_5	0.185 (0.149)	−0.050 (0.157)	0.124 (0.157)	0.124 (0.145)	−0.128 (0.097)
b_1, b_2, b_3, b_4, b_5	−0.455 (0.148)	0.379 (0.218)	−0.208 (0.227)	−0.020 (0.207)	0.131 (0.146)
Serial Corr	$\chi^2(4) = 3.81$		Normality	$\chi^2(2) = 108.9$	
Savings					
c_1, c_2, c_3, c_4, c_5	−0.164 (0.149)	−0.808 (0.152)	−0.092 (0.154)	0.363 (0.149)	−0.089 (0.099)
d_1, d_2, d_3, d_4, d_5	0.739 (0.153)	0.135 (0.212)	−0.059 (0.217)	−0.357 (0.206)	0.267 (0.153)
Serial Corr	$\chi^2(4) = 4.07$		Normality	$\chi^2(2) = 22.8$	

Granger Causality tests for the US (p-values in brackets)

	Total Consumption	Non-durable, $\lambda = 1.33$
t-test (VAR-1)	-4.201 0.0004	-6.725 (0.0000)
Wald test (VAR-2)	33.292 (0.0000)	22.049 (0.0000)
Wald test (VAR-5)	33.293 (0.0000)	41.405 (0.00003)

3.11.2 UK data

VAR(1) : Total Consumption

Labour Income			
a_1, b_1	-0.055 (0.075)	-0.127 (0.034)	
Serial Corr	$\chi^2(4) = 1.99$	Normality	$\chi^2(2) = 7.18$
Savings			
c_1, d_1	-0.295 (0.102)	0.786 (0.046)	
Serial Corr	$\chi^2(4) = 17.4$	Normality	$\chi^2(2) = 5.58$

VAR(1): Non-durable Consumption, ($\lambda = 1.33$)

Labour Income			
a_1, b_1	−0.037 (0.077)	−0.137 (0.039)	
Serial Corr	$\chi^2(4) = 2.27$	Normality	$\chi^2(2) = 6.01$
Savings			
c_1, d_1	−0.267 (0.098)	0.787 (0.048)	
Serial Corr	$\chi^2(4) = 21.8$	Normality	$\chi^2(2) = 12.9$

VAR(3) : Total Consumption

Labour Income			
a_1, a_2, a_3	−0.054 (0.085)	0.078 (0.085)	−0.089 (0.079)
b_1, b_2, b_3	−0.123 (0.065)	−0.089 (0.073)	0.085 (0.067)
Serial Corr	$\chi^2(4) = 2.00$	Normality	$\chi^2(2) = 7.09$
Savings			
c_1, c_2, c_3	−0.849 (0.106)	0.012 (0.104)	−0.273 (0.100)
d_1, d_2, d_3	0.497 (0.081)	0.229 (0.091)	0.134 (0.086)
Serial Corr	$\chi^2(4) = 7.68$	Normality	$\chi^2(2) = 18.7$

VAR(3): Non-durable Consumption, ($\lambda = 1.33$)

Labour Income			
a_1, a_2, a_3	−0.009 (0.091)	0.113 (0.089)	−0.087 (0.081)
b_1, b_2, b_3	−0.175 (0.075)	−0.072 (0.082)	0.125 (0.078)
Serial Corr	$\chi^2(4) = 1.59$	Normality	$\chi^2(2) = 5.62$
Savings			
c_1, c_2, c_3	−0.016 (0.107)	−0.409 (0.103)	−0.215 (0.096)
d_1, d_2, d_3	0.454 (0.088)	0.317 (0.096)	0.096 (0.091)
Serial Corr	$\chi^2(4) = 5.61$	Normality	$\chi^2(2) = 66.0$

VAR(4) : Total Consumption

Labour Income				
a_1, a_2, a_3, a_4	−0.0636 (0.096)	0.095 (0.087)	−0.412 (0.096)	0.044 (0.088)
b_1, b_2, b_3, b_4	−0.121 (0.071)	−0.121 (0.082)	0.039 (0.078)	0.095 (0.072)
Serial Corr	$\chi^2(4) = 4.87$	Normality	$\chi^2(2) = 5.04$	
Savings				
c_1, c_2, c_3, c_4	−0.079 (0.114)	0.009 (0.118)	−0.281 (0.107)	0.110 (0.101)
d_1, d_2, d_3, d_4	0.511 (0.081)	0.225 (0.099)	0.148 (0.98)	−0.019 (0.082)
Serial Corr	$\chi^2(4) = 5.81$	Normality	$\chi^2(2) = 21.1$	

VAR(4): Non-durable Consumption, ($\lambda = 1.33$)

Labour Income				
a_1, a_2, a_3, a_4	−0.024 (0.094)	0.128 (0.089)	−0.035 (0.089)	0.047 (0.081)
b_1, b_2, b_3, b_4	−0.166 (0.078)	−0.109 (0.096)	0.0764 (0.859)	0.094 (0.081)
Serial Corr	$\chi^2(4) = 5.22$	Normality	$\chi^2(2) = 4.27$	
Savings				
c_1, c_2, c_3, c_4	−0.025 (0.111)	−0.031 (0.105)	−0.181 (0.105)	0.025 (0.094)
d_1, d_2, d_3, d_4	0.459 (0.088)	0.292 (0.109)	0.0638 (0.103)	0.061 (0.096)
Serial Corr	$\chi^2(4) = 5.62$	Normality	$\chi^2(2) = 59.5$	

Granger Causality for the UK

	Total Consumption	Non-durable, $\lambda = 1.33$
t-test (VAR-1)	−3.665 (0.0003)	−3.547 (0.00045)
Wald test (VAR-3)	15.505 (0.001)	25.342 (0.00001)
Wald test (VAR-4)	15.592 (0.004)	17.474 (0.002)

3.12 Appendix 4: A Simple Test for the Failure of the Evolution of Savings Equation, Is Excess Smoothness to Blame?

The conclusions about the failures of the consumption hypotheses have left the failure of approximation (3.10) unexplained. For the approximation in that equation to be true two factors must be satisfied: that all the logarithmic approximations do in fact hold and that consumers are on their evolution of assets constraint. The last assumption has the added assumption that the rate of interest should be constant through time. Reverting to a model in levels does provide some useful information about the logarithmic approximations used and whether their use should be continued in further research.

From the definition of savings and the evolution of assets equation $A_t - (1+r)A_{t-1} = (1+r)(y_{t-1} - c_{t-1}) + \psi_t$ we have

$$\Delta c_t + \left(\frac{r}{1+r}\right)\psi_t = -\left[s_t - \Delta y_t - (1+r)s_{t-1}\right] \qquad (3.44)$$

It is often assumed in the literature that unanticipated capital gains, ψ_t, are equal to zero. The importance of this assumption together with assumptions about other unobservable variables in this framework such as transitory income and consumption cannot be taken lightly because they enable the testing of REPI. Those same assumptions enable a test of the truth of equation (3.44) because both sides are now observable. If this equation holds, then one can conclude that the logarithmic approximations are the reason behind the failure of expression (3.10). If (3.44) fails however, then either the definition of the evolution of assets, or the assumptions about the unobservable variables in the framework can be accountable for those failures. In tables 29 and 30, the means and specially the standard deviations of the components on the right hand-side of (3.44) are compared to the mean and standard deviation of the change in consumption for a rate of interest of 10%[25]. In those tables, the right hand-side of equation (3.44) is termed $\Delta\eta_t$.

The evidence collected in tables 29 and 30 does not favour equation (3.44). For both the US and the UK, we find that the mean of $\Delta\eta_t$ *is less than the mean of* Δc_t in absolute value and more importantly, that *the standard deviation of* $\Delta\eta_t$ *is always greater than the standard*

[25] The results are not very sensitive to the rate of interest used.

Table 3.29: Means and Standard Deviations for the US, eq. (3.37)

US	Mean	s.d	Scaling Factor
Δc_t	0.067	0.082	
$\Delta \eta_t$	-0.06	0.095	
$\lambda \Delta c_{nt}$	0.068	0.065	1.33
$\lambda \Delta \eta_t$	-0.05	0.077	1.33

Table 3.30: Means and Standard Deviations for the UK, eq. (3.37)

UK	Mean	s.d.	Scaling Factor
Δc_t	5.539	12.507	
$\Delta \eta_t$	3.909	18.453	
$\lambda \Delta c_{nt}$	6.264	11.468	1.33
$\lambda \Delta \eta_{nt}$	-2.863	16.283	1.33

deviation of the change in consumption. For the US the standard deviation of total consumption and non-durable consumption are, respectively, 15% and 20% lower than the standard deviation of $\Delta \eta_t$ and its equivalent measure for non-durables. These figures compare more favourably for the logarithmic case, where the differences are only of 13% and 10% respectively (see table 14). The same story occurs for the UK, although in this case the differences are greater. In levels, we find that the standard deviations of $\Delta \eta_t$ for the total measure of savings and the non-durable measure are respectively 47% and 42% higher than their consumption equivalents. For the logarithmic approximations, these figures drop to 40% and 36% respectively (see table 15). Thus, logarithmic approximations should not be blamed for the failure of equation (3.10); the failure of the definition asset constraint could explain the nature of these results. This means that using our definition of savings may be an incorrect means for testing the REPI because equation (3.41) would not then be an accurate statement about the permanent income hypothesis. Hence it would not make much sense to use such a definition of savings in a VAR framework to gather accurate information about the expectations that agents have about future labour income -which overcame the superior information problem. With this evidence we may draw two conclusions i) the test for smoothness that examined the ratio of the predicted theoretical and actual innovations standard deviations may not be an accurate form to measure the degree of smoothness in the data; ii) more research should quantify the importance of unanticipated capital gains and transitory consumption in this framework. A finding that unanticipated capital gains are the cause of the discrepancies reported above must surely lead to more research on those variables that were previously thought to be insignificant in this literature such as transitory consumption, transitory income and unanticipated capital gains.

Chapter 4

Theoretical Implications of Imperfect Information, Excess Sensitivity and Partial Adjustment

4.1 Introduction

Since the publication of Hall's (1978) claim that the consumption behaviour of forward-looking, rational agents could be accurately approximated by a random walk, a large amount of evidence at the aggregate level has been amassed against that assertion[1]. A number of studies have examined possible extensions that could in principle explain the failure of REPI at the aggregate level. In this chapter, we combine three recent extensions to develop more general specifications of the consumption function. With certain restrictions, these general specifications can be tested against the three extensions themselves and REPI to gain more understanding about the behaviour of the consumption function.

In this chapter we produce a number of different models for consumption behaviour that can be tested using time series data. These are the result of combining three different papers; Attfield et al.'s (1992) *partial adjustment paper*, Flavin's (1993) *excess sensitivity paper*, and Goodfriend's (1992) and Pischke's (1995) *information papers*. We shall briefly explain these

[1] See the previous chapter.

papers before explaining the new models developed in this chapter.

4.2 A Review of the Three Models that Are Used Throughout this Chapter

4.2.1 Partial Adjustment (PIH2)

The costs of adjusting a variable may not be specific to the time the adjustment takes place but may extend to other periods too. In Attfield et al.'s model consumers are slower to adjust to permanent income changes (these are prompted by changes in labour income given the assumption that no unanticipated capital gains exist) than assumed by the REPI. The reasons behind slow adjustment maybe due to inertia, habit formation or liquidity constraints. In particular, the model considers the 'time absorbing costs of planning to ensure that actual consumption equals its desired level. [...] actual consumption in any period will depend not only on the level of permanent income in that period but also on the decomposition of permanent income into that component which became predictable the period before and so on.' [Attfield et al., pp. 1206]. With the tendency to discount the future, it is expected that the planning effort devoted to the immediate future will exceed the efforts assigned to plan for the future as rational agents equate the marginal cost (time and not being to able to enjoy consumption) of planning for period $t+i$ with the marginal benefit (obtaining a better forecast). The obvious implication of this type of adjustment cost is that the proportion of permanent income that is unpredictable well before the current consumption decision was made, is accountable for making actual consumption deviate from desired consumption. Hence, if a high proportion of permanent income were predictable well in advance, then actual consumption would be close to its desired (or permanent income) level. The model is written formally as

$$c_t = E_{t-n}y_t^p + \sum_{0}^{n-1} \gamma_i \Delta E_{t-i} y_t^p \tag{4.1}$$

where n is the time span where adjustment is less than full and the $\gamma's$ are partial adjustment coefficients. We expect to observe $0 \leq \gamma_0 \leq \gamma_1 \leq \cdots \leq \gamma_{n-1} < 1$ since it is assumed that it takes time to adjust to recent news about the labour income process. Note that if all the $\gamma's$

were equal to one, REPI would result. Given Flavin's (1981) definition of permanent income and the standard process for assets, $A_t = (1+r)A_{t-1} + y_{t-1} - c_{t-1}$ it is straightforward to show that the following equation holds

$$y_t^p = [1+r]\, y_{t-1}^p - rc_{t-1} + \omega_t \tag{4.2}$$

where $\omega_t = r\sum_{j=0}^{\infty} \delta^{j+1} \Delta E_t y_{t+j}$, $\delta^i = (1+r)^{-i}$. Permanent income at time t depends on the level of permanent income in the previous time period minus the annuity value of the previous period's consumption level plus any new information about the labour income process. From (4.2) and (4.1) we have $\Delta E_t y_t^p = \omega_t$, $\Delta E_{t-1} y_t^p = \zeta_0 \omega_{t-1}$ and for $j \geqq 2$,

$$E_{t-j+1} y_t^p - E_{t-j} y_t^p = \left[\prod_0^{j-2} \zeta_i\right] \omega_{t-j+1}$$

where $\zeta_i = [1 + r(1-\gamma_i)]$. The change in consumption is given by[2]

$$\begin{aligned}\Delta c_t &= \gamma_0 \omega_t + (\gamma_1 \zeta_0 - \gamma_0)\, \omega_{t-1} + (\gamma_2 \zeta_1 - \gamma_1)\, \zeta_0 \omega_{t-2} + \cdots \\ &\quad + \left\{ (\gamma_{n-1}\zeta_{n-2} - \gamma_{n-2}) \left[\prod_0^{n-3} \zeta_i\right] \right\} \omega_{t-n+1} + \left\{ (\zeta_{n-1} - \gamma_{n-1}) \left[\prod_0^{n-2} \zeta_i\right] \right\} \omega_{t-n}\end{aligned} \tag{4.3}$$

The innovation in consumption follows an $MA(n)$ process. The number of error terms relate to all the information gathered by the consumer and how she adjusts to it during the n periods used to predict what permanent income ought to be at time t. Equation (4.3) is able to explain the excess sensitivity and smoothness phenomena. Write (4.3) as

$$\Delta c_t = \sum^{n} \phi_i \omega_{t-i}$$

[2] Since

$$\Delta c_t = E_{t-n} y_t^p - E_{t-n-1} y_{t-1}^p + \sum_0^{n-1} \gamma_j \Delta E_{t-j} y_t^p - \sum_0^{n-1} \gamma_j \Delta E_{t-j-1} y_{t-1}^p$$

where $\phi_0 = 1 - \beta_0$, $\phi_j = [\beta_{j-1}/\rho - \beta_j]$ for $0 < j < n$, $\phi_n = \beta_{n-1}/\rho$ and $\beta_0 = 1 - \gamma_0$, $\beta_i = [1 - \gamma_i]\, \zeta_{i-1}\beta_{i-1}/\left(1 - \gamma_{i-1}\right)$ for $i > 0$ and note that since ω is white noise,

$$var\left(\Delta c_t\right) = var\left(\omega_t\right) \sum^{n} \phi_i^2$$

Excess smoothness will arise if $\sum^n \phi_i^2 < 1$. Excess sensitivity arises in this model because lagged shocks to permanent income which are likely to be correlated to lagged innovations in labour income are shown to influence the current change in consumption.

4.2.2 Excess Sensitivity

Flavin considers a specific alternative hypothesis to REPI where consumption exhibits excess sensitivity to current income. Excess sensitivity to current income is introduced with the assumption that the marginal propensity to consume out of transitory income is non-zero. Her model is written as

$$c_t = \beta y_t^T + y_t^p \tag{4.4}$$

where $0 < \beta < 1$, permanent income is defined as in Flavin (1981) and y_t^T denotes transitory income which is defined as the residual $y_t^T \equiv \left[y_t + \left(\frac{r}{1+r}\right) A_t\right] - y_t^p$. Flavin (1985) has suggested the finding that $\beta > 0$ does not mean that agents are myopic; households can be rational and forward looking as excess sensitivity of consumption to current income may be the result of liquidity constraints that prevent individuals from realising their consumption plans. We see this if we define the change in consumption as[3]

$$\Delta c_t = \beta \Delta y_t + (1 - \beta)\, \Delta y_t^p \tag{4.5}$$

The change in consumption due to an innovation in permanent income is less than one to one (i.e. $1 - \beta$). We can explain this if we assume imperfect capital markets, i.e. that liquidity constraints exist. Since consumers may not be able to obtain funds after an increase in their

[3]Where we assume no unanticipated capital gains so $\Delta y_t^p = \left(\frac{r}{1+r}\right) \sum_{i=0}^{\infty} \left(\frac{r}{1+r}\right)^i (E_t - E_{t-1})\, y_{t+i}$.

permanent income, they then have to resort to consume a proportion of their total income (by definition, that proportion is part of their transitory income). One presumes that the higher β is the more constrained the consumer would be. Excess sensitivity is present if $\beta \neq 0$ and excess smoothness if

$$var(\Delta c_t) = \beta^2 var(\Delta y_t) + 2\beta(1-\beta)\,cov(\Delta y_t, \Delta y_t^p) + (1-\beta)^2\, var(\Delta y_t^p) < var(\Delta y_t^p) \quad (4.6)$$

4.2.3 Information-Aggregation

Within the context of the true REPI, Pischke examines the behaviour agents may have when they gather information about their labour income to predict what their permanent income will be. Pischke develops a framework where the microfoundations of permanent income are still considered but where care is taken to explain possible aggregation issues. The aggregation problem is introduced in the context of information gathering under the presumption that aggregate information plays 'little role in household decisions since the economic environment in which individuals operate differs sharply from the economy as it is described by aggregate data.' [pp. 806]. Pischke finds that 'the optimal consumption response calculated on the basis of individual income processes differs substantially from the predictions of a representative agent model calibrated with aggregate data.' [pp. 806]. He pays attention to two models; one where agents may simply not care enough about aggregate information because ignoring it is not very costly for them (i.e. there is incomplete aggregate information) and another one where agents may lack the information on contemporaneous aggregate variables (this is a model originally developed by Goodfriend (1992)).

We explain these informational issues in terms of a simple framework which assumes that all individuals have identical labour income processes, but each agent faces a different realisation of that process every time period. For the simplest framework, it is assumed that income consists of a random walk with innovations that are common to all individuals and a white noise component with shocks that are uncorrelated across individuals. Mathematically,

$$\Delta y_{it} = \varepsilon_t + u_{it} - u_{it-1} \quad (4.7)$$

where i subscripts denote individual variables while no subscripts refer to aggregate variables. The agent is assumed to be able to observe all components; that is, the agent can distinguish between ε_t and u_{it}. It is also assumed that individual shocks (the u's) are mutually uncorrelated and will sum to zero for a large population

$$\Delta y_t = \frac{1}{n}\sum^{n} \Delta y_{it} = \varepsilon_t \tag{4.8}$$

which is the income process that the time series analyst observes from the aggregate data.

Unobservable Aggregate Shocks

It is assumed that the individual is unable to distinguish between aggregate and individual income components, i.e. the agent cannot distinguish between ε_t and u_{it}, so that for the micro agent her income process takes this form

$$\Delta y_{it} = \eta_{it} - \theta\eta_{it-1} \tag{4.9}$$

where $\eta = f(\varepsilon_t, u_{it})$ and $\theta = -\left(1 - \sqrt{1-4\rho^2}\right)/2\rho$ where $\rho = -\sigma_u^2/\left(\sigma_\varepsilon^2 + 2\sigma_u^2\right)$ is the autocorrelation coefficient and σ_ε^2 and σ_u^2 denote the variances of the aggregate and individual errors respectively. Given Flavin's definition of permanent income and (4.9), the change in individual consumption is given by

$$\Delta c_{it} = (1-\theta\delta)\,\eta_{it} \equiv A\eta_{it} = A\frac{\Delta y_{it}}{1-\theta L} \tag{4.10}$$

the last equality comes from (4.9) and the invertibility of $1-\theta L$. Using (4.8) we can obtain the aggregate consumption change as

$$\Delta c_t = \frac{1}{n}\sum^{n}\Delta c_{it} = A\frac{1}{n}\sum^{n}\frac{\Delta y_{it}}{1-\theta L} = A\frac{\Delta y_t}{1-\theta L} = A\frac{\varepsilon_t}{1-\theta L} \tag{4.11}$$

so consumption innovations follow an $AR(1)$ process, $\Delta c_t = A\varepsilon_t + \theta\Delta c_{t-1}$.[4]

[4] In chapter 2 we explained how this model generated excess sensitivity and smoothness.

Lagged Information About Aggregate Shocks

This is Goodfriend's model. In this model, agents are able to observe individual and aggregate shocks, but the latter are only observed with a one period delay. The income process for the individual now takes the form

$$\Delta y_{it} = v_{it} - u_{it-1} \tag{4.12}$$

where $v_{it} = \varepsilon_t + u_{it}$. At time t, the consumer is unable to distinguish between the aggregate and individual components and it can only observe v_{it} (i.e. it cannot distinguish between ε_t and u_{it}). Thus, when making a rational consumption decision, the consumer will attempt to attribute part of the current period innovation to each component given the relative histories of both shocks (these are represented by their variances). The optimal consumption response will have two parts in this case; one that is associated with a response to the innovation in the agent's income and a term that corrects for the error made in predicting both components of income in the previous period. The response to an income shock at time t would be

$$\omega v_{it} + (1-\omega)\frac{r}{1+r}v_{it} = \frac{\omega + r}{1+r}v_{it} \tag{4.13}$$

where $\omega = \sigma_\varepsilon^2 / \left(\sigma_\varepsilon^2 + \sigma_u^2\right)$ is a term that reflects the histories of the individual and aggregate components and represents the proportion of the total (new) income innovation the agent expects to be aggregate. The first term represents the adjustment to what is thought to be a permanent innovation in income and the second term represents a transitory innovation. This is because it is believed that 'individual income processes are much less persistent than aggregate income'. [Pischke, pp. 806]

Agents are likely to make mistakes when they predict the proportion of an aggregate and individual shock following an innovation in income. Given that the size of the aggregate shock can be observed the next time period and since any error in the signal extraction problem will affect permanent income and therefore consumption, rational agents will correct mistakes made previously. The (negative) of the error made in predicting the aggregate shock is given by

$\xi_{it-1} = \varepsilon_{t-1} - \omega v_{it-1} = (1-\omega)\,\varepsilon_{t-1} - \omega u_{it-1}$[5] so that the optimal response to any errors made the previous time period is

$$(1+r)\left[\xi_{it-1} + \frac{r}{1+r}\left(-\xi_{it-1}\right)\right] = \xi_{it-1} \tag{4.14}$$

Putting together both responses, (4.13) and (4.14), yields the optimal consumption response to innovations in labour income

$$\Delta c_{it} = \frac{\omega + r}{1+r} v_{it} + (1-\omega)\,\varepsilon_{t-1} - \omega u_{it-1} \tag{4.15}$$

Upon aggregation we obtain:

$$\Delta c_t = \frac{1}{n}\sum^{n} \Delta c_{it} = \frac{\omega + r}{1+r}\varepsilon_t + (1-\omega)\,\varepsilon_{t-1} \tag{4.16}$$

Note that the model of lagged information will always yield an $MA(1)$ process for the change in consumption if two conditions are met, i) aggregate information is available after one time period and ii) agents then correct the errors made in predicting the individual and aggregate components of each income shock[6].

4.3 Introducing Imperfect Information into the Excess Sensitivity Hypothesis

The first model that we examine introduces Pischke's imperfect information to the Flavin framework. Since the Flavin model could be considered a model of liquidity constraints, it is an interesting exercise to examine the reactions agents have when these find it hard to distinguish between individual and aggregate shocks to income.

[5] The negative of the error in predicting the individual component is given by $u_{it-1} - (1-\omega)\,v_{it-1} = -(1-\omega)\,\varepsilon_{t-1} + \omega u_{it-1}\ (= -\xi_{it-1})$

[6] In chapter 2 we explained how this model generated excess sensitivity and smoothness.

4.3.1 Incomplete Information

A Simple Income Process

When the simple income process of no information (4.9) is applied to Flavin's model of excess sensitivity given by equation (4.5), the behaviour of consumption changes at the individual level is

$$\Delta c_{it} = \beta \Delta y_{it} + (1-\beta)\left(1 - \frac{\theta}{1+r}\right)\eta_{it} \qquad (4.17)$$

since $r\sum_{j=0}^{\infty}\delta^{j+1}\Delta E_{it}y_{it+j} = \theta(\delta)\eta_{it}$ where $\delta = \frac{1}{1+r}$ and $\theta(L) = 1-\theta L$. This equation has the same features as Flavin's framework; provided $\beta \neq 0$ then excess sensitivity exists. The conditions for excess smoothness to occur are the same as those given by inequality (4.6)[7]. Hence the features of (4.17) at the individual level are dominated by the model of excess sensitivity. The interesting implications of the no information case arise at the aggregate level. To find the change in average per capita consumption, use (4.9) to substitute into (4.17) and sum over individuals

$$\Delta c_t = \frac{1}{n}\sum^{n}\Delta c_{it} = \beta\Delta y_t + (1-\beta)\left(1-\frac{\theta}{1+r}\right)\frac{\varepsilon_t}{1-\theta L}$$

provided $(1-\theta L)$ is invertible. This last expression can be rewritten as

$$(1-\theta L)\Delta c_t = \beta(1-\theta L)\Delta y_t + (1-\beta)A\varepsilon_t \qquad (4.18)$$

$A = \left(1-\frac{\theta}{1+r}\right)$. The effect on permanent income warranted by an innovation in labour income is reflected by the error term. As liquidity constraints may exist, consumers may not be able to borrow to consume what they believe is their best estimate of the value of their permanent

[7] $var(\Delta c_{it}) = \beta^2 var(\Delta y_{it}) + 2\beta(1-\beta)A\,cov(\Delta y_{it},\eta_{it}) + (1-\beta)^2 A^2 var(\eta_{it}) < A^2 var(\eta_{it})$ where $A \equiv \left(1-\frac{\theta}{1+r}\right)$. For the income process in question we have, $var(\Delta y_{it}) = (1+\theta^2)\sigma_\eta^2$, $cov(\Delta y_{it},\eta_{it}) = var(\eta_{it}) = \sigma_\eta^2$ so for the inequality to be satisfied it must meet the condition

$$\frac{\beta^2(1+\theta^2)}{A^2} + \frac{2\beta(1-\beta)}{A} + (1-\beta)^2 < 1$$

income. Assume that a positive shock to labour income occurs. Since the labour income process is difference-stationary, a positive shock to labour income will increase permanent income. Because liquidity constraints exist, the agent cannot borrow to smooth consumption and therefore realise its desired consumption level in the face of increased permanent income. The consumer is therefore forced to 'borrow' from its transitory income to achieve a level of consumption that will be close to its desired level. Hence, part of its current consumption change is related to a proportion of (labour) income.

The θL term introduces further dynamics to the model. This term is related to the imperfect information faced by agents and its significance becomes clearer when we explain the intuition behind the autoregressive term for the change in consumption. Due to imperfect information agents do not distinguish between aggregate and individual income and so any type of income shock is likely to lead individuals to be surprised the following time period thereby leading them to change their consumption as a function of their previous consumption change. Suppose a positive aggregate income shock occurs. All agents will see their income increase but they will assume that part of that income shock is individual in nature and therefore transitory. Their rational response will be to increase their consumption but not as much as warranted by the size of the (permanent) aggregate income shock. But because the shock is persistent, in the following period they will be surprised that their income is higher than expected and consumption is therefore increased. The same arguments affect the income component too; the lagged income term is not only multiplied by the excess sensitivity parameter β but also by θ indicating that agents are slow to respond to past aggregate innovations not only because they are constrained but also because they choose to ignore aggregate information. Take the example of a positive aggregate shock. Agents do not adjust to the shock as much as they should have done because they believed that part of the shock was transitory. As they adjusted to what they believed was the persistent component, agents faced liquidity constraints and were not able to borrow. Hence they borrowed from their transitory income component - which they thought was larger because they believed that their permanent income was smaller - by a proportion β (this is represented by the term $\beta\Delta y_{t-1}$). Next period, as agents are surprised to find that their permanent income in the previous period was higher than previously thought, and they realise that they had borrowed against their transitory component when this component was smaller than thought, then their overall reaction will be to adjust for the mistakes made before. They will reduce the proportion out of their transitory income and increase the proportion of consumption that comes from their permanent income (i.e. $\theta\Delta c_{t-1} - \beta\theta\Delta y_{t-1}$).

Note that if $\beta = 0$ the standard Pischke model is obtained. This is a testable restriction that can be imposed to our more general consumption specification.

This model is able to explain the failures of the REPI encountered by Campbell (1987), Campbell and Deaton (1989), Flavin (1993) and in our previous chapter. This model is also capable of explaining the failures of the excess sensitivity hypothesis encountered in the previous chapter. First, start with the definition of savings at the aggregate level; $s_t = -(1-\beta)\sum_1^\infty \delta^j E_t \Delta y_{t+j} = \left(\frac{1-\beta}{1-\theta L}\right)(1-A)\varepsilon_t$[8]. Campbell's first-order VAR representation of income changes and savings has the form,

$$\begin{bmatrix} \Delta y_t \\ s_t \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ \left(\frac{1-\beta}{1-\theta L}\right)\theta & (1+r) \end{bmatrix} \begin{bmatrix} \Delta y_{t-1} \\ s_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_t \\ (1-\beta)\left(\frac{\theta\delta}{1-\theta L}\right)\varepsilon_t \end{bmatrix}$$

Clearly, this VAR framework violates the REPI restrictions, $a_1 = c_1$ and $d_1 = (1+r)\,b_1$, and the excess sensitivity restrictions, $a_1 = (1-\beta)\,c_1$ and $d_1 = (1+r) + (1-\beta)\,b_1$ in (4.19)

$$\begin{bmatrix} \Delta y_t \\ s_t \end{bmatrix} = \begin{bmatrix} a_1 & b_1 \\ c_1 & d_1 \end{bmatrix} \begin{bmatrix} \Delta y_{t-1} \\ s_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{st} \end{bmatrix} \tag{4.19}$$

Because these restrictions do not hold, it is not possible to recover the term of the innovation in permanent income from the VAR system and hence overcome the problem of superior information examined by Campbell and Flavin[9]. The model can nonetheless explain the sensitivity and smoothness phenomena. It is straight-forward to see that sensitivity will arise if $\beta \neq 0$ and smoothness if inequality (4.6) holds.

[8] This result comes from the fact that at the individual level, $s_{it} = -(1-\beta)\,r\sum_1^\infty \delta^j E_{it}\Delta y_{it+j} = (1-\beta)\,\delta\theta\eta_{it}$. Aggregation yields, $s_t = (1-\beta)\,\delta\theta\eta_t = \left(\frac{1-\beta}{1-\theta L}\right)(1-A)\,\varepsilon_t$.

[9] See the Appendix for an explanation.

More General Income Processes

We now assume a more general labour income process where individual i's change in labour income has both aggregate and individual-specific components given by their respective Wold representations,

$$\Delta y_{it} = \phi(L)\,\varepsilon_t + \theta(L)\,u_{it} \tag{4.20}$$

where $\phi(L) = \sum_{i=0}^{\infty}\phi_i L^i$ and $\theta(L) = \sum_{i=0}^{\infty}\theta_i L^i$. Unlike the simple income process case we do not impose the condition that aggregate shocks are permanent and individual shocks are transitory in nature. We still impose the condition that individual shocks will sum to zero when aggregated, thus average per capita labour income evolves according to the following process

$$\Delta y_t = \phi(L)\,\varepsilon_t \tag{4.21}$$

Given imperfect information, the process for individual income changes has a Wold representation

$$\Delta y_{it} = A(L)\,\eta_{it} \tag{4.22}$$

η_{it} is a white noise error, $A(L) = \sum_{i=0}^{\infty} a_i L^i$ and $a_i = f\left(\varepsilon_t, u_{it}, \phi(L), \theta(L), \sigma_\varepsilon^2, \sigma_u^2\right)$. If we aggregate this last equation we have

$$\Delta y_t = A(L)\,\overline{\eta}_t = \phi(L)\,\varepsilon_t \tag{4.23}$$

where $\overline{\eta}_t \left(\equiv \frac{1}{n}\sum^n \eta_{it}\right)$is the mean of η_{it} and where we have equated (4.20) to (4.22) and summed over individuals to obtain the last equality. Note that the *average* innovation, $\overline{\eta}_t$, can be expressed as an *infinite* order autoregressive process for the innovation in labour income. But the innovation in the labour income process can also be re-expressed given our definitions above as a finite order moving average representation in the errors. Thus the *average innovation* can be re-expressed in terms of an infinite order moving average process for *aggregate errors.* Given Flavin's excess sensitivity equation (4.5) and that the innovation in permanent income is given by $A\left(\frac{r}{1+r}\right)\eta_{it}$ provided the labour process is (4.22), the change in individual consumption is

$$\Delta c_{it} = \beta\Delta y_{it} + (1-\beta)\,A\left(\frac{r}{1+r}\right)\eta_{it} \tag{4.24}$$

Summing over individuals, using (4.23) and assuming $A(L)$ is invertible, yields the aggregate consumption equation

$$\Delta c_t = \beta\Delta y_t + (1-\beta)\,A\left(\frac{r}{1+r}\right)\overline{\eta}_t = \beta\Delta y_t + (1-\beta)\,A\left(\frac{r}{1+r}\right)\frac{\phi(L)\,\varepsilon_t}{A(L)}$$

which we can rewrite as

$$A(L)\,\Delta c_t = \beta A(L)\,\Delta y_t + (1-\beta)\,A\left(\frac{r}{1+r}\right)\phi(L)\,\varepsilon_t \tag{4.25}$$

The number of lags in consumption and income are determined by the number of lags in the individual income process. The interpretation of this equation is similar to that one given for the simple income process case. The $A(L)$ component of this equation comes again from the lack of aggregate information that agents have, for consumption it expresses the failure to forecast accurately what the innovation in permanent income is, and for the income term it reflects the failure to forecast the size of transitory income. The error component shows the failure to adjust to the (correct) amount of permanent income due to the liquidity constraint[10] and also the further effect that imperfect information has through the lags $\phi(L)$ on the aggregate error term. These lags represent information deficiencies on the part of the consumer since all the previous aggregate labour income shocks had not been adjusted to by the agent in the correct manner. We note that Pischke's model results if $\beta = 0$ and Flavin's model would result if $A(L) = \phi(L) + \theta(L)$ (i.e. if there were perfect information in the economy so that the consumer would be able to distinguish between aggregate and individual components). For excess sensitivity and smoothness to arise, the same conditions as the simple income process are required. This is one of the equations that will be tested using time series data.

[10] All the perceived innovations in permanent income are multiplied by the coefficient $(1-\beta)$ meaning that individuals are not able to adjust fully to the innovation in permanent income.

4.3.2 Lagged Information About Aggregate Shocks

A Simple Income Process

Agents are able to observe individual and aggregate shocks, but the latter are only observed with a one period delay. The derivation of this model is somewhat more involved, but it always renders an $MA(1)$ process for the errors. We assume the simple income process (4.12) and that equation (4.5) holds. The consumption response to new information consists of two parts:

1. The response to new innovations in income. The innovation in permanent income without unanticipated capital gains is by definition $\Delta y_{it}^p = r\sum_{j=0}^{\infty}\delta^{j+1}\Delta E_{it}y_{t+j}$ which given imperfect information yields (4.13). Thus the (rational) response of consumption to a current innovation in income v_{it} is therefore

$$\beta\Delta y_{it} + (1-\beta)\left(\omega v_{it} + r\delta\left(1-\omega\right)v_{it}\right)$$

   since $\omega v_{it} \approx \varepsilon_t$, $(1-\omega)\,v_{it} \approx u_{it}$ and where $\omega = \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2+\sigma_u^2}$.

2. Response to new information about the innovation in last period's income. The optimal response to any errors made the previous time period is

$$(1+\beta)(1+r)\left[\xi_{it-1} + \frac{r}{1+r}\mu_{it-1}\right]$$

$$= (1+r)(1+\beta)\left[1-\frac{r}{1+r}\right]\xi_{it-1}$$

   this response is now multiplied by $(1-\beta)$ because the adjustment to permanent income innovations was multiplied by that factor the previous time period.

Adding the two responses gives the total consumption response

$$\begin{aligned}\Delta c_{it} &= \beta\Delta y_{it} + (1-\beta)\left(\omega v_{it} + r\delta\left(1-\omega\right)v_{it}\right) \\ &\quad + (1-\beta)(1+r)\left[1-\frac{r}{1+r}\right]\xi_{it-1}\end{aligned} \tag{4.26}$$

If $\beta = 0$ we would have obtained a martingale with respect to the history of individual income and consumption, i.e. equation (4.15). Hence, as in the no information case, the individual consumption equation has features that stem from Flavin's framework. We see that aggregation across all individuals in the economy yields

$$\begin{aligned}\Delta c_t &= \beta \Delta y_t + (1-\beta)(\omega \varepsilon_t + r\delta(1-\omega)\varepsilon_t) \\ &\quad + (1-\beta)(1-\omega)\varepsilon_{t-1} \\ &= \beta \Delta y_t + (1-\beta)\left(\frac{\omega + r}{1+r}\right)\varepsilon_t \\ &\quad + (1-\beta)(1-\omega)\varepsilon_{t-1}\end{aligned} \tag{4.27}$$

We have an $MA(1)$ model for consumption changes and an innovation in income term. This model can explain the orthogonality failures in the REPI encountered by Campbell, Campbell and Deaton and Flavin and the failures of the excess sensitivity hypothesis encountered in the previous chapter. Campbell's first-order VAR representation of income changes and savings has the form[11],

$$\begin{bmatrix} \Delta y_t \\ s_t \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ -(1-\beta)(1-\omega) & (1+r) \end{bmatrix} \begin{bmatrix} \Delta y_{t-1} \\ s_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_t \\ (1-\beta)\left(1-\frac{\omega+r}{1+r}\right)\varepsilon_t \end{bmatrix}$$

[11] Take the following equality

$$s_t \equiv \Delta y_t + (1+r)s_{t-1} - \Delta c_t$$

and substitute our definition of consumption and labour income

$$\begin{aligned} s_t &= \Delta y_t + (1+r)s_{t-1} - \beta \Delta y_t \\ &\quad - (1-\beta)\left(\frac{\omega+r}{1+r}\right)\varepsilon_t \\ &\quad - (1-\beta)(1-\omega)\varepsilon_{t-1} \\ &= (1-\beta)\Delta y_t + (1+r)s_{t-1} \\ &\quad - (1-\beta)\left(\frac{\omega+r}{1+r}\right)\varepsilon_t \\ &\quad - (1-\beta)(1-\omega)\varepsilon_{t-1} \end{aligned}$$

From our definition of labour income, $\Delta y_t = \varepsilon_t$, write the last term as $-(1-\beta)(1-\omega)\Delta y_{t-1}$ and note that the error term can be re-expressed using current labour income innovations as $(1-\beta)\left(1-\frac{\omega+r}{1+r}\right)\varepsilon_t$, thus we have the savings term in the VAR.

Clearly, this VAR framework violates the orthogonality restrictions on equation (4.19), $a_1 = c_1$ and $d_1 = (1+r)\, b_1$, and the excess sensitivity restrictions, $a_1 = (1-\beta)\, c_1$ and $d_1 = (1+r) + (1-\beta)\, b_1$. Therefore the information problem cannot be overcome. The model is nonetheless capable of explaining the sensitivity and smoothness phenomena. Again it is straight-forward to see that sensitivity will arise if $\beta \neq 0$ and smoothness if inequality (4.6) holds.

A More General Income Process

The overall results do not change much when more complex income processes are considered provided information about the aggregate labour income shock becomes available after one period. Take the following income process to be the true one at the individual level,

$$\Delta y_{it} = v_{it} + \overline{\phi}\,(L)\, \varepsilon_t + \overline{\theta}\,(L)\, u_{it} \tag{4.28}$$

where $\overline{\phi}\,(L) = \sum_{i=1}^{\infty} \phi_i L^i$ and $\overline{\theta}\,(L) = \sum_{i=1}^{\infty} \theta_i L^i$. The response of consumption to new information is made up of two parts

1. The response to new innovations in income. Again, individuals are not able to distinguish between contemporary aggregate and individual income innovations. Since the actual innovation in income is $v_{it} = \varepsilon_t + u_{it}$ we may view the following as

   $$\phi_j \varepsilon_t + \theta_j u_{it} = \phi_j \omega v_{it} + \theta_j\,(1-\omega)\, v_{it} \;\; i = 0, 1, 2, ... \;\; where \;\; \phi_0 = \theta_0 = 1$$

   since $\omega v_{it} \approx \varepsilon_t$; $\omega = \sigma_\varepsilon^2 / \left(\sigma_\varepsilon^2 + \sigma_u^2\right)$. Recall that the innovation in permanent income is $\Delta y_{it}^p = \frac{r}{1+r} \sum_{j=0}^{\infty} \left(\frac{1}{1+r}\right)^j \Delta E_{it} y_{it+j}$ and given the full information income process (4.20), we have

   $$\frac{r}{1+r} \sum_{j=0}^{\infty} \left(\frac{1}{1+r}\right)^j \Delta E_{it} y_{it+j} =$$

   $$= r\delta\, [\omega v_{it} + (1-\omega)\, v_{it}] + r\delta^2\, [\omega v_{it} + (1-\omega)\, v_{it}] + r\delta^2\, [\phi_1 \omega v_{it} + \theta_1\,(1-\omega)\, v_{it}] +$$

$$r\delta^3 \left[\omega v_{it} + (1-\omega) v_{it}\right] + r\delta^3 \left[\phi_1 \omega v_{it} + \theta_1 (1-\omega) v_{it}\right] + r\delta^3 \left[\phi_2 \omega v_{it} + \theta_2 (1-\omega) v_{it}\right] + \cdots$$

$$= \left[\omega + (1-\omega)\right] v_{it} + \delta \left[\phi_1 \omega + \theta_1 (1-\omega)\right] v_{it} + \delta^2 \left[\phi_2 \omega + \theta_2 (1-\omega)\right] v_{it} + \cdots$$

$$= \left[\phi(\delta) \omega + \theta(\delta)(1-\omega)\right] v_{it}$$

Given the excess sensitivity hypothesis (4.5), the adjustment to an innovation in current income is

$$\beta \Delta y_{it} + (1-\beta)\left[\phi(\delta)\omega + \theta(\delta)(1-\omega)\right] v_{it}$$

2. Response to new information about the innovation in last period's income.

i) The error made in predicting the permanent component last time period was

$$\xi_{it-1} = \phi(\delta)\varepsilon_{t-1} - \phi(\delta)\omega v_{it-1}$$

The second part on the right hand side in the above equation denotes the adjustment made by the individual in the last time period. The first part denotes what the adjustment should have been.

ii) The error made in predicting the transitory component last time period was

$$\mu_{it-1} = \theta(\delta) u_{it-1} - \theta(\delta)(1-\omega) v_{it-1}$$

$$= \theta(\delta)\left[-(1-\omega)\varepsilon_{t-1} + \omega u_{it-1}\right]$$

The response of consumption to the errors made in the last time period is

$$(1+r)\left[\xi_{it-1} + \mu_{it-1}\right] =$$

$$= (1+r)\{\phi(\delta)[(1-\omega)\varepsilon_{t-1} - \omega u_{it-1}] + \theta(\delta)[-(1-\omega)\varepsilon_{t-1} + \omega u_{it-1}]\}$$

$$= (1+r)\{\phi(\delta) - \theta(\delta)\}\zeta_{it-1}$$

where $\zeta_{it-1} = (1-\omega)\varepsilon_{t-1} - \omega u_{it-1}$[12]. Adding 1) and 2) together

$$\begin{aligned} \Delta c_{it} &= \beta \Delta y_{it} + (1-\beta)[\phi(\delta)\omega + \theta(\delta)(1-\omega)](\varepsilon_t + u_{it}) \\ &\quad + (1+r)(1-\beta)\{\phi(\delta) - \theta(\delta)\}((1-\omega)\varepsilon_{t-1} - \omega u_{it-1}) \end{aligned} \tag{4.29}$$

Aggregating yields

$$\begin{aligned} \Delta c_t &= \beta \Delta y_t + (1-\beta)[\phi(\delta)\omega + \theta(\delta)(1-\omega)]\varepsilon_t \\ &\quad + (1+r)(1-\beta)\{\phi(\delta) - \theta(\delta)\}(1-\omega)\varepsilon_{t-1} \end{aligned} \tag{4.30}$$

The significance of the terms ε_t and ε_{t-1} is a characteristic of the model of lagged information when aggregate information is available after one time period and errors can be corrected. Hence the significance and interpretation of both errors; ε_t relates to the optimal response to new information this period and ε_{t-1} relates to corrections made in predicting aggregate shocks the previous time period. The introduction of the term for the innovation in income stems from the excess sensitivity hypothesis and its interpretation as the inability of consumers to optimally adjust to their changes in permanent income. This is one of the equations that will be estimated using time series data.

[12]Note that there are no interest terms $\left(\frac{r}{1+r}\right)$ in front of the response to the individual error made in the previous time period (μ_{it-1}). This is due to the fact that the individual component as defined in the general income process may have permanent effects upon labour income since it is not necessarily preceeded by $(1-L)$.

4.4 Introducing Imperfect Information to a Model of Partial Adjustment

4.4.1 Incomplete Information

A Simple Income Process

As in previous sections we assume that when agents are not able to distinguish between aggregate and individual shocks to income, their income process looks like (4.9). Given this income process, the 'evolution' of permanent income equation is

$$y_{it}^{p} = [1+r]\, y_{it-1}^{p} - rc_{it-1} + (1-\theta\delta)\,\eta_{it} \tag{4.31}$$

since $r\sum_{0}^{\infty}\delta^{\tau+1}\Delta E_t y_{it+\tau} = (1-\theta\delta)\,\eta_{it}$. From (4.1) the change in consumption can be expressed as

$$\Delta c_{it} = E_{it-n}y_{it}^{p} - E_{it-n-1}y_{it-1}^{p} + \sum_{0}^{n-1}\gamma_j \Delta E_{it-j}y_{it}^{p} - \sum_{0}^{n-1}\gamma_j \Delta E_{it-j-1}y_{it-1}^{p} \tag{4.32}$$

and given (4.31) the following are easily derived:

$$\Delta E_{it}y_{it}^{p} = (1-\theta\delta)\,\eta_{it}$$

$$\Delta E_{it-1}y_{it}^{p} = [(1+r) - r\gamma_0]\left[(1-\theta\delta)\,\eta_{it-1}\right]$$

$$\Delta E_{it-2}y_{it}^{p} = [(1+r) - r\gamma_1]\,[(1+r) - r\gamma_0]\left[(1-\theta\delta)\,\eta_{it-2}\right]$$

and

$$\Delta E_{it-1}y_{it-1}^{p} = (1-\theta\delta)\,\eta_{it-1}$$

$$\Delta E_{it-2}y_{it-1}^{p} = [(1+r) - r\gamma_0]\left[(1-\theta\delta)\,\eta_{it-2}\right]$$

$$\Delta E_{it-3}y^{p}_{it-1}=\left[(1+r)-r\gamma_1\right]\left[(1+r)-r\gamma_0\right]\left[(1-\theta\delta)\,\eta_{it-3}\right]$$

and

$$E_{it-n}y^{p}_{it}-E_{it-n-1}y^{p}_{it-1}=\prod_{0}^{n-1}\zeta_j\,(1-\theta\delta)\,\eta_{it-n}$$

Substituting these expressions into (4.32) yields the change in individual consumption when agents do not distinguish between aggregate and individual information

$$\begin{aligned}\Delta c_{it}&=\gamma_0\,(1-\theta\delta)\,\eta_{it}+\left[\gamma_1\zeta_0-\gamma_0\right](1-\theta\delta)\,\eta_{it-1}+\left[\gamma_2\zeta_1-\gamma_1\right]\zeta_0\,(1-\theta\delta)\,\eta_{it-2}+\cdots\\&+\left[\gamma_{n-1}\zeta_{n-2}-\gamma_{n-2}\right](1-\theta\delta)\prod_{j=0}^{n-3}\zeta_j\eta_{it-n+1}+\left[\zeta_{n-1}-\gamma_{n-1}\right](1-\theta\delta)\prod_{j=0}^{n-2}\zeta_j\eta_{it-n}\end{aligned}$$

where $\zeta_j=\left[1+r\left(1-\gamma_j\right)\right]$. The principal flavour of the partial adjustment model prevails at the individual level. The change in consumption is still a moving average process determined by the time span required to ensure that any component of current permanent income has its full effect on current consumption. Aggregation yields interesting results, use the income process to rewrite $\eta_{it}=\frac{\Delta y_{it}}{(1-\theta L)}$ provided $(1-\theta L)$ is invertible. We substitute this last expression into the consumption equation to obtain:

$$\begin{aligned}\Delta c_{it}&=\gamma_0\,(1-\theta\delta)\,\tfrac{\Delta y_{it}}{(1-\theta L)}+\left[\gamma_1\zeta_0-\gamma_0\right](1-\theta\delta)\,\tfrac{\Delta y_{it-1}}{(1-\theta L)}\\&+\left[\gamma_2\zeta_1-\gamma_1\right]\zeta_0\,(1-\theta\delta)\,\tfrac{\Delta y_{it-2}}{(1-\theta L)}+\cdots+\left[\gamma_{n-1}\zeta_{n-2}-\gamma_{n-2}\right](1-\theta\delta)\prod_{j=0}^{n-3}\zeta_j\tfrac{\Delta y_{it-n+1}}{(1-\theta L)}\\&+\left[\zeta_{n-1}-\gamma_{n-1}\right](1-\theta\delta)\prod_{j=0}^{n-2}\zeta_j\tfrac{\Delta y_{it-n}}{(1-\theta L)}\end{aligned}\tag{4.33}$$

Using (4.23) and aggregating (4.33) gives

$$\begin{aligned}(1-\theta L)\,\Delta c_t &= \gamma_0\,(1-\theta\delta)\,\Delta y_t + [\gamma_1\zeta_0 - \gamma_0]\,(1-\theta\delta)\,\Delta y_{t-1} \\ &\quad + [\gamma_2\zeta_1 - \gamma_1]\,\zeta_0\,(1-\theta\delta)\,\Delta y_{t-2} + \cdots \\ &\quad + \left[\gamma_{n-1}\zeta_{n-2} - \gamma_{n-2}\right](1-\theta\delta)\prod_{0}^{n-3}\zeta_j\Delta y_{t-n+1} \\ &\quad + \left[\zeta_{n-1} - \gamma_{n-1}\right](1-\theta\delta)\prod_{0}^{n-2}\zeta_j\Delta y_{t-n}\end{aligned}$$

which we can write as

$$\begin{aligned}&(1-\theta L)\,\Delta c_t = \gamma_0\,(1-\theta\delta)\,\varepsilon_t + [\gamma_1\zeta_0 - \gamma_0]\,(1-\theta\delta)\,\varepsilon_{t-1} \\ &+ [\gamma_2\zeta_1 - \gamma_1]\,\zeta_0\,(1-\theta\delta)\,\varepsilon_{t-2} + \cdots \\ &+ \left[\gamma_{n-1}\zeta_{n-2} - \gamma_{n-2}\right](1-\theta\delta)\prod_{j=0}^{n-3}\zeta_j\varepsilon_{t-n+1} \\ &+ \left[\zeta_{n-1} - \gamma_{n-1}\right](1-\theta\delta)\prod_{j=0}^{n-2}\zeta_j\varepsilon_{t-n}\end{aligned} \tag{4.34}$$

We now have an $ARMA(1,n)$ model for the change in consumption. Imperfect information introduces an AR term to the innovation in consumption, a characteristic of models of no information and it therefore has a similar interpretation to that one provided above: the AR component is the result of the lack of aggregate information that agents are presumed to have about the economy and of their subsequent adjustments to consumption given further surprises to their expected labour income. The MA components are independent of that information; they come from the costly adjustments that have to be endured following shocks to income (and the information about future levels of labour income and hence permanent income that is inherent in those shocks). In this model, adjustment is necessary and it is an ongoing process[13]: as information becomes available and is processed by agents, previous adjustments will turn out to be incorrect because agents were unable to distinguish between aggregate and individual income shocks. Hence, to be close to their desired level of consumption agents must account for previous surprises associated with their lack of knowledge about aggregate events. These

[13] An interesting intuition may be that consumers may have to change their consumption because their previous habits proved to be wrong.

surprises would have resulted in a level of consumption last period that would not have been close to the desired level. Hence adjustment is necessary and ongoing.

More General Income Processes

The properties encountered in equation (4.34) are found in the more general case. As before, assume income has the following form, $\Delta y_{it} = A(L)\,\eta_{it}$ hence, $r\sum_{0}^{\infty}\delta^{\tau+1}\Delta E_{it}y_{it+\tau} = A(\delta)\,\eta_{it}$ and the evolution of permanent income is

$$y_{it}^{p} = [1+r]\,y_{it-1}^{p} - rc_{it-1} + A(\delta)\,\eta_{it}$$

From this equation the change in individual consumption is

$$\begin{aligned}\Delta c_{it} &= \gamma_0 A(\delta)\,\eta_{it} + [\gamma_1\zeta_0 - \gamma_0]\,A(\delta)\,\eta_{it-1} + [\gamma_2\zeta_1 - \gamma_1]\,\zeta_0 A(\delta)\,\eta_{it-2} \\ &+ \cdots + [\gamma_{n-1}\zeta_{n-2} - \gamma_{n-2}]\,A(\delta)\prod_{j=0}^{n-3}\zeta_j\eta_{it-n+1} + [\zeta_{n-1} - \gamma_{n-1}]\,A(\delta)\prod_{j=0}^{n-2}\zeta_j\eta_{it-n}\end{aligned} \tag{4.35}$$

The ζ_i's are defined as before. An econometrician working at the aggregate level would find interesting results: assuming $A(L)$ is invertible and aggregating the individual components as before, yields the consumption equation

$$\begin{aligned}\Delta c_t &= \gamma_0\frac{\Delta y_t}{A(L)} + [\gamma_1\zeta_0 - \gamma_0]\,\frac{\Delta y_{t-1}}{A(L)} + [\gamma_2\zeta_1 - \gamma_1]\,\zeta_0\frac{\Delta y_{t-2}}{A(L)} + \cdots \\ &+ [\gamma_{n-1}\zeta_{n-2} - \gamma_{n-2}]\prod_{0}^{n-3}\zeta_j\frac{\Delta y_{t-n+1}}{A(L)} + [\zeta_{n-1} - \gamma_{n-1}]\prod_{0}^{n-2}\zeta_j\frac{\Delta y_{t-n}}{A(L)}\end{aligned} \tag{4.36}$$

which we can re-write as

$$\begin{aligned}A(L)\,\Delta c_t &= \gamma_0\phi(L)\,\varepsilon_t + [\gamma_1\zeta_0 - \gamma_0]\,\phi(L)\,\varepsilon_{t-1} + [\gamma_2\zeta_1 - \gamma_1]\,\zeta_0\phi(L)\,\varepsilon_{t-2} + \cdots \\ &+ [\gamma_{n-1}\zeta_{n-2} - \gamma_{n-2}]\,\phi(L)\prod_{j=0}^{n-3}\zeta_j\varepsilon_{t-n+1} + [\zeta_{n-1} - \gamma_{n-1}]\,\phi(L)\prod_{j=0}^{n-2}\zeta_j\varepsilon_{t-n}\end{aligned} \tag{4.37}$$

The interesting result is that we have an $ARMA(p,q)$ where the value of p is determined by the number of lags in the individual income process. This is due to the fact that when agents observe a shock and adjust accordingly, they will be surprised again the following time period and will have to readjust their consumption pattern accordingly. The value of q is the sum of the n lags associated with adjustments to information that are less than full, plus the d lags in the (individual) income process i.e. $q = n+d$. As we have seen before, the AR components are a characteristic of the model of incomplete information. The (extra) lags from the income process play a role in the MA components or the change in consumption (even when adjustment was supposed to be for n periods) because of the information deficiencies; all the previous shocks to income still play a role as individuals had not adjusted to them in the correct manner when these shocks occurred. This is one of the equations that can be tested using time series data.

4.4.2 Lagged Information About Aggregate Shocks

A Simple Income Process

Assume that the income process is given by (4.12). Therefore, the information that the consumer gets every period is made up of two parts:

1. v is the innovation contained in individual income at time t. Consumers cannot distinguish at any particular period how an innovation in income is made up of the aggregate and individual components. The agent attributes part of the current period innovation to each component given their relative variances. Errors are therefore made when computations take place.

2. The consumer gets information from the lagged aggregate shock, so that she will be able to correct the error made in the previous period when she attributed the innovation to its corresponding components.

These two concepts are introduced into Attfield et al.'s model using the evolution of permanent income equation (4.2). Introducing lagged aggregate information changes the evolution of permanent income which now takes the following form

$$y_{it}^{p} = (1+r)\left[y_{it-1}^{p} + \vartheta_{it-1}\right] - rc_{it-1} + r\delta \sum_{0}^{\infty} \delta^{\tau} \Delta E_{it} y_{it+\tau} \qquad (4.38)$$

where ϑ_{it-1} denotes the sum of the *negative* of the errors made in predicting the aggregate and individual innovations in income and $r\delta \sum_{0}^{\infty} \delta^{\tau} \Delta E_{it} y_{it+\tau}$ continues to represent the optimal response to income innovations. Note that this form is consistent with Pischke/Goodfriend's framework[14].

Why this form? It is clear that permanent income depends on the unpredictable innovations to the labour income process. These innovations are the standard innovations encountered in the permanent income literature and represent the innovation in permanent income when there are no unanticipated capital gains. What is different now (as it is in Pischke and Goodfriend's papers) is that at time t, agents are not able to distinguish what proportion of the overall innovation in labour income is economy-wide and thus more persistent and which one is individual. Rational agents attempt to ascertain the exact proportion of that innovation that is expected to be the aggregate one. That information will affect permanent income.

Two factors drive the dynamics of the evolution in permanent income:

a) $r\delta \sum_{0}^{\infty} \delta^{\tau} \Delta E_{it} y_{it+\tau} = \left[\omega v_{it} + (1-\omega)\left(\frac{r}{1+r}\right) v_{it}\right]$

where $\omega = \frac{\sigma_{\varepsilon}^2}{\sigma_{\varepsilon}^2 + \sigma_u^2}$, thus $\omega v_{it} \approx \varepsilon_t$ and $(1-\omega) v_{it} \approx u_{it}$.

b) At the same time, information about the accuracy of the calculations made the previous time period becomes available. This information allows the consumer to readjust its behaviour whenever a mistake is made in the signal extraction problem the previous time period. Why would the consumer wish to readjust his or her behaviour? In an infinite horizon problem, and provided the consumer reacts differently to (more persistent) aggregate and individual

[14] Assume that the permanent income hypothesis holds;

$$c_{it} = y_{it}^{p}$$

we have

$$y_{it}^{p} = (1+r)\left[y_{it-1}^{p} + \vartheta_{it-1}\right] - rc_{it-1} + r\delta \sum_{0}^{\infty} \delta^{\tau} \Delta E_t y_{it+\tau}$$

and so we can rewrite this last equation as

$$\begin{aligned} c_{it} &= (1+r)\left[c_{it-1} + \vartheta_{it-1}\right] - rc_{it-1} + r\delta \sum_{0}^{\infty} \delta^{\tau} \Delta E_t y_{it+\tau} \\ &= c_{it-1} + \xi_{it-1} + r\delta \sum_{0}^{\infty} \delta^{\tau} \Delta E_t y_{it+\tau} \end{aligned}$$

and so we have Pischke's equation (18) for the simple income case (4.12) as $(1+r)\vartheta_{it-1} = \xi_{it-1}$.

innovations in income, it is clear that the consumer will correct any mistakes made when undertaking the signal extraction problem. If the consumer thought that the innovation the previous time period was mostly a permanent one, the consumer would have expected his or her permanent income to rise and would have, as a result, consumed more. As new information becomes available the next time period, the consumer is able to see that in the previous time period he or she might have under-consumed. This under-consumption increases his or her permanent income forever and so the agent corrects his or her behaviour in order to maximise utility. It is less clear what the behaviour of the agent would be in the case of finite lifetimes.

How do we define this error ϑ_{it-1}? Here we follow Pischke; by defining the negative of the error made when predicting the aggregate component the previous time period as

$$\xi_{it-1} = \varepsilon_{t-1} - \omega v_{it-1} = (1-\omega)\,\varepsilon_{t-1} - \omega u_{it-1}$$

and define the negative of the error in predicting the individual component the previous time period as

$$\begin{aligned} \mu_{it-1} &= u_{it-1} - (1-\omega)\,v_{it-1} \\ &= -\xi_{it-1} \end{aligned}$$

We therefore have

$$\begin{aligned} (1+r)\,\vartheta_{it-1} &= (1+r)\left[\xi_{it-1} + \left(\frac{r}{1+r}\right)\mu_{it-1}\right] \\ &= \xi_{it-1} \end{aligned}$$

so that the evolution of permanent income equation becomes

$$y_{it}^{p} = (1+r)\,y_{it-1}^{p} + \xi_{it-1} - rc_{it-1} + \left[\omega v_{it} + (1-\omega)\left(\frac{r}{1+r}\right)v_{it}\right]$$

From this equation the following are easily derived

$$\Delta E_{it} y_{it}^{p} = \left[\omega v_{it} + (1-\omega)\left(\frac{r}{1+r}\right)v_{it}\right] + \xi_{it-1}$$

$$\Delta E_{it-1} y^p_{it} = \zeta_0 \left\{ \left[\omega v_{it-1} + (1-\omega) \left(\frac{r}{1+r} \right) v_{it-1} \right] + \xi_{it-2} \right\}$$

$$\Delta E_{it-2} y^p_{it} = \zeta_1 \zeta_0 \left\{ \left[\omega v_{it-2} + (1-\omega) \left(\frac{r}{1+r} \right) v_{it-2} \right] + \xi_{it-3} \right\}$$

since ξ_{it-j-1} only becomes known at $t-j$, and

$$\Delta E_{it-1} y^p_{it-1} = \left[\omega v_{it-1} + (1-\omega) \left(\frac{r}{1+r} \right) v_{it-1} \right] + \xi_{it-2}$$

$$\Delta E_{it-2} y^p_{it-1} = \zeta_0 \left\{ \left[\omega v_{it-2} + (1-\omega) \left(\frac{r}{1+r} \right) v_{it-2} \right] + \xi_{it-3} \right\}$$

$$\Delta E_{it-3} y^p_{it-1} = \zeta_1 \zeta_0 \left\{ \left[\omega v_{it-3} + (1-\omega) \left(\frac{r}{1+r} \right) v_{it-3} \right] + \xi_{it-4} \right\}$$

and

$$E_{it-n} y^p_{it} - E_{it-n-1} y^p_{it-1} = \prod_0^{n-1} \zeta_j \left\{ \left[\omega v_{it-n} + (1-\omega) \left(\frac{r}{1+r} \right) v_{it-n} \right] + \xi_{it-n-1} \right\}$$

Hence, individual consumption changes are given by

$$\begin{aligned} \Delta c_{it} = {} & \gamma_0 \left\{ \left[\omega v_{it} + (1-\omega) \left(\frac{r}{1+r} \right) v_{it} \right] + \xi_{it-1} \right\} \\ & + [\zeta_0 \gamma_1 - \gamma_0] \left\{ \left[\omega v_{it-1} + (1-\omega) \left(\frac{r}{1+r} \right) v_{it-1} \right] + \xi_{it-2} \right\} \\ & + \zeta_0 [\zeta_1 \gamma_2 - \gamma_1] \left\{ \left[\omega v_{it-2} + (1-\omega) \left(\frac{r}{1+r} \right) v_{it-2} \right] + \xi_{it-3} \right\} + \cdots \\ & + [\zeta_{n-2} \gamma_{n-1} - \gamma_{n-2}] \prod_0^{n-3} \zeta_j \left\{ \left[\omega v_{it-n+1} + (1-\omega) \left(\frac{r}{1+r} \right) v_{it-n+1} \right] + \xi_{it-n} \right\} \\ & + [\zeta_{n-1} - \gamma_{n-1}] \prod_0^{n-2} \zeta_j \left\{ \left[\omega v_{it-n} + (1-\omega) \left(\frac{r}{1+r} \right) v_{it-n} \right] + \xi_{it-n-1} \right\} \end{aligned} \tag{4.39}$$

The overall flavour of the partial adjustment model is maintained in as far as the changes in consumption follow an MA process. The adjustment in consumption to new information does not depend on the innovation in labour income alone, but also depends on whatever new information becomes available about the accuracy of the previous period's signal extraction problem (i.e. the ξ_{it} terms). This information is new and should not be related to previous innovations in the labour income process. Note that the consumption equation is now an $MA(n+1)$ process instead of an $MA(n)$.

Aggregation is fairly straightforward, as the individual/transitory components sum to zero[15]

$$\begin{aligned}
\Delta c_t &= \gamma_0 \left\{ \left[\left(\frac{\omega + r}{1+r} \right) \varepsilon_t \right] + (1-\omega)\, \varepsilon_{t-1} \right\} \\
&\quad + [\zeta_0 \gamma_1 - \gamma_0] \left\{ \left[\left(\frac{\omega + r}{1+r} \right) \varepsilon_{t-1} \right] + (1-\omega)\, \varepsilon_{t-2} \right\} \\
&\quad + \zeta_0 [\zeta_1 \gamma_2 - \gamma_1] \left\{ \left[\left(\frac{\omega + r}{1+r} \right) \varepsilon_{t-2} \right] + (1-\omega)\, \varepsilon_{t-3} \right\} + \cdots \\
&\quad + [\zeta_{n-2} \gamma_{n-1} - \gamma_{n-2}] \prod_0^{n-3} \zeta_j \left\{ \left[\left(\frac{\omega + r}{1+r} \right) \varepsilon_{t-n+1} \right] + (1-\omega)\, \varepsilon_{t-n} \right\} \\
&\quad + [\zeta_{n-1} - \gamma_{n-1}] \prod_0^{n-2} \zeta_j \left\{ \left[\left(\frac{\omega + r}{1+r} \right) \varepsilon_{t-n} \right] + (1-\omega)\, \varepsilon_{t-n-1} \right\}
\end{aligned} \tag{4.40}$$

The reaction to aggregate shocks in the economy differs with respect to Attfield et al.'s model, in the signal extraction coefficient ω. Only if $\omega = 1$ we have the original response to labour income innovations that was present in the partial adjustment model. The lower this coefficient is, the lower the adjustment to recent innovations to aggregate income. This means that adjustment is even slower in this case compared to the original model of Attfield et al.. The overall flavour of the original Attfield et al. model seems to remain intact although we now have an $MA(n+1)$ instead of an $MA(n)$ process.

[15] Recall that

$$v_{it} = \varepsilon_t + u_{it}$$

and that

$$\xi_{it-1} = (1-\omega)\, \varepsilon_{t-1} - \omega u_{it-1}$$

More General Income Processes

The main flavour of the model is not affected by the introduction of a more general income process like (4.28). Assume that the evolution of permanent income equation is still the same

$$y^p_{it} = (1+r)\left[y^p_{it-1} + \vartheta_{it-1}\right] - rc_{it-1} + r\delta\sum_0^\infty \delta^\tau \Delta E_t y_{it+\tau}$$

where $r\delta\sum_0^\infty \delta^\tau \Delta E_t y_{it+\tau} = \left[\phi(\delta)\omega + (1-\omega)\theta(\delta)\right] v_{it}$. The error in predicting the aggregate component is given by $\xi_{it-1} = \phi(\delta)\varepsilon_{t-1} - \phi(\delta)\omega v_{it-1} = \phi(\delta)\zeta_{it-1}$and the error in predicting the individual component is $\mu_{it-1} = \theta(\delta) u_{it-1} - (1-\omega)\theta(\delta) v_{it-1} = \theta(\delta)\zeta_{it-1}$ where $\zeta_{it-1} = [(1-\omega)\varepsilon_{t-1} - \omega v_{it-1}]$. This time

$$\vartheta_{it-1} = \left[\phi(\delta) - \theta(\delta)\right]\zeta_{it-1}$$

and so we have

$$y^p_{it} = (1+r)\left[y^p_{it-1} + \left[\phi(\delta) - \theta(\delta)\right]\zeta_{it-1}\right] - rc_{it-1} + \left[\phi(\delta)\omega + (1-\omega)\theta(\delta)\right] v_{it}$$

(note $y^p_{it-1} = (1+r)\left[y^p_{it-2} + [\phi(\delta) - \theta(\delta)]\zeta_{it-2}\right] - rc_{it-2} + [\phi(\delta)\omega + (1-\omega)\theta(\delta)]v_{it-1}$). Thus, individual consumption changes are given by

$$
\begin{aligned}
\Delta c_{it} \;=\; & \gamma_0\left\{[\phi(\delta)\omega + (1-\omega)\theta(\delta)]v_{it} + (1+r)[\phi(\delta) - \theta(\delta)]\zeta_{it-1}\right\} \\
& + [\zeta_0\gamma_1 - \gamma_0][\phi(\delta)\omega + (1-\omega)\theta(\delta)]v_{it-1} \\
& + [\zeta_0\gamma_1 - \gamma_0](1+r)[\phi(\delta) - \theta(\delta)]\zeta_{it-2} \\
& + \zeta_0[\zeta_1\gamma_2 - \gamma_1][\phi(\delta)\omega + (1-\omega)\theta(\delta)]v_{it-2} \\
& + \zeta_0[\zeta_1\gamma_2 - \gamma_1](1+r)[\phi(\delta) - \theta(\delta)]\zeta_{it-3} + \cdots \\
& + \left[\zeta_{n-2}\gamma_{n-1} - \gamma_{n-2}\right]\prod_0^{n-3}\zeta_j[\phi(\delta)\omega + (1-\omega)\theta(\delta)]v_{it-n+1} \\
& + \left[\zeta_{n-2}\gamma_{n-1} - \gamma_{n-2}\right]\prod_0^{n-3}\zeta_j(1+r)[\phi(\delta) - \theta(\delta)]\zeta_{it-n} \\
& + \left[\zeta_{n-1} - \gamma_{n-1}\right]\prod_0^{n-2}\zeta_j[\phi(\delta)\omega + (1-\omega)\theta(\delta)]v_{it-n} \\
& + \left[\zeta_{n-1} - \gamma_{n-1}\right]\prod_0^{n-2}\zeta_j(1+r)[\phi(\delta) - \theta(\delta)]\zeta_{it-n-1}
\end{aligned}
\tag{4.41}
$$

Aggregation yields

$$
\begin{aligned}
\Delta c_t \;=\; & \gamma_0\{[\phi(\delta)\omega + (1-\omega)\theta(\delta)]\varepsilon_t + (1+r)[\phi(\delta) - \theta(\delta)](1-\omega)\varepsilon_{t-1}\} \\
& + [\zeta_0\gamma_1 - \gamma_0]\{[\phi(\delta)\omega + (1-\omega)\theta(\delta)]\varepsilon_{t-1} + (1+r)[\phi(\delta) - \theta(\delta)](1-\omega)\varepsilon_{t-2}\} \\
& + \zeta_0[\zeta_1\gamma_2 - \gamma_1]\{[\phi(\delta)\omega + (1-\omega)\theta(\delta)]\varepsilon_{t-2} + (1+r)[\phi(\delta) - \theta(\delta)](1-\omega)\varepsilon_{t-3}\} + \cdots \\
& + \left[\zeta_{n-2}\gamma_{n-1} - \gamma_{n-2}\right]\prod_0^{n-3}\zeta_j\{[\phi(\delta)\omega + (1-\omega)\theta(\delta)]\varepsilon_{t-n+1}\} \\
& + \left[\zeta_{n-2}\gamma_{n-1} - \gamma_{n-2}\right]\prod_0^{n-3}\zeta_j\{(1+r)[\phi(\delta) - \theta(\delta)](1-\omega)\varepsilon_{t-n}\} \\
& + \left[\zeta_{n-1} - \gamma_{n-1}\right]\prod_0^{n-2}\zeta_j\{[\phi(\delta)\omega + (1-\omega)\theta(\delta)]\varepsilon_{t-n}\} \\
& + \left[\zeta_{n-1} - \gamma_{n-1}\right]\prod_0^{n-2}\zeta_j\{(1+r)[\phi(\delta) - \theta(\delta)](1-\omega)\varepsilon_{t-n-1}\}
\end{aligned}
\tag{4.42}
$$

In this case we also have an $MA(n+1)$ model for the change in consumption. Excess sensitivity and smoothness ought to occur. Note that in contrast to the no information case, the specific form of the income dynamics does not play a role in models of lagged information (the simple income case yields a specification that is similar to the more general income case); what matters is the relative persistence of aggregate and individual shocks as measured by $\phi(1)$ and $\theta(1)$. This is a result that Pischke found in his work (pp. 818) and that is also consistent with the introduction of lagged information to the excess sensitivity framework. We shall test this equation using time series data.

4.5 Introducing the Excess Sensitivity Hypothesis to a Model of Partial Adjustment

In this section we examine the effects of the excess sensitivity hypothesis in a model of partial adjustment. The excess sensitivity model presumes that agents over-react to current income, whilst the partial adjustment model asserts that agent's decisions depend on how far away information became available. The interesting question is whether the consumer's behaviour would depend on current income or on past information?

A way of introducing the excess sensitivity hypothesis into the model of partial adjustment would be to assume that consumption has the following form

$$c_t = \beta y_t^T + E_{t-n} y_t^p + \sum_{i=0}^{n-1} \gamma_i \Delta E_{t-i} y_t^p \qquad (4.43)$$

where β is the marginal propensity to consume out of transitory income, y_t^T and y_t^p refer to transitory and permanent income respectively. Assume that $0 \leq \gamma_0 \leq \gamma_1 \leq \cdots \leq \gamma_{n-1} \leq 1$. The model suggests that consumers react more forcefully to those components of current income that were predictable earlier, but at the same time their consumption is dependent on transitory income as consumers may be constrained in their borrowing. The coefficient β may serve as a measure of the inability of consumers to borrow money; the closer it is to one, the more likely are consumers to be liquidity constrained. Transitory income is defined as the residual

$$y_t^T \equiv (y_t + rA_t) - y_t^p \tag{4.44}$$

and given the definition of permanent income, we can express (4.44) as

$$y_t^T = y_t - \frac{r}{1+r}\sum_{\tau=0}^{\infty}\delta^{\tau} E_t y_{t+\tau}$$

The following equation applies regardless of the time series process that governs labour income: $y_t^p = [1+r]\, y_{t-1}^p - rc_{t-1} + r\sum_0^{\infty}\delta^{\tau+1}\Delta E_t y_{t+\tau}$ where the last term continues to indicate innovations in information about the labour income process. It is then relatively straight-forward to show that

$$\Delta E_t y_t^p = r\sum_0^{\infty}\delta^{\tau+1}\Delta E_t y_{t+\tau} \tag{4.45}$$

$$\Delta E_{t-1} y_t^p = \zeta_0 r\sum_0^{\infty}\delta^{\tau+1}\Delta E_{t-1} y_{t-1+\tau} - r\beta\Delta E_{t-1} y_{t-1}^T \tag{4.46}$$

$$\Delta E_{t-2} y_t^p = \zeta_1\left\{\zeta_0 r\sum_0^{\infty}\delta^{\tau+1}\Delta E_{t-2} y_{t-2+\tau} - r\beta\Delta E_{t-2} y_{t-2}^T\right\} - r\beta\Delta E_{t-2} y_{t-1}^T \tag{4.47}$$

$$\begin{aligned}\Delta E_{t-3} y_t^p = \zeta_2&\left\{\zeta_1\left\{\zeta_0 r\sum_0^{\infty}\delta^{\tau+1}\Delta E_{t-3} y_{t-3+\tau} - r\beta\Delta E_{t-3} y_{t-3}^T\right\} - r\beta\Delta E_{t-3} y_{t-2}^T\right\}\\ &-r\beta\Delta E_{t-3} y_{t-1}^T\end{aligned} \tag{4.48}$$

and

$$\Delta E_{t-1} y_{t-1}^p = r\sum_0^{\infty}\delta^{\tau+1}\Delta E_{t-1} y_{t+\tau-1} \tag{4.49}$$

$$\Delta E_{t-2} y_{t-1}^{P} = \zeta_0 r \sum_{0}^{\infty} \delta^{\tau+1} \Delta E_{t-2} y_{t+\tau-2} - r\beta \Delta E_{t-2} y_{t-2}^{T} \tag{4.50}$$

$$\Delta E_{t-3} y_{t-1}^{P} = \zeta_1 \left\{ \zeta_0 r \sum_{0}^{\infty} \delta^{\tau+1} \Delta E_{t-3} y_{t-3+\tau} - r\beta \Delta E_{t-3} y_{t-3}^{T} \right\} - r\beta \Delta E_{t-3} y_{t-2}^{T} \tag{4.51}$$

$$\begin{aligned} \Delta E_{t-4} y_{t-1}^{P} = \zeta_2 \left\{ \zeta_1 \left\{ \zeta_0 r \sum_{0}^{\infty} \delta^{\tau+1} \Delta E_{t-4} y_{t-4+\tau} - r\beta \Delta E_{t-4} y_{t-3}^{T} \right\} - r\beta \Delta E_{t-4} y_{t-3}^{T} \right\} \\ -r\beta \Delta E_{t-4} y_{t-2}^{T} \end{aligned} \tag{4.52}$$

again, $\zeta_i = [(1+r) - r\gamma_i]$. Note that the transitory income terms enter these equations through the consumption term in the evolution of permanent income equation. In all cases, given a difference-stationary labour income process, $y_t = A(L)\omega_t$, we have

$$r \sum_{0}^{\infty} \delta^{\tau+1} \Delta E_{t-i} y_{t+\tau-i} = A(\delta)\, \omega_{t-i} \ \forall i \tag{4.53}$$

Note that the transitory income component can be written as

$$y_{t-i}^{T} = -r \sum_{1}^{\infty} \delta^{\tau} E_{t-i} \Delta y_{t+\tau-i} \quad \forall i$$

We solve this model using different income processes. Throughout, it is assumed that labour income is difference stationary and that no components apart from labour income itself will aid in the predictions for future labour income.

4.5.1 A Simple Income Process

Consider the following income process

$$\Delta y_t = \varepsilon_t$$

From this income process and the definition of transitory income, we have that

$$y_{t-i}^{T} = 0 \quad \forall i$$

Transitory income is therefore always zero and so the excess sensitivity hypothesis collapses to become the standard model of partial adjustment in this case.

4.5.2 IMA (1,1)

The income process is

$$\Delta y_t = \varepsilon_t + \phi\varepsilon_{t-1}$$

the coefficient in the $IMA(1,1)$ process maybe positive or negative. This process has more memory than the one that we have examined above, and that has significant implications for the model in question. Note the following apply now

$$y_{t-i}^{T} = -r\delta\phi\varepsilon_{t-i} \; \forall i$$

and

$$r\sum_{0}^{\infty}\delta^{\tau+1}\Delta E_{t-i}y_{t+\tau-i} = (1+\phi\delta)\,\varepsilon_{t-i} \quad \forall i$$

Thus, we obtain the following equations in general:

$$\Delta E_{t-j}y_t^p = \prod_{1}^{j-1}\zeta_i\left[\zeta_0\,(1+\phi\delta)\,\varepsilon_{t-j} + \delta\beta\phi r^2\varepsilon_{t-j}\right]$$

and

$$\Delta E_{t-j}y_{t-1}^p = \prod_{1}^{j-2}\zeta_i\left[\zeta_0\,(1+\phi\delta)\,\varepsilon_{t-j} + \delta\beta\phi r^2\varepsilon_{t-j}\right]$$

Also, we have

$$E_{t-n}y_t^p - E_{t-n-1}y_{t-1}^p = \zeta_{n-1}\Delta E_{t-n}y_{t-1}^p$$

With these equations, we can construct the change in consumption as

$$\begin{aligned}
\Delta c_t &= -\beta\phi\delta r\varepsilon_t + \beta\phi\delta r\varepsilon_{t-1} + \zeta_{n-1}\prod_1^{n-2}\zeta_j\left[\zeta_0\left(1+\phi\delta\right)\varepsilon_{t-n} + \delta\beta\phi r^2\varepsilon_{t-n}\right] \\
&\quad +\gamma_0\left(1+\phi\delta\right)\varepsilon_t + \gamma_1\left[\zeta_0\left(1+\phi\delta\right)\varepsilon_{t-1} + \delta\beta\phi r^2\varepsilon_{t-1}\right] \\
&\quad +\gamma_2\left[\zeta_1\left\{\zeta_0\left(1+\phi\delta\right)\varepsilon_{t-2} + \delta\beta\phi r^2\varepsilon_{t-2}\right\}\right] + \cdots \\
&\quad +\gamma_{n-1}\prod_1^{n-2}\zeta_j\left[\zeta_0\left(1+\phi\delta\right)\varepsilon_{t-n+1} + \delta\beta\phi r^2\varepsilon_{t-n+1}\right] - \gamma_0\left(1+\phi\delta\right)\varepsilon_{t-1} \\
&\quad -\gamma_1\left[\zeta_0\left(1+\phi\delta\right)\varepsilon_{t-2} + \delta\beta\phi r^2\varepsilon_{t-2}\right] - \gamma_2\left[\zeta_1\left\{\zeta_0\left(1+\phi\delta\right)\varepsilon_{t-3} + \delta\beta\phi r^2\varepsilon_{t-3}\right\}\right] \\
&\quad -\gamma_{n-2}\prod_1^{n-3}\zeta_j\left[\zeta_0\left(1+\phi\delta\right)\varepsilon_{t-n+1} + \delta\beta\phi r^2\varepsilon_{t-n+1}\right] \\
&\quad -\gamma_{n-1}\prod_1^{n-2}\zeta_j\left[\zeta_0\left(1+\phi\delta\right)\varepsilon_{t-n} + \delta\beta\phi r^2\varepsilon_{t-n}\right]
\end{aligned}$$

We can re-write this equation as

$$\begin{aligned}
\Delta c_t &= \gamma_0\left(1+\phi\delta\right)\varepsilon_t - \beta\phi\delta r\varepsilon_t \\
&\quad + \left[1+\gamma_1 r\right]\beta\phi\delta r\varepsilon_{t-1} + \left[\gamma_1\zeta_0 - \gamma_0\right]\left(1+\phi\delta\right)\varepsilon_{t-1} \\
&\quad + \left[\gamma_2\zeta_1 - \gamma_1\right]\zeta_0\left(1+\phi\delta\right)\varepsilon_{t-2} + \left[\gamma_2\zeta_1 - \gamma_1\right]\delta\beta\phi r^2\varepsilon_{t-2} \\
&\quad + \left[\gamma_3\zeta_2 - \gamma_2\right]\zeta_1\zeta_0\left(1+\phi\delta\right)\varepsilon_{t-3} + \left[\gamma_3\zeta_2 - \gamma_2\right]\zeta_1\delta\beta\phi r^2\varepsilon_{t-3} + \cdots \\
&\quad + \left[\gamma_{n-1}\zeta_{n-2} - \gamma_{n-2}\right]\left[\prod_0^{n-3}\zeta_j\left(1+\phi\delta\right) + \prod_1^{n-3}\zeta_j\delta\beta\phi r^2\right]\varepsilon_{t-n+1} \\
&\quad + \left[\zeta_{n-1} - \gamma_{n-1}\right]\left[\prod_0^{n-2}\zeta_j\left(1+\phi\delta\right) + \prod_0^{n-2}\zeta_j\delta\beta\phi r^2\right]\varepsilon_{t-n}
\end{aligned} \tag{4.54}$$

The equation is similar to Attfield et al.'s original one as it follows an $MA(n)$ process. Excess sensitivity will occur in this model - as it did in Attfield et al.'s original model - because we have lagged shocks to permanent income that are likely to be correlated with lagged innovations in labour income. Since excess smoothness occurs when $var(\Delta c_t) < var(\varepsilon_t)$ and we have from (4.54) that

$$var(\Delta c_t) = var(\varepsilon_t)\sum_{0}^{n}\lambda_i^2 \; as \; cov(\varepsilon_t,\ \varepsilon_{t-i}) = 0 \; i = 0, \cdots, \ n$$

where λ is a function of the coefficients of the $MA(n)$ process, then smoothness will occur if $\sum_0^n \lambda_i^2 < 1$. The difference between (4.54) and Attfield et al.'s equation (4.3) is that we now have the transitivity income effect 'reinforcing' the original partial adjustment result provided ϕ is positive. It is unclear whether this model yields more sensitivity and smoothness than the original partial adjustment model, this depends on the characteristics of two coefficients; the MA coefficient in the income process and the excess sensitivity coefficient β. Finally, note that if we impose the following restrictions on (4.54): i) $\beta = 0$, we have the original partial adjustment model, ii) $\gamma_0 = \gamma_1 = \cdots = \gamma_{n-1} = 1$, so that we have the excess sensitivity hypothesis in its purest form and iii) $\beta = 0$ and $\gamma_0 = \gamma_1 = \cdots = \gamma_{n-1} = 1$ so that we have the original REPI.

It is interesting at this point to introduce the Pischke model of no information since it is only a trivial extension to (4.54). Assume that the individual income process (4.9) holds. As we have seen, the individual consumption response to a differenced stationary $MA(1)$ income process would be

$$\begin{aligned}\Delta c_{it} &= \gamma_0(1+\theta\delta)\eta_{it} - \beta\theta\delta r\eta_{it} \\ &+ [1+\gamma_1 r]\,\beta\theta\delta r\eta_{it-1} + [\gamma_1\zeta_0 - \gamma_0](1+\theta\delta)\eta_{it-1} \\ &+ [\gamma_2\zeta_1 - \gamma_1]\,\zeta_0(1+\theta\delta)\eta_{it-2} + [\gamma_2\zeta_1 - \gamma_1]\,\delta\beta\theta r^2\eta_{it-2} \\ &+ [\gamma_3\zeta_2 - \gamma_2]\,\zeta_1\zeta_0(1+\theta\delta)\eta_{it-3} + [\gamma_3\zeta_2 - \gamma_2]\,\delta\beta\theta r^2\eta_{it-3} + \cdots \\ &+ [\gamma_{n-1}\zeta_{n-2} - \gamma_{n-2}]\left[\prod_{0}^{n-3}\zeta_j(1+\theta\delta) + \prod_{1}^{n-3}\zeta_j\delta\beta\theta r^2\right]\eta_{it-n+1} \\ &+ [\zeta_{n-1} - \gamma_{n-1}]\left[\prod_{0}^{n-2}\zeta_j(1+\theta\delta) + \prod_{0}^{n-2}\zeta_j\delta\beta\theta r^2\right]\eta_{it-n}\end{aligned} \tag{4.55}$$

Using the same methods as before we can aggregate the consumption equation to obtain

$$\begin{aligned}
\Delta c_t &= \gamma_0 (1+\theta\delta) \frac{\varepsilon_t}{(1+\theta L)} + \beta\theta\delta r \frac{\varepsilon_t}{(1+\theta L)} \qquad (4.56)\\
&+ [\gamma_1\zeta_0 - \gamma_0] (1+\theta\delta) \frac{\varepsilon_{t-1}}{(1+\theta L)} + [1+\gamma_1 r] \beta\theta\delta r \frac{\varepsilon_{t-1}}{(1+\theta L)} \\
&+ [\gamma_2\zeta_1 - \gamma_1] \zeta_0 (1+\theta\delta) \frac{\varepsilon_{t-2}}{(1+\theta L)} + [\gamma_2\zeta_1 - \gamma_1] \delta\beta\theta r^2 \frac{\varepsilon_{t-2}}{(1+\theta L)} \\
&+ [\gamma_3\zeta_2 - \gamma_2] \zeta_1\zeta_0 (1+\theta\delta) \frac{\varepsilon_{t-3}}{(1+\theta L)} + [\gamma_3\zeta_2 - \gamma_2] \delta\beta\theta r^2 \frac{\varepsilon_{t-3}}{(1+\theta L)} + \cdots \\
&+ \left[\gamma_{n-1}\zeta_{n-2} - \gamma_{n-2}\right] \left[\prod_0^{n-3} \zeta_j (1+\theta\delta) + \prod_1^{n-3} \zeta_j \delta\beta\theta r^2\right] \frac{\varepsilon_{t-n+1}}{(1+\theta L)} \\
&+ \left[\zeta_{n-1} - \gamma_{n-1}\right] \left[\prod_0^{n-2} \zeta_j (1+\theta\delta) + \prod_0^{n-2} \zeta_j \delta\beta\theta r^2\right] \frac{\varepsilon_{t-n}}{(1+\theta L)}
\end{aligned}$$

The change in consumption now follows an $ARMA(1,n)$. The interpretation of this model is somewhat involved; the n lags in this model originate from the speed at which information becomes available to agents each time period - this being up to $n-1$ periods in advance - and how fast agents respond to that new information. The autoregressive term originates because the information that becomes available cannot be broken down into its aggregate and individual components so that agents are subsequently surprised about the information they received the previous time period. This surprise makes agents reassess their beliefs about permanent income at any time period given the information that was already available to them the previous time period and even before that.

4.5.3 IMA(1,2)

Consider the income process

$$\Delta y_t = \varepsilon_t + \phi_1\varepsilon_{t-1} + \phi_2\varepsilon_{t-2}$$

and so the transitory income component is

$$\begin{aligned}
y^T_{t-i} &= -\left(r\delta\phi_1\varepsilon_{t-i} + r\delta^2\phi_2\varepsilon_{t-i} + r\delta\phi_2\varepsilon_{t-i-1}\right) \\
&= -\left[(\psi_1 + \psi_2)\,\varepsilon_{t-i} + \psi_3\varepsilon_{t-i-1}\right] \; \forall i
\end{aligned}$$

We also have the following result

$$
\begin{aligned}
r\sum_{0}^{\infty}\delta^{\tau+1}\Delta E_{t-i}y_{t+\tau-i} &= \left(1+\phi_1\delta+\phi_2\delta^2\right)\varepsilon_{t-i} \quad \forall i \\
&= \psi_0\varepsilon_{t-i}
\end{aligned}
$$

In general we can write the following

$$
\Delta E_{t-j}y_t^p = \prod_{0}^{j-1}\zeta_i\psi_0\varepsilon_{t-j} + \prod_{1}^{j-1}\zeta_i r\beta\left(\psi_1+\psi_2\right)\varepsilon_{t-j} + \prod_{2}^{j-1}\zeta_i r\beta\psi_3\varepsilon_{t-j}
$$

$$
\Delta E_{t-j}y_{t-1}^p = \prod_{0}^{j-2}\zeta_i\psi_0\varepsilon_{t-j} + \prod_{1}^{j-2}\zeta_i r\beta\left(\psi_1+\psi_2\right)\varepsilon_{t-j} + \prod_{2}^{j-2}\zeta_i r\beta\psi_3\varepsilon_{t-j}
$$

Therefore the change in consumption is given by

$$
\begin{aligned}
\Delta c_t = {} & \beta\left[-\left(\psi_1+\psi_2\right)\varepsilon_t-\psi_3\varepsilon_{t-1}\right]-\beta\left[-\left(\psi_1+\psi_2\right)\varepsilon_{t-1}-\psi_3\varepsilon_{t-2}\right] \\
& +\zeta_{n-1}\left[\prod_{0}^{n-2}\zeta_i\psi_0+\prod_{1}^{n-2}\zeta_i r\beta\left(\psi_1+\psi_2\right)+\prod_{2}^{n-2}\zeta_i r\beta\psi_3\right]\varepsilon_{t-n} \\
& +\gamma_0\psi_0\varepsilon_t+\gamma_1\left[\zeta_0\psi_0\varepsilon_{t-1}+r\beta\left(\psi_1+\psi_2\right)\varepsilon_{t-1}\right] \\
& +\gamma_2\left[\zeta_1\left\{\zeta_0\psi_0\varepsilon_{t-2}+r\beta\left(\psi_1+\psi_2\right)\varepsilon_{t-2}\right\}+r\beta\psi_3\varepsilon_{t-2}\right] \\
& +\gamma_3\zeta_2\left\{\zeta_1\left\{\zeta_0\psi_0\varepsilon_{t-3}+r\beta\left(\psi_1+\psi_2\right)\varepsilon_{t-3}\right\}+r\beta\psi_3\varepsilon_{t-3}\right\}+\cdots \\
& +\gamma_{n-2}\left[\prod_{0}^{n-3}\zeta_i\psi_0+\prod_{1}^{n-3}\zeta_i r\beta\left(\psi_1+\psi_2\right)+\prod_{2}^{n-3}\zeta_i r\beta\psi_3\right]\varepsilon_{t-n+2} \\
& +\gamma_{n-1}\left[\prod_{0}^{n-2}\zeta_i\psi_0+\prod_{1}^{n-2}\zeta_i r\beta\left(\psi_1+\psi_2\right)+\prod_{2}^{n-2}\zeta_i r\beta\psi_3\right]\varepsilon_{t-n+1} \\
& -\gamma_0\psi_0\varepsilon_{t-1}-\gamma_1\left[\zeta_0\psi_0\varepsilon_{t-2}+r\beta\left(\psi_1+\psi_2\right)\varepsilon_{t-2}\right] \\
& -\gamma_2\left[\zeta_1\left\{\zeta_0\psi_0\varepsilon_{t-3}+r\beta\left(\psi_1+\psi_2\right)\varepsilon_{t-3}\right\}+r\beta\psi_3\varepsilon_{t-3}\right] \\
& -\gamma_3\zeta_2\left\{\zeta_1\left\{\zeta_0\psi_0\varepsilon_{t-4}+r\beta\left(\psi_1+\psi_2\right)\varepsilon_{t-4}\right\}+r\beta\psi_3\varepsilon_{t-4}\right\}-\cdots \\
& -\gamma_{n-1}\left[\prod_{0}^{n-3}\zeta_i\psi_0+\prod_{1}^{n-3}\zeta_i r\beta\left(\psi_1+\psi_2\right)+\prod_{2}^{n-3}\zeta_i r\beta\psi_3\right]\varepsilon_{t-n}
\end{aligned}
$$

which we can rewrite as

$$
\begin{aligned}
\Delta c_t &= \gamma_0\psi_0\varepsilon_t - \beta(\psi_1+\psi_2)\varepsilon_t \\
&+ [\gamma_1\zeta_0 - \gamma_0]\psi_0\varepsilon_{t-1} + (1+r\gamma_1)r\beta(\psi_1+\psi_2)\varepsilon_{t-1} - \beta\psi_3\varepsilon_{t-1} \\
&+ [\gamma_2\zeta_1 - \gamma_1]\zeta_0\psi_0\varepsilon_{t-2} + [\gamma_2\zeta_1 - \gamma_1]r\beta(\psi_1+\psi_2)\varepsilon_{t-2} + (1+r\gamma_2)\beta\psi_3\varepsilon_{t-2} \\
&+ [\gamma_3\zeta_2 - \gamma_2]\zeta_1\zeta_0\psi_0\varepsilon_{t-3} + [\gamma_3\zeta_2 - \gamma_2]r\beta(\psi_1+\psi_2)\varepsilon_{t-3} + [\gamma_3\zeta_2 - \gamma_2]r\beta\psi_3\varepsilon_{t-3} \\
&+ \cdots + [\gamma_{n-1}\zeta_{n-2} - \gamma_{n-2}]\left[\prod_0^{n-3}\zeta_i\psi_0 + \prod_1^{n-3}\zeta_i r\beta(\psi_1+\psi_2) + \prod_2^{n-3}\zeta_i r\beta\psi_3\right]\varepsilon_{t-n+1} \\
&+ [\gamma_{n-1} - \zeta_{n-1}]\left[\prod_0^{n-2}\zeta_i\psi_0 + \prod_1^{n-2}\zeta_i r\beta(\psi_1+\psi_2) + \prod_2^{n-2}\zeta_i r\beta\psi_3\right]\varepsilon_{t-n}
\end{aligned}
\tag{4.57}
$$

The principal flavour of the partial adjustment model persists; consumption changes follow an $MA(n)$ process. This comes from the theoretical assumption that agents react to the information that becomes available to them up to $n-1$ periods in advance.

Regardless of the type of income process we observe, whilst the theoretical specification for the consumption function does not involve expectational adjustments to the transitory income components - see (4.32), (4.43) - we found them in all the equations above. This is because any response to transitory income does affect permanent income and hence the desired level of consumption. As a final point note that the more terms that we add to the income process, the more effect will the transitory income component have in reinforcing or working in opposite direction to the innovations in permanent income. Any of these $MA(n)$ models can be estimated using time series data.

4.6 Conclusions

The aim of this chapter was to combine a number of models that have successfully explained, from a theoretical and empirical point of view, the failures of the rational expectations permanent income hypothesis. We combined the models of Attfield et al., Flavin and Pischke/Goodfriend to develop consumption specifications that can be as general as possible and which can be estimated using time series data. We do this for the following reasons:

1. To obtain equations with rich dynamics which will enhance our understanding of con-

sumption behaviour and which can also explain the phenomena of excess sensitivity and smoothness.

2. More importantly, from an econometric perspective, to enable us to impose appropriate restrictions on these general specifications to discriminate between the original models in order to understand whether partial adjustment, excess sensitivity or imperfect information can explain the failures of the REPI best, or whether a hybrid combination should be preferred. We undertake these tasks in the next chapter.

4.7 Appendix: An Explanation of Superior Information

Take the first-order VAR representation

$$\begin{bmatrix} \Delta y_t \\ s_t \end{bmatrix} = \begin{bmatrix} a_1 & b_1 \\ c_1 & d_1 \end{bmatrix} \begin{bmatrix} \Delta y_{t-1} \\ s_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{st} \end{bmatrix}$$

where $\varepsilon_{yt} \equiv \Delta y_t - E(\Delta y_t \mid \Omega_{t-1})$, $\varepsilon_{ypt} = \left(\frac{r}{1+r}\right)\sum_{\tau=0}^{\infty}\delta^{\tau}(E_t - E_{t-1})\,y_{t+\tau}$ and where Ω_{t-1} denotes the information set used by the econometrician. Applying the formula for the expectational revision in the present discounted value of future labour income to the VAR system gives

$$\tilde{\varepsilon}_{ypt} = \left(\frac{r}{1+r}\right)\sum_{\tau=0}^{\infty}\delta^{\tau}(E_t - E_{t-1})\,y_{t+\tau} = r\delta\begin{bmatrix} 1 & 0 \end{bmatrix}[I - A(\delta)]^{-1}\begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{yt} - \varepsilon_{ypt} \end{bmatrix}$$

where $\tilde{\varepsilon}_{ypt}$ denotes the econometrician's inference about the innovation in permanent income from the VAR model[16]. If the parameters of the VAR satisfy the REPI restrictions, we have that $r\delta\begin{bmatrix} 1 & 0 \end{bmatrix}[I - A(\delta)]^{-1} = \begin{bmatrix} 1 & -1 \end{bmatrix}$ and

$$\tilde{\varepsilon}_{ypt} = r\delta\begin{bmatrix} 1 & 0 \end{bmatrix}[I - A(\delta)]^{-1}\begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{yt} - \varepsilon_{ypt} \end{bmatrix} = \begin{bmatrix} 1 & -1 \end{bmatrix}\begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{yt} - \varepsilon_{ypt} \end{bmatrix} = \varepsilon_{ypt}$$

Therefore, the expectational revision in permanent income estimated by the econometrician in terms of the VAR is the same as the 'true' expectational revision in permanent income defined relative to the agent's complete information set *if the REPI restrictions hold.* However, some of our models in the text do not satisfy the REPI restrictions[17] and this therefore suggests that the innovation in permanent income cannot be recovered from a bivariate specification for the

[16] In the last expression we have substituted $\varepsilon_{st} = \varepsilon_{yt} - \varepsilon_{ypt}$. To see this consider,
$\varepsilon_{yt} - \varepsilon_{ypt} = \Delta y_t - E(\Delta y_t \mid \Omega_{t-1}) - \left(\frac{r}{1+r}\right)\sum_{\tau=0}^{\infty}\delta^{\tau}(E_t - E_{t-1})\,y_{t+\tau}$
$= (E_t - E_{t-1})\,y_t - \left(\frac{r}{1+r}\right)\sum_{\tau=0}^{\infty}\delta^{\tau}(E_t - E_{t-1})\,y_{t+\tau}$
$= E_t y_t - \left(\frac{r}{1+r}\right)\sum_{\tau=0}^{\infty}\delta^{\tau}E_t y_{t+\tau} - E_{t-1}y_t + \left(\frac{r}{1+r}\right)\sum_{\tau=0}^{\infty}\delta^{\tau}E_{t-1}y_{t+\tau}$
$= s_t - E_{t-1}s_t = s_t - E(s_t \mid \Omega_{t-1}) = \varepsilon_{st}$

[17] See chapter 3 for a description of these restrictions.

change in income and savings. This same argument is valid for the excess sensitivity model too. In that model, *provided the excess sensitivity restrictions*[18] *hold*, the innovation in permanent income is given by

$$\widetilde{\varepsilon}_{ypt} = \begin{bmatrix} 1 & \frac{-1}{1-\beta} \end{bmatrix} \begin{bmatrix} \varepsilon_{yt} \\ (1-\beta)\left(\varepsilon_{yt} - \varepsilon_{ypt}\right) \end{bmatrix} = \varepsilon_{ypt}$$

where we have used an argument equivalent to the REPI case[19] and so in principle it would be possible to recover the innovation in permanent income from the VAR if the excess sensitivity restrictions held.

[18] *See chapter 3 for a description of these restrictions.*

[19] Note that in the excess sensitivity model, savings are defined as

$$s_t = -(1-\beta)\, r \sum_{1}^{\infty} \delta^{\tau} E_t \Delta y_{t+\tau}.$$

Chapter 5

Imperfect Information, Excess Sensitivity and Partial Adjustment: Evidence from US and UK Data

5.1 Introduction

This chapter is split into two sections. In the first section, we estimate the equations that were developed in the previous chapter and we impose certain restrictions to examine which of these equations is able to explain the behaviour of US and UK consumption best. All these equations are estimated in levels and no logarithmic approximations are used. In keeping with previously reported results in the literature, we use the same definitions for consumption that were used by Campbell (1987). In the second part of this chapter we use the same data sets as before but we consider the shorter period that begins in the first quarter of 1973. We re-estimate all the equations from the first part and examine the same tests for smoothness and sensitivity that were reported in our first empirical chapter. We do this because it has been argued that from 1973 onwards, the growth rates in the US and UK have been substantially lower than before.

5.2 Empirical Methods and Results

5.2.1 Five Equations to be Estimated

When we combined the papers of Attfield et al. (1992), Goodfriend (1992), Flavin (1993) and Pischke (1995) we obtained five testable equations. Of these five testable specifications, we see that three equations are (statistically) encompassed within a more general one (not necessarily the most general theoretical specification) leaving another equation to be a competing alternative. When we introduced Pischke's imperfect information into Flavin's model of excess sensitivity we obtained the following specification:

$$A(L)\,\Delta c_t = \beta A(L)\,\Delta y_t + (1-\beta)\,A(\delta)\,\phi(L)\,\varepsilon_t \tag{5.1}$$

where the number of lags in income and consumption *must be the same and not necessary equal to the number of lags in the error term.* This specification encompasses the next three.

The next testable equation did result from the assumption that agents use lagged information to correct for optimization errors (i.e. Goodfriend's model) in the excess sensitivity model,

$$\begin{aligned}\Delta c_t &= \beta\Delta y_t + (1-\beta)\left[\phi(\delta)\,\omega + \theta(\delta)\,(1-\omega)\right]\varepsilon_t \\ &\quad + (1+r)\,(1-\beta)\left[\phi(\delta) - \theta(\delta)\right](1-\omega)\,\varepsilon_{t-1}\end{aligned} \tag{5.2}$$

where we have *no lags in consumption and income, and only one lag in the error process.* This equation is a special case of (5.1) with $A(L) = 1$ and $A(\delta)\,\phi(L) = \mu_0 + \mu_1 L$, where $\mu_0 = [\phi(\delta)\,\omega + \theta(\delta)\,(1-\omega)]$ and $\mu_1 = (1+r)\,[\phi(\delta) - \theta(\delta)]\,(1-\omega)$.

When we combined the excess sensitivity hypothesis and the model of partial adjustment model developed by Attfield et al. we obtained an $MA(n)$ process,

$$
\begin{aligned}
\Delta c_t &= \gamma_0\psi_0\varepsilon_t - \beta(\psi_1+\psi_2)\varepsilon_t \qquad (5.3)\\
&+[\gamma_1\zeta_0-\gamma_0]\psi_0\varepsilon_{t-1} + (1+r\gamma_1)r\beta(\psi_1+\psi_2)\varepsilon_{t-1} - \beta\psi_3\varepsilon_{t-1}\\
&+[\gamma_2\zeta_1-\gamma_1]\zeta_0\psi_0\varepsilon_{t-2} + [\gamma_2\zeta_1-\gamma_1]r\beta(\psi_1+\psi_2)\varepsilon_{t-2} + (1+r\gamma_2)\beta\psi_3\varepsilon_{t-2}\\
&+[\gamma_3\zeta_2-\gamma_2]\zeta_1\zeta_0\psi_0\varepsilon_{t-3} + [\gamma_3\zeta_2-\gamma_2]r\beta(\psi_1+\psi_2)\varepsilon_{t-3} + [\gamma_3\zeta_2-\gamma_2]r\beta\psi_3\varepsilon_{t-3}\\
&+\cdots+[\gamma_{n-1}\zeta_{n-2}-\gamma_{n-2}]\left[\prod_0^{n-3}\zeta_i\psi_0 + \prod_1^{n-3}\zeta_i r\beta(\psi_1+\psi_2) + \prod_2^{n-3}\zeta_i r\beta\psi_3\right]\varepsilon_{t-n+1}\\
&+[\gamma_{n-1}-\zeta_{n-1}]\left[\prod_0^{n-2}\zeta_i\psi_0 + \prod_1^{n-2}\zeta_i r\beta(\psi_1+\psi_2) + \prod_2^{n-2}\zeta_i r\beta\psi_3\right]\varepsilon_{t-n}
\end{aligned}
$$

which we view as a special case of (5.1) with subsequent restrictions $A(\delta)\phi(L) = \mu_0 + \mu_1 L + \cdots + \mu_n L^n$ where $\mu_0 \neq 1$ and every term in $A(L)$ is equal to zero.

Assuming that agents use lagged information to correct for optimization errors but are quite sluggish to adjust to those shocks resulted in the following,

$$
\begin{aligned}
\Delta c_t &= \gamma_0[\phi(\delta)\omega + (1-\omega)\theta(\delta)]\varepsilon_t + \rho[\phi(\delta)-\theta(\delta)](1-\omega)\varepsilon_{t-1} \qquad (5.4)\\
&+[\zeta_0\gamma_1-\gamma_0][\phi(\delta)\omega+(1-\omega)\theta(\delta)]\varepsilon_{t-1} + \rho[\phi(\delta)-\theta(\delta)](1-\omega)\varepsilon_{t-2}\\
&+\zeta_0[\zeta_1\gamma_2-\gamma_1][\phi(\delta)\omega+(1-\omega)\theta(\delta)]\varepsilon_{t-2} + \rho[\phi(\delta)-\theta(\delta)](1-w)\varepsilon_{t-3}+\cdots\\
&+[\zeta_{n-2}\gamma_{n-1}-\gamma_{n-2}]\prod_0^{n-3}\zeta_j[\phi(\delta)\omega+(1-\omega)\theta(\delta)]\varepsilon_{t-n+1} + \rho[\phi(\delta)-\theta(\delta)](1-\omega)\varepsilon_{t-n}\\
&+[\zeta_{n-1}-\gamma_{n-1}]\prod_0^{n-2}\zeta_j[\phi(\delta)\omega+(1-\omega)\theta(\delta)]\varepsilon_{t-n} + \rho[\phi(\delta)-\theta(\delta)](1-\omega)\varepsilon_{t-n-1}
\end{aligned}
$$

We have an $MA(n+1)$ model for the change in consumption. This equation is a special case of (5.1) with restrictions $A(z) = 1$ and $A(\delta)\phi(L) = \mu_0 + \mu_1 L + \cdots + \mu_{n+1}L^{n+1}$ where $\mu_0 \neq 1$ and $\beta = 0$. Furthermore, we can see that (5.3) is a special case of equation (5.4), but because n is not know *a priori*, it is not possible to statistically distinguish between them. Thus, (5.3) and (5.4) are for empirical purposes observationally equivalent.

Finally, the competing alternative came as the result of introducing Pischke's imperfect information into Attfield et al.:

$$A(L)\Delta c_t = \gamma_0 \Delta y_t + [\gamma_1 \zeta_0 - \gamma_0] \Delta y_{t-1} + [\gamma_2 \zeta_1 - \gamma_1] \zeta_0 \Delta y_{t-2} + \cdots \\ + [\gamma_{n-1}\zeta_{n-2} - \gamma_{n-2}] \prod_{j=0}^{n-3} \zeta_j \Delta y_{t-n+1} + [\zeta_{n-1} - \gamma_{n-1}] \prod_{j=0}^{n-2} \zeta_j \Delta y_{t-n} \tag{5.5}$$

In this equation, *the number of lags of income and consumption need not be the same.* In all previous specifications, we found that the number of lags for the change in labour income and consumption were the same, hence (5.5) may not be necessarily embed within equation (5.1). We may, therefore, consider this equation to be the competing alternative to (5.1). Note that compared to the four other equations above, (5.5) does not have (stochastic) error terms. There are two reasons why we do not substitute the specification for the aggregate income process into this equation: first, we do not want to 'contaminate' the equation with the introduction of an incorrect aggregate labour income process and second, we do want to lose the information inherent in the model of partial adjustment (i.e. (5.5) can tell us the value of n, the time span where adjustment to 'news' is than full). If we were to substitute the labour income process into (5.5) we would end up with an $ARMA(p,q)$ where q is the sum of the n adjustment terms (the labour income terms) and the d error terms which constitute the aggregate labour income term $\Delta y_t = \phi(L)\varepsilon_t$ (see equation (4.36) in chapter 4). The way in which we introduce stochastic terms to (5.5) is to assume measurement errors in one or all of the variables involved.

5.2.2 Econometric Methods

It is apparent that, statistically speaking, there exists two competing data generating processes, one of which is a general specification of three other equations which can be tested in a general to specific manner. Our aim is to choose the best fitting equation from the five presented earlier. We therefore begin with the estimation of the two most general and competing alternatives - equations (5.1) and (5.5) - and then test down until a more parsimonious specification results. We have to be aware of an econometric issue that must be dealt with before estimation. In aggregate data, it is likely that labour income and consumption are jointly endogenous. To overcome this problem and obtain consistent estimates, instrumental variables are used for Δy_t. The instruments used for the change in labour income are the same as those that Demery and Duck (1999, 2000) used for the US and UK. For the US, Demery and Duck used 2 to 6 lags for the change in labour income, 2 to 6 lags for the change in exports, government consumption, defence expenditure, government wages and salaries and net factor income from abroad. For the UK, the instruments were 2 to 6 lags for the change of labour income, lags 2 to 6 for the change in government consumption, public corporation investment, general government investment, central government consumption and the exports of goods and services. Note that all instruments start at $t-2$ and not at $t-1$. Demery and Duck claim, following an argument by Christiano et al. (1991) that consistent estimates can only be obtained using instrumental variables dated at $t-2$ and earlier because 'if individual planning takes place continuously the time-averaged quarterly first-difference in consumption will be correlated with changes in labour income lagged one period'.[1999, pp. 379].

For both countries the data we use are quarterly, seasonally-adjusted observations on labour income, total consumption expenditure and consumption expenditure on non-durables. For the US the series run from 1959:4 to 1996:1 whereas for the UK the series run from 1963:1 to 1996:2[1]. The definitions of consumption and labour income follow those that were adopted by Blinder and Deaton (1985) for the US and Attfield et al. (1990) for the UK. The unit root tests from chapter 3 apply and so we know that all the series involved are $I(1)$.

5.2.3 Tests on the Five Equations

In the first stage we seek to obtain the two best fitting competing alternatives[2]. None of the specifications were restricted in any particular way and so equation (5.1) was estimated without requiring the coefficients of the lags in consumption and income to have the common component $A(L)$ although only equations where the number of lags in income were the same as the number of lags in consumption were estimated. We began with the estimation of the most general specifications which included, to start with, ten lags in income and ten lags in consumption for equation (5.5) and ten lags in income, ten lags in consumption and ten lags in the error term for equation (5.1). The Pagan (1974) estimation method was used to obtain the

[1] Some of the instrumental variables are not available prior to 1963:1.

[2] All the errors were normalised before estimating all the equations. For instance, equation (5.1) was estimated as $A(L)\Delta c_t = \beta A(L)\Delta y_t + \lambda(L)v_t$ where $v_t = (1-\beta)A(\delta)\varepsilon_t$ and $\lambda(z) = 1 + \lambda_1 L + \lambda_2 L^2 + \cdots$.

estimates for all the equations[3].

In the second stage all the equations were examined for serial correlation and those specifications that did not exhibit any serial correlation were then chosen as possible explanations for the change in consumption. To obtain the specific order for all the lag terms amongst those remaining equations we performed Likelihood Ratio tests for the joint significance of the estimated coefficients and we also looked at the Akaike Information Criteria (AIC)[4]. We also look at a Bartlett adjustment[5] to the likelihood ratio test. The adjusted likelihood ratio (ALR) is calculated from $ALR = LR/b$, where LR denotes the likelihood ratio test, $b = \frac{[n+\frac{1}{2}(p+p^{o})+1]}{n}$ and n is the number of (usable) observations, p is the number of estimated coefficients in the unrestricted equation and p^0 is the number of estimated coefficients in the restricted equation[6]. Table 1 shows the preferred consumption specifications for the US and the UK using this criteria.

L_i denotes the maximum likelihood estimate of the equation and aic the Akaike Information Criteria. Higher order lags were also examined, but we found that the specifications reported in table 1 fitted the data best[7]. From the table we see, for example, that US total consumption appears to be defined best as having one lag in income and consumption and two lags in the error term for equation (5.1). For equation (5.5), US total consumption appears to be defined by having no lags in the change in income and three lags in the change in consumption. From this table we also observe that for all the consumption measures in the US and the UK, Flavin's incomplete information equation did not embed the equation which introduced imperfect information to partial adjustment and so there is no (direct) way of discriminating between the two equations. Note further that for all specifications of consumption for both the US and the UK,

[3]The Pagan estimation method is the default method in Shazam when the errors are AR(2), MA(2) or have higher order lags. Other methods did not change the nature of the results when we estimated lower order lags as the same coefficients were obtained. For a brief explanation of the Pagan method, see the Shazam manual (v.8.0) page 146.

[4]This was defined as

$$AIC = T \ln(\text{residual sum of squares}) + 2n$$

where n = number of parameters estimated; T = number of usable observations.

[5]The Barlett adjustment is undertaken on the LR test because we have a finite sample and the LR test is an asymptotic test.

[6](See C.L.F. Attfield (1995) Journal of Econometrics for more details).

[7]The UK specifications for total consumption in table 5.1 do not pass the Jarque-Bera test for normality. Alternative specifications were not able to overcome this problem.

Table 5.1: Preferred Specifications for Equations (5.1) and (5.5)

US	Total	Nondurable OLS
eqn (5.1)	1 lag in Δy_t, 1 lag in Δc_t, 2 lags in v_t, $L_1 = -632.596$ $aic = 1582.192$ Ser Corr $\chi^2(15) = 5.48$ Normality $\chi^2(2) = 5.66$	1 lag in Δy_t, 1 lag in Δc_{t-i}, 5 lags in v_t, $L_1 = -584.130$ $aic = 1491.259$ Ser Corr $\chi^2(15) = 5.91$ Normality $\chi^2(2) = 1.83$
eqn (5.5)	0 lags in Δy_t, 3 lags in Δc_t, $L_5 = -633.195$ $aic = 1581.389$ Ser Corr $\chi^2(15) = 4.69$ Normality $\chi^2(2) = 4.72$	1 lag in Δy_t, 6 lags in Δc_t $L_5 = -584.061$ $aic = 1491.121$ Ser Corr $\chi^2(15) = 8.72$ Normality $\chi^2(2) = 1.65$
UK		
eqn (5.1)	1 lag in Δy_t, 1 lag in Δc_t, 3 lags in v_t, $L_1 = -491.139$ $aic = 1240.098$ Ser Corr $\chi^2(15) = 8.21$ Normality $\chi^2(2) = 16.01$	1 lag in Δy_t, 1 lag in Δc_t, 1 lag in v_t, $L_1 = -483.189$ $aic = 1220.197$ Ser Corr $\chi^2(15) = 10.44$ Normality $\chi^2(2) = 2.02$
eqn (5.5)	1 lag in Δy_t, 3 lags in Δc_t, $L_5 = -492.159$ $aic = 1240.135$ Ser Corr $\chi^2(15) = 12.23$ Normality $\chi^2(2) = 16.58$	1 lag in Δy_t, 3 lags in Δc_t, $L_5 = -483.039$ $aic = 1219.896$ Ser Corr $\chi^2(15) = 10.58$ Normality $\chi^2(2) = 2.25$

equations (5.1) and (5.5) never have the same number of lags in consumption. This means that equations (5.1) and (5.5) never have the same common component $A(L)$ which we would have expected if individuals faced labour income processes where they could not distinguish between individual and aggregate information. Imposing the appropriate restrictions using a Likelihood Ratio test enables us to discriminate between the more general and unrestricted version of equation (5.1) and equations (5.2), (5.3) and (5.4) to examine which specification would fit the data better[8]. Table 2 reports the results of introducing these restrictions to equation (5.1).

From Table 2 we see that all the restrictions are rejected at the 5% significance level even after performing the Barlett adjustment to the likelihood ratio test. Thus the *unrestricted* version of equation (5.1) can explain the behaviour of US and UK consumers better than equations (5.2), (5.3) and (5.4) can. Taking the view that all of these equations are more general specifications than Hall's random walk equation for consumption, we can conclude that the unrestrictive version of equation (5.1) is in itself a better explanation of consumption changes than Hall's specification.

From table 2 we see that it is possible to reject the appropriate restrictions imposed on equation (5.1) at the 5% significance level and these restrictions involve tests that examine the statistical significance of the autoregressive components of consumption and the lags in income. Since both terms are statistically different from zero, our data suggests that imperfect information appears to be a better theoretical explanation for consumption behaviour than lagged information (i.e. that Pischke's model is superior to Goodfriend's.) This is not, however, evidence that equation (5.1) can explain the behaviour of consumption accurately in the UK and the US because we have only estimated an unrestricted version of such equation and we have not tested against other (and more general) specifications. However, (5.1) implies a set of restrictions that can help determine the significance of the equation and which we now consider and report in Table 3. First, to test whether the excess sensitivity-incomplete information hypothesis is capable of explaining consumer behaviour in the US and the UK, note that the restriction that the income coefficients are equal to the consumption coefficients but scaled by the marginal propensity to consume out of transitory income β can be imposed. This test is

[8]Note that equations (5.3) and (5.4) are observationally equivalent since the order of the MA process is not known *a priory*. Hence the absence of (5.3) in Table 2.

Table 5.2: Tests of (Unrestricted) (5.1) against (5.2) and (5.4)

US	Total	Nondurable OLS
eqn (5.1)	1 lag in Δy_t, 1 lag in Δc_t, 2 lags in v_t, $L_1 = -632.596$	1 lag in Δy_t, 1 lag in Δc_{t-i}, 5 lags in v_t, $L_1 = -584.130$
eqn (5.2)	0 lags in Δy_t, 1 lag in v_t $L_2 = -649.676$ Ser Corr $\chi^2(15) = 24.43$ Normality $\chi^2(2) = 1.69$ Restriction: LR Test $\chi^2(3) = 34.16$	0 lags in Δy_t, 1 lag in v_t $L_2 = -612.946$ Ser Corr $\chi^2(15) = 63.51$ Normality $\chi^2(2) = 1.03$ Restriction: LR Test $\chi^2(6) = 57.632$
eqn (5.4)	2 lags in v_t $L_4 = -663.081$ Ser Corr $\chi^2(15) = 43.1$ Normality $\chi^2(2) = 13.96$ Restriction: LR Test $\chi^2(2) = 60.97$	5 lags in v_t $L_4 = -604.099$ Ser Corr $\chi^2(15) = 50.28$ Normality $\chi^2(2) = 13.54$ Restriction: LR Test $\chi^2(2) = 39.938$
UK		
eqn (5.1)	1 lag in Δy_t, 1 lag in Δc_t, 3 lags in v_t, $L_1 = -491.139$	1 lag in Δy_t, 1 lag in Δc_t, 1 lag in v_t, $L_1 = -483.189$
eqn (5.2)	0 lags in Δy_t, 1 lag in v_t $L_2 = -504.289$ Ser Corr $\chi^2(15) = 23.23$ Normality $\chi^2(2) = 12.40$ Restriction: LR Test $\chi^2(4) = 26.3$	0 lags in Δy_t, 1 lag in v_t $L_2 = -495.230$ Ser Corr $\chi^2(15) = 24.23$ Normality $\chi^2(2) = 0.82$ Restriction: LR Test $\chi^2(2) = 24.082$
eqn (5.4)	3 lags in v_t $L_4 = -495.528$ Ser Corr $\chi^2(15) = 17.65$ Normality $\chi^2(2) = 24.69$ Restriction: LR Test $\chi^2(2) = 8.778$	1 lag in v_t $L_4 = -505.545$ Ser Corr $\chi^2(15) = 28.92$ Normality $\chi^2(2) = 2.13$ Restriction: LR Test $\chi^2(2) = 44.712$

a 'weak test' for testing the validity of our theory since an even stronger test for (5.1) would require that restrictions on the error term be imposed at the same time as the restrictions of the weak-test. These further restrictions involve a test on the coefficients of the error terms; these ought to be a function of $A(\delta)$ which should in turn be equal to the coefficients on lagged consumption and (lagged) labour income (i.e. $A(L)$ and $\beta A(L)$ respectively). This however requires *a) knowledge* of the ϕ_i coefficients which can be obtained by estimating the correct form for the aggregate labour income process $\Delta y_t = \phi(L)\varepsilon_t$ and *b)* the further *assumption of a constant rate of interest.* The major problem with this stronger type of test is that the labour income process has to be modelled correctly and imposing a certain functional form may affect the nature of the results[9].

Table 3 reports the results of imposing the weak restriction with the likelihood ratio test, and also the size of the coefficient β which is obtained from the estimation of the restricted model[10].

All of the appropriate restrictions are rejected at the 5% significance level for all consumption specifications in both countries. Moreover, note that whilst the excess sensitivity coefficient is insignificant for the US it is significant for the UK although the values of β differ in the single equation specification from those calculated from the *VARs*. If we consider (5.1) to be a more general specification than Pischke's imperfect information hypothesis, then the results encountered in Table 3 are consistent with Demery and Duck's rejections of the Pischke model

[9]Demery and Duck (1999, 2000) perform a similar type of test to examine the time series implications of the Pischke (1995) model. They estimate the following logarithmic approximation of the Pischke model,

$$\frac{\Delta c_t}{y_{t-1}} \approx \lambda + \kappa\alpha(L)\gamma(L)\Delta\log y_t + \sum_{j=1}^{q_{max}} \theta_j \frac{\Delta c_{t-j}}{y_{t-j-1}} + \theta(L)(1+\pi L)\tilde{\omega}_t$$

where $\kappa = \frac{r(1+\mu)}{r-\mu}\left(\frac{\theta(\delta)}{\alpha(\delta)\gamma(\delta)}\right)$. The restrictions that Demery and Duck impose are:

1) There ought to be q_{max} coefficients on the lagged dependent variables and $q_{max}+1$ moving average error coefficients.

2) Provided one can obtain accurate estimates of the aggregate labour income process, $\alpha(L)(\Delta\log y_t - \mu) = \beta(L)\varepsilon_t$, one could identify the parameters of (the individual component) $\gamma(L)$ from the present and lagged coefficients of labour income. This imposes a set of restrictions on the consumption equation.

3) Given an estimated value for the intercept κ, that μ is known and that a value for r can be imposed a priori, one can test whether this estimated value of κ is explained by the estimated coefficients in $\theta(L)$ - the lagged consumption coefficients -, $\alpha(L)$ and $\gamma(L)$ -the income coefficients - from the consumption equation.

[10]The significance of the excess sensitivity parameter determines whether the hybrid model is better than the Pischke model. Since we have found that the lagged consumption terms are significant, we can conclude that the hybrid model is superior to the model of excess sensitivity since the model of excess sensitivity does not have any autoregressive components in consumption changes.

Table 5.3: Test of Weak Restrictions Implied Equation (5.1)

US	Total Consumption	NonDurable OLS
LR:	$\chi^2(1)$ =22.036	$\chi^2(1)$ =9.561
ALR:	$\chi^2(1)$ =21.225	$\chi^2(1)$ =8.713
β(s.e.)	0.129 (0.117)	0.041 (0.0952)
UK		
LR:	$\chi^2(1)$ =6.997	$\chi^2(1)$ =9.062
ALR:	$\chi^2(1)$ =6.649	$\chi^2(1)$ =8.813
β(s.e.)	0.518 (0.173)	0.517 (0.181)

Table 5.4: Estimation Results for Competing Alternative (5.5)

US	Total Consumption	Nondurable OLS
γ_0(s.e.)	0.348(0.129)	0.222(0.098)
γ_1(s.e.)		0.438(0.1297)
a_1(s.e.)	0.163(0.081)	0.381(0.0847)
a_2(s.e.)	0.155(0.083)	-0.0314(0.0924)
a_3(s.e.)	0.319(0.0807)	0.288(0.0907)
a_4(s.e.)		-0.162(0.0879)
a_5(s.e.)		-0.072(0.0895)
a_6(s.e.)		0.216(0.0808)
$\sum a_i$	0.637	0.633
UK		
γ_0(s.e.)	0.376(0.1528)	0.444(0.142)
γ_1(s.e.)	0.747(0.1938)	0.789(0.187)
a_1(s.e.)	-0.126(0.0886)	0.019(0.0880)
a_2(s.e.)	0.173(0.0848)	0.268(0.0855)
a_3(s.e.)	0.263(0.0841)	
$\sum a_i$	0.310	0.287

at the 5% significance level for both the US and the UK. Furthermore, these results are also in line with the findings which we encountered in chapter 3.

In Table 4 we turn to equation (5.5) as we present the parameter estimates of this equation. The estimated values of the $\gamma's$ are all significantly less than one and they show the pattern predicted by the theory of partial adjustment[11]. Table 4 also reports the coefficients of the lagged consumption terms which are associated with the coefficients of the individual income process when agents cannot distinguish between aggregate and individual information[12]. Unlike Attfield et al., we cannot use the same orthogonality test they performed to check the validity

[11]We obtain the standard errors for each of the partial adjustment coefficients by following the methods suggested by Goldberger (1964) pp. 122-5.

[12]The a_i's in table 5.4 correspond to the individual elements of $A(L)$.

of this model, because no other income and consumption terms were found to be significant in our estimation stage. Moreover, given that all the estimated equations had no serial correlation, introducing lagged error terms would prove to be inconsequential.

Apart from total consumption in the US (where $\gamma_1 = 0$) adjustment appears to be complete within two periods (i.e. $\gamma_i = 0$, $i \geq 2$). This suggests that adjustment is slightly faster in our framework than the yearly adjustment Attfield et al. found to take place in their quarterly data for the UK and the US[13]. Furthermore, we find all the adjustment coefficients to be smaller than those reported by Attfield et al. and also for the coefficients to be larger in the UK than the US, thus implying consumers in the UK do adjust more fully to innovations in labour income compared to their US counterparts. An explanation for all of these results may be found within the confines of our equation and specifically in the number of lags in consumption[14]. Lagged consumption terms represent surprise innovations agents encounter if they are not capable of distinguishing between their individual and aggregate labour income components and thus represent sluggish behaviour in the partial adjustment framework: responses to an innovation in labour income do not take place within two periods as the partial adjustment framework would have us believe but they are an ongoing process which could take up to an infinite number of periods[15]. Closer examination of the lagged non-durable consumption terms indicates that the *individual* (imperfect information) income processes in the US consumption have more lags (and appear to have an overall greater effect on consumption) than those in the UK thus potentially introducing more complex adjustment dynamics and perhaps explaining why the partial adjustment coefficients have a smaller size in the US than in the UK; consumers in the US continue to be surprised up to four times more than UK consumers thereby deciding to adjust less to each innovation.

[13]Attfield et al. found that four adjustment coefficients were statistically significant for the US and five for the UK.

[14]Again, since the lagged terms in consumption are significant, this implies that the model of imperfect information is better at explaining the behaviour of US and UK consumers than the model of lagged information.

[15]We can see this if we invert the lag operator in front of the change in consumption which will give us an infinite number of income terms.

5.2.4 Conclusions

In this section we used aggregate time series data for the UK and the US to examine five models that resulted from marrying imperfect information, partial adjustment and excess sensitivity to one another. Of those five models, we consider two of those to be general competing alternatives and the other three to be embed within one of the other two models. Our results suggest that the standard martingale specification for consumption can be rejected in favour of one of the two general models although we are not able to determine which of the two models is superior. We also find that the data does tend to be kinder to the model of imperfect information above the model of lagged information.

The Davidson and McKinnon tests cannot be easily applied to our case since our residuals are autocorrelated of an order higher than one and no critical values have been tabulated for this - see Bernanke, Bohn and Reiss (1988) for an attempt to look at models with serial correlation. Bernanke et al., give at least three reasons why 'these non-nested testing principles have not been more widely applied. First, non-nested tests have not been developed for time-series models that possess general mixtures of serial correlation, lagged dependent variables, and endogenous variables. Second, there is some evidence that the asymptotic critical values of existing non-nested test procedures reject the null hypothesis too often in finite samples. Third, in more complicated practical applications, especially when one wishes to implement non-nested tests based upon maximum-likelihood techniques, these tests can be quite burdensome computationally'[pp. 294]. Bernanke et al.'s paper explores 'the usefulness of non-nested testing procedures for linear regression models with first-order serially correlated errors' [pp. 294]. Comparing time-series investment models on quarterly US business investment data, Bernanke et al. find that when using 'non-nested tests that take into account serial correlation in the residuals all the models are rejected by at least one of the other models. However, in a Monte Carlo study [...] we find a significant bias in the distribution of these tests toward rejection of the true model'. [pp. 320]. Bernanke et al. recommend the need to exercise caution in rejecting non-nested models with highly serially correlated errors [pp. 320]. They also point out that there is a great computational burden associated with computing these GLS non-nested tests and that 'investigators may wish to evaluate empirically the adequacy of the asymptotic distribution through Monte Carlo or Bootstrap experiments' [pp. 320] . Thus to be able to compare (5.1) and (5.5) using non-nested procedures we would have to check the asymptotic distribution through Monte Carlo or Bootstrap experiments which is beyond the scope of this thesis.

5.3 An Examination of all our Results from 1973 onwards

5.3.1 Excess Smoothness and Sensitivity: Single Equation Tests

To compare the implications of the post 1973 data sets with our previous results, we begin with the estimation of an $ARIMA(1,1,0)$ for labour income. For the US we obtained the following results

$$\Delta y_t = \underset{(0.001)}{0.048} - \underset{(0.105)}{0.097}\Delta y_{t-1} \qquad \begin{array}{l} \sigma_\varepsilon = 0.121 \\ R^2 = 0.0094 \end{array}$$

Assuming a rate of interest of 10%, the predicted standard error in the innovation in consumption ought to be $\hat{\sigma}_{\Delta c} = 0.919 \times 0.121 = 0.111$. The standard deviation of total consumption and consumption excluding durables obtained from the data for the post 1973 period, are $\sigma_{\Delta c} = 0.092$ and $\sigma_{\Delta cn} = 0.052$ respectively. Thus, this simple framework states that excess smoothness appears to exist in the innovation of consumption. However, closer inspection reveals a poor fit for the labour income equation and the Chow and Goldfeld-Quandt tests show that there is some evidence of a structural break in this equation.

For the UK, the estimated $ARIMA(1,1,0)$ yielded the following results,

$$\Delta y_t = \underset{(1.681)}{4.926} - \underset{(0.106)}{0.061}\Delta y_{t-1} \qquad \begin{array}{l} \sigma_\varepsilon = 15.449 \\ R^2 = 0.0037 \end{array}$$

The predicted standard deviation of the change in consumption for a rate of interest of 10% is according to REPI $\hat{\sigma}_{\Delta c} = 0.948 \times 15.449 = 14.640$. The standard deviation for the change in total consumption is $\sigma_{\Delta c} = 14.483$ and for non-durable consumption is $\sigma_{\Delta cn} = 9.714$. Thus, there seems to be evidence of excess smoothness in the UK. Once more, an examination of the equation for the change in labour income reveals not only a poor fit, but the Goldfeld-Quandt test continues to reject this specification (the results for the Chow test are more favourable,

Table 5.5: Means, Actual and Predicted Standard Deviations for the US

US	Mean	s.d.	Scaling Factor
$\Delta \log c_t$	1.735	2.629	
$\frac{\Delta c_t}{y_{t-1}}$	1.994	2.979	
$\Delta \eta_t$	-1.93	3.525	
$\Delta \log c_{nt}$	1.538	1.762	
$\lambda \frac{\Delta c_{nt}}{y_{t-1}}$	1.764	2.013	1.192 (Mean)
$\lambda \Delta \eta_t$	-1.701	2.707	1.192
$\lambda \frac{\Delta c_{nt}}{y_{t-1}}$	1.902	2.170	1.285 (OLS)
$\lambda \Delta \eta_t$	-1.316	2.735	1.285

there is some evidence of structural breaks for only a couple of years in the late 1970s).

We run the equivalent labour income equations in logarithmic terms to compare the equations with our previous results and the results of Campbell and Deaton in the belief that these may capture the properties of the data better than the data in levels. For the US, the equivalent results are,

$$\Delta \log y_t = \underset{(0.420)}{1.411} - \underset{(0.106)}{0.0492} \Delta \log y_{t-1} \qquad \begin{array}{l} \sigma_\varepsilon = 3.764 \\ R^2 = 0.0024 \end{array}$$

Given that the sample average quantity rate of growth of labour income in the US for the post 1973 period is 1.364% per annum, and with a rate of interest of 10%, then the predicted standard deviation for the innovation in consumption is $\dot{\sigma}_{\Delta c} = 0.989 \times 3.764 = 3.722$. The standard deviations for total consumption and non-durable consumption are reported in Table 5 for the US. For the UK, the equivalent labour income equation in logarithmic form is

$$\Delta \log y_t = \underset{(0.673)}{1.810} - \underset{(0.153)}{0.0731} \Delta \log y_{t-1} \qquad \begin{array}{l} \sigma_\varepsilon = 6.233 \\ R^2 = 0.0053 \end{array}$$

Given a sample average quantity rate of growth of labour income in the UK for the post 1973 period of 1.721% per annum, and with a rate of interest of 10%, the predicted standard deviation for the innovation in consumption is $\dot{\sigma}_{\Delta c} = 0.981 \times 6.233 = 6.116$. The standard deviations for total consumption and non-durable consumption are reported in Table 6 for the UK.

For the US, we note that all the standard deviations are less than the predicted 3.7, thus

Table 5.6: Means, Actual and Predicted Standard Deviations for the UK

UK	Mean	s.d.	Scaling Factor
$\Delta \log c_t$	1.937	4.953	
$\frac{\Delta c_t}{y_{t-1}}$	2.317	5.763	
$\Delta \eta_t$	-2.145	7.708	
$\Delta \log c_{nt}$	1.782	3.574	
$\lambda \frac{\Delta c_{nt}}{y_{t-1}}$	2.125	4.194	1.099 (Mean)
$\lambda \Delta \eta_t$	-1.957	6.553	1.099
$\lambda \frac{\Delta c_{nt}}{y_{t-1}}$	2.429	4.793	1.256 (OLS)
$\lambda \Delta \eta_t$	-1.283	6.691	1.256

suggesting that there is evidence of excess smoothness. For the UK, the results continue to be mixed in as far as the predicted standard deviation of consumption changes continues to be in between the consumption ratio and the equivalent term $\Delta \eta_t$. The post 1973 sample does not produce results that are different from before. We now turn to the (superior) VAR analysis.

5.3.2 Excess Smoothness and Sensitivity Tests: Campbell and Deaton (1989) and Flavin's (1993)Tests

Logarithmic Data

To determine the lag length of the VAR specification we examine the likelihood ratio test, the Akaike Information Criteria (AIC) and the Schwartz Bayesian Criteria (SBC). For the US, the AIC and SBC favoured a $VAR(1)$ system for total consumption although the $VAR(2)$ measure was close. For both non-durable measures a $VAR(2)$ system was preferred by the AIC whilst the SBC suggested a $VAR(1)$. For the US, both systems for all consumption measures are reported in the tests. For the UK, the AIC suggested a $VAR(2)$ for all measures of consumption whilst the SBC suggested a $VAR(1)$. Again we report both measures. The Campbell and Deaton tests for excess smoothness are reported in Tables 7 (for the US) and 8 (for the UK). Tables 9 (US) and 10 (UK) report the LR tests undertaken by Flavin[16].

The results of the Wald tests for the US are for the most part consistent with our previous results. Although the $VAR(1)$ systems do accept the restrictions imposed by REPI, the ratio

[16]The weak implication of the REPI, that of Granger causality is satisfied for all consumption measures in both countries.

Table 5.7: Tests for Excess Smoothness in the US

	WaldTest *(p − Value)*	*Predicted Innovation*	*Actual Innovation*	*Ratio (s.e.)*
Total Consumption				
VAR(1)	1.699 (0.429)	3.735	3.477	0.931 (0.247)
VAR(2)	7.246 (0.123)	4.321	3.388	0.784 (0.216)
Non-Durable $\lambda = 1.19$				
VAR(1)	2.249 (0.325)	2.682	2.659	0.991 (0.106)
VAR(2)	12.395 (0.015)	3.567	2.531	0.709 (0.242)
Non-Durable $\lambda = 1.28$				
VAR(1)	3.282 (0.194)	2.636	2.671	1.013 (0.108)
VAR(2)	12.254 (0.016)	3.510	2.559	0.729 (0.3059)

Table 5.8: Tests for Excess Smoothness in the UK

	WaldTest *(p − Value)*	*Predicted Innovation*	*Actual Innovation*	*Ratio (s.e.)*
Total Consumption				
VAR(1)	5.711 (0.058)	7.047	7.435	1.055 (0.237)
VAR(2)	15.054 (0.004)	7.842	7.083	0.903 (0.304)
Non-Durable $\lambda = 1.099$				
VAR(1)	5.076 (0.079)	6.278	6.348	1.011 (0.199)
VAR(2)	15.919 (0.003)	6.967	6.014	0.863 (0.287)
Non-Durable $\lambda = 1.28$				
VAR(1)	6.559 (0.038)	6.377	6.431	1.085 (0.214)
VAR(2)	17.950 (0.001)	6.970	6.085	0.873 (0.334)

Table 5.9: LR Tests for Orthogonality and Sensitivity for the US

	Orthogonality	Sensitivity	*value of* β (*s.e.*)
Total Consumption			
VAR(1)	$\chi^2(2)$ =1.774	$\chi^2(1)$ =1.153	−0.318 (0.470)
VAR(2)	$\chi^2(4)$ =7.289	$\chi^2(3)$ =7.057	0.145 (0.293)
Non-Durable $\lambda = 1.19$			
VAR(1)	$\chi^2(2)$ =2.415	$\chi^2(1)$ =0.879	−0.327 (0.324)
VAR(2)	$\chi^2(4)$ =12.244	$\chi^2(3)$ =12.061	0.177 (0.394)
Non-Durable $\lambda = 1.28$			
VAR(1)	$\chi^2(2)$ =1.161	$\chi^2(1)$ =1.157	−0.366 (0.288)
VAR(2)	$\chi^2(4)$ =12.164	$\chi^2(3)$ =12.125	−0.069 (0.358)

of the actual to the predicted standard deviations reported in the last column continues to be less than one for most cases[17]. The $VAR(2)$ systems always reject the restrictions associated with REPI. For the UK, the results are similar; the $VAR(2)$ cases reject the permanent income hypothesis and the ratio is less than one. The story is different for the $VAR(1)$ cases where the restrictions are met and the ratio is always greater than one.

The LR tests are consistent with the Wald test results of Tables 7 and 8. Again the $VAR(1)$ framework fails to reject the orthogonality conditions at the 5% significance level for both countries. Since the excess sensitivity condition is a more general restriction than the REPI one, it can never be rejected for this $VAR(1)$ case. However, the sensitivity coefficient is always negative and insignificant thus suggesting that for the $VAR(1)$ case REPI provides a better representation for consumption behaviour. For the $VAR(2)$ case, the orthogonality restrictions are always rejected for both countries with the exception of the total consumption measure in the US. The sensitivity restrictions fail for both non-durable consumption measures in the US but cannot be rejected for all consumption measures in the UK and for total consumption in the US. However, the sensitivity parameter continues to be negative and insignificant in all cases, thus rejecting the excess sensitivity hypothesis proposed by Flavin. Our results continue

[17]The ratio is (statistically) close to one however.

Table 5.10: LR Tests for Orthogonality and Sensitivityfor the UK

	Orthogonality	Sensitivity	*value of* β (*s.e.*)
Total Consumption			
VAR(1)	$\chi^2(2) = 5.647$	$\chi^2(1) = 1.436$	−0.890 (0.619)
VAR(2)	$\chi^2(4) = 14.985$	$\chi^2(3) = 6.237$	−1.767 (1.225)
Non-Durable $\lambda = 1.099$			
VAR(1)	$\chi^2(2) = 5.029$	$\chi^2(1) = 1.499$	−0.739 (0.542)
VAR(2)	$\chi^2(4) = 15.75$	$\chi^2(3) = 5.749$	−1.748 (1.262)
Non-Durable $\lambda = 1.25$			
VAR(1)	$\chi^2(2) = 3.08$	$\chi^2(1) = 2.377$	−0.739 (0.491)
VAR(2)	$\chi^2(4) = 17.305$	$\chi^2(3) = 7.541$	−1.629 (1.051)

to confirm what we found in chapter 3; that with the observations for the later part of the 1980s and 1990s, the excess sensitivity hypothesis fails. What is puzzling however, is the significance of the REPI restrictions for the $VAR(1)$ case as these cannot reject REPI.

Levels Data

To complete our analysis of the post 1973 data set against the results in chapter 3, we examine the equivalent restrictions for the levels data, the results being reported in Tables 11 (US) and 12 (UK). To determine the appropriate lag length of the VAR we continue to use the same methods as before. The AIC and SBC suggested a $VAR(2)$ system for both measures of consumption for the US, although in the UK, a $VAR(2)$ for total consumption and a $VAR(1)$ for non-durable consumption were preferred[18].

From Tables 11 and 12, we see that for the US, the levels data strongly rejects both the REPI and sensitivity hypotheses for both total and non-durable consumption measures. This is consistent with our results in Chapter 3 and Campbell's findings. For the UK, the story is puzzling once more; total consumption fails to reject both REPI and the sensitivity hypotheses,

[18]The weak implication of the REPI, that of Granger causality from saving to labour income, although not reported here, is satisfied for all measures of consumption for both countries apart from the $VAR(1)$ total consumption measure for the UK.

Table 5.11: LR Tests for Orthogonality and Sensitivity in the US

	Orthogonality	Sensitivity	*Value of β* *(s.e.)*
Total Consumption			
VAR(2)	$\chi^2(2)$ =7.793	$\chi^2(1)$ =7.275	−0.183 (0.267)
Non-Durable OLS			
VAR(2)	$\chi^2(4)$ =34.699	$\chi^2(3)$ =11.386	−0.337 (0.430)

Table 5.12: LR Tests for Orthogonalityand Sensitivity in the UK

	Orthogonality	Sensitivity	*Value of β* *(s.e.)*
Total Consumption			
VAR(1)	$\chi^2(2)$ =5.145	$\chi^2(1)$ =1.614	−0.874 (0.642)
Non-Durable OLS			
VAR(2)	$\chi^2(4)$ =14.322	$\chi^2(3)$ =6.589	−1.620 (1.163)

whilst non-durable consumption fails to reject the excess sensitivity hypothesis only. Note further that all the excess sensitivity coefficients are negative for both the US and the UK once more. The results are consistent with our previous findings.

5.3.3 How do the Equations Developed in Chapter 4 Fare Over the Post 1973 Sample Period?

We continue to use the same methods that we used earlier in this chapter to estimate the equations from chapter 4. Table 13 shows the preferred consumption specifications for the US and the UK[19]. What is notable is that the results for the UK are very similar to those reported earlier but all equation specifications differ in the US. The number of lags in $A(L)$ which represent the coefficients in the IMA process for labour income when imperfect information exists, increase or stay the same for the case of excess sensitivity/imperfect information, whilst for the partial adjustment case, these lags decrease. We note as before, that all equations (5.1) and (5.5) in the US and the UK never have the same common component $A(L)$.

Only in one case does the (unrestricted) excess sensitivity/imperfect information equation

[19]Note that in the post 1973 period, many of the equations suffer problems of non-normality in the residuals.

Table 5.13: Preferred Specifications for Equations (5.1) and (5.5)

US	Total	Nondurable OLS
eqn (5.1)	9 lags in Δy_t, 9 lags in Δc_t, 11 lags in v_t Ser Corr $\chi^2(15) = 3.84$ Normality $\chi^2(2) = 6.52$ $L_1 = -372.865$ $aic = 972.707$	3 lags in Δy_t, 3 lags in Δc_t, 3 lags in v_t Ser Corr $\chi^2(15) = 3.69$ Normality $\chi^2(2) = 15.7$ $L_1 = -361.871$ $aic = 895.983$
eqn (5.5)	0 lags in Δy_t, 3 lags in Δc_t, Ser Corr $\chi^2(15) = 4.59$ Normality $\chi^2(2) = 5.13$ $L_5 = -398.213$ $aic = 951.343$	1 lag in Δy_t, 4 lags in Δc_t, Ser Corr $\chi^2(15) = 11.34$ Normality $\chi^2(2) = 11.13$ $L_5 = -365.682$ $aic = 895.605$
UK		
eqn (5.1)	1 lag in Δy_t, 1 lag in Δc_t, 3 lags in v_t Ser Corr $\chi^2(15) = 9.09$ Normality $\chi^2(2) = 15.79$ $L_1 = -330.465$ $aic = 806.737$	1 lag in Δy_t, 1 lag in Δc_t, 1 lag in v_t Ser Corr $\chi^2(15) = 16.6$ Normality $\chi^2(2) = 4.67$ $L_1 = -320.004$ $aic = 781.815$
eqn (5.5)	1 lag in Δy_t, 4 lags in Δc_t, Ser Corr $\chi^2(15) = 6.56$ Normality $\chi^2(2) = 11.36$ $L_5 = -330.625$ $aic = 809.057$	1 lag in Δy_t, 2 lags in Δc_t, Ser Corr $\chi^2(15) = 15.75$ Normality $\chi^2(2) = 5.04$ $L_5 = -320.124$ $aic = 782.055$

Table 5.14: Tests of (Unrestricted) (5.1) against (5.2) and (5.4)

US	Total	Nondurable OLS
eqn (5.1)	9 lags in Δy_t, 9 lags in Δc_t, 11 lags in v_t $L_1 = -372.865$	3 lags in Δy_t, 3 lags in Δc_t, 3 lags in v_t $L_1 = -361.871$
eqn (5.2)	0 lags in Δy_t, 1 lag in v_t Ser Corr $\chi^2(15) = 21.63$ Normality $\chi^2(2) = 2.49$ Restriction: LR Test $LR : \chi^2(28) = 78.454$	0 lags in Δy_t, 1 lag in v_t Ser Corr $\chi^2(15) = 39.99$ Normality $\chi^2(2) = 0.94$ Restriction: LR Test $LR : \chi^2(8) = 55.794$
eqn (5.4)	11 lags in v_t Ser Corr $\chi^2(15) = 5.64$ Normality $\chi^2(2) = 29.35$ Restriction: LR Test $LR : \chi^2(22) = 63.444$	3 lags in v_t Ser Corr $\chi^2(15) = 13.76$ Normality $\chi^2(2) = 12.66$ Restriction: LR Test $LR : \chi^2(6) = 20.28$
UK		
eqn (5.1)	1 lag in Δy_t, 1 lag in Δc_t, 3 lags in v_t $L_1 = -330.465$	1 lag in Δy_t, 1 lag in Δc_t, 1 lag in v_t $L_1 = -320.004$
eqn (5.2)	0 lags in Δy_t, 1 lag in v_t Ser Corr $\chi^2(15) = 26.94$ Normality $\chi^2(2) = 13.24$ Restriction: LR Test $LR : \chi^2(4) = 20.28$	0 lags in Δy_t, 1 lag in v_t Ser Corr $\chi^2(15) = 22.31$ Normality $\chi^2(2) = 0.39$ Restriction: LR Test $LR : \chi^2(2) = 105.684$
eqn (5.4)	3 lags in v_t Ser Corr $\chi^2(15) = 29.51$ Normality $\chi^2(2) = 19.71$ Restriction: LR Test $LR : \chi^2(2) = 111.208$	1 lag in v_t Ser Corr $\chi^2(15) = 23.12$ Normality $\chi^2(2) = 0.68$ Restriction: LR Test $LR : \chi^2(2) = 121.288$

encompass the partial adjustment/imperfect information case. This is for US total consumption. Performing a Likelihood ratio test to determine which of the two specifications can explain the behaviour of the data best, we can see from the table that the $LR = 50.696$ is greater than the critical value of the $\chi^2(26) = 38.885$ at the 5% level. Even scaling by the Barlett factor of $\frac{1}{1.202}$ does not change this rejection of the appropriate restriction. In Table 14 using a Likelihood ratio test, we impose the restrictions that enable us to discriminate between more general and unrestricted version of the imperfect information/excess sensitivity equation and equations (5.2), (5.3) and (5.4).

All the restrictions are rejected at the 5% significance level even after performing the Barlett

Table 5.15: Tests on the Unrestricted Equation (5.1)

US	Total Consumption	NonDurable OLS
LR	$\chi^2(9)$ =35.886	$\chi^2(3)$ =14.184
ALR	$\chi^2(9)$ =27.652	$\chi^2(3)$ =12.816
β(s.e.)	0.328 (0.188)	-0.022 (0.094)
UK		
LR	$\chi^2(1)$ =6.913	$\chi^2(1)$ =9.153
ALR	$\chi^2(1)$ =6.417	$\chi^2(1)$ =8.496
β(s.e.)	0.743 (0.229)	0.613(0.235)

adjustment to the Likelihood ratio test. This continues to suggest that the Pischke model is better at explaining the behaviour of consumption for the US and the UK than the Goodfriend model since the lags in consumption are statistically significant. However, this is not evidence that equation (5.1) is the best fitting equation because it still requires the further (weak) restriction that the coefficients of the income terms be equal to the negative of the consumption coefficients scaled by the marginal propensity to consume out of transitory income. Table 15 reports both the results of imposing these (weak) restrictions with the Likelihood Ratio test and also the size and significance of the β coefficient obtained from the estimation of the restricted model.

In table 15 all the restrictions fail at the 5% level but all the β coefficients are significant (and have increased in size). For the US, all tests apart from the ALR test for non-durable consumption fail even at the 1% significance level and even in that case the coefficient of the marginal propensity to consume out of transitory income is negative. These results continue to be in line with those found earlier and are thus still consistent with Demery and Duck's rejections of the Pischke Imperfect Information model at the 5% .

In table 16, we turn to a closer inspection of equation (5.5) as we present the parameter estimates for this equation.

The dynamics are very similar to the full sample period (specially for the UK) and the size of the coefficients has changed little. As before, apart from total consumption in the US (where $\gamma_1 = 0$) adjustment appears to be complete after two quarters. We find the adjustment coefficients to be smaller than those which Attfield et al. found for the US and the UK, although these estimates are higher (for all measures apart from US total consumption) than the ones which we reported earlier. It seems that the speed of adjustment has increased for the post 1973

Table 5.16: Estimation Results for Competing Alternative (5.5)

US	Total	Nondurable OLS
γ_0(s.e.)	0.324(0.178)	0.223(0.116)
γ_1(s.e.)		0.582(0.156)
a_1(s.e.)	0.112(0.101)	0.377(0.101)
a_2(s.e.)	0.193(0.104)	-0.127(0.096)
a_3(s.e.)	0.383(0.101)	0.473(0.096)
a_4(s.e.)		-0.206(0.099)
$\sum a_i$	0.688	0.517
UK		
γ_0(s.e.)	0.493(0.206)	0.545(0.173)
γ_1(s.e.)	0.895(0.249)	0.870(0.221)
a_1(s.e.)	-0.168(0.107)	0.010(0.106)
a_2(s.e.)	0.173(0.101)	0.284(0.102)
a_3(s.e.)	0.297(0.101)	
$\sum a_i$	0.302	0.294

period. We also find, as before, that the partial adjustment coefficients appear to be higher for the UK than for the US, but the number of lags in the innovations of consumption continue to be higher in the US compared to the UK. We can therefore come to the same conclusions which we made when we examined the whole sample in the previous section.

5.3.4 Conclusions

The post 1973 data set produces no significant differences to the results which have either been found in the literature or in earlier parts of this thesis. Perhaps the notable exception to this conclusion is that REPI appears to be supported in a number of specific cases, namely when REPI restrictions are imposed on a $VAR(1)$ system in the US and UK[20]. One aspect that continues to be puzzling and should perhaps be addressed in the future by the literature is the significance of the excess sensitivity hypothesis. Whilst the restrictions that are imposed on the VAR analysis cannot usually reject Flavin's theory, the excess sensitivity coefficient is for the most part negative or insignificant. An explanation for such negative coefficient could be a starting point for future research.

The model of imperfect information appears to provide a superior explanation of consumption behaviour in the US and UK than the model of lagged information. Our hybrid model

[20]An explanation for this result could perhaps be found with financial liberalization; this could have made borrowing easier for consumers thus easing their liquidity constraints

(5.1) is also closer to being accepted for the UK economy at the 5% significance level than before, although the US equivalent specification is not. Finally, our study demonstrates that the significance of lagged consumption terms cannot be ignored[21] and is an important component of consumption behaviour for both the full sample and post 1973 sample in both the US and the UK[22].

[21] Both of the competing alternatives had significant lags in consumption changes.

[22] Galí (1991) has shown that models of consumption innovations that have lags in the dependent variable have excess smoothness. Since the data suggests that lagged consumption innovations are significant, excess smoothness appears to be the rule rather than the exception.

Chapter 6

Imperfect Information and the Aggregate Stochastic Implications of the Life Cycle Hypothesis

6.1 Introduction

In a recent paper Clarida (1991) has examined the aggregate implications of the Modigliani-Brumberg (1980) life cycle hypothesis (LCH hereafter) and in particular the implications that it had on the 'first and second moment properties of changes in per capita consumption' (pp. 865). Clarida demonstrates how under certain circumstances, a life cycle model is capable of explaining both the Deaton paradox and the excess sensitivity phenomenon. Clarida also shows that a drift in per capita consumption should exist and that it is a function of the drift in per capita income. However, when persistence is introduced to the income process, Clarida's model is not able to explain the excess smoothness phenomenon. Pischke (1995) has also shown that a model that incorporates imperfect information is capable of explaining the phenomena of excess sensitivity and smoothness. In this chapter we incorporate the concepts of imperfect information into the life cycle hypothesis through two different and simple income processes[1]. We favour the introduction of the imperfect information approach above the lagged

[1] In the derivations that follow we attempt to keep the assumptions that are made as close to the original papers as possible.

information model of Goodfriend on the evidence of the last chapter and Demery and Duck's (1999, 2000) conclusions for US and UK aggregate time series data where it was shown that lagged innovations in consumption were significant.

6.2 Introducing Imperfect Information into LCH: A Simple Difference Stationary Income Process with Drift

We consider a life cycle model which consists of n overlapping generations comprised of x members each. Individuals live for n periods and receive labour income y_t during the first $w < n$ periods of life (these are termed the 'working periods'). During the remaining $n - w$ periods (the 'retirement periods'), no labour income is received and consumption must be financed through previous accumulation of assets. The rate of interest, r, is assumed constant and capital markets are assumed perfect.

Labour income received by each worker is specific to that worker and it is presumed to have the following specification: $y_{nit} = g + y_{nit-1} + \varepsilon_t + u_{nit} - u_{nit-1}$where the i subscripts denote labour income innovations that are specific to the individual in generation n only and g is a drift term that is specific to that generation. Terms with a subscript t only are aggregate terms that are common to all agents in the economy. Note that we are assuming as in Pischke's simple income case that aggregate shocks are more persistent than individual ones. The average labour income for the generation is $(\Delta y_{nt} - g) = \frac{1}{x}\sum_1^x (\Delta y_{nit} - g) = \varepsilon_t$ since it is assumed that the individual specific innovations vanish upon aggregation of the generation. It is assumed that the average labour income received by each generation Δy_{nt} is identical. Total labour income in the economy is therefore $\sum_1^w \sum_1^x (\Delta y_{nit} - g) = w \cdot x \cdot (\Delta y_t - g)$ and per capita labour income is equal to total labour income divided by the number of persons in this economy;

$$\frac{\sum_1^w \sum_1^x (\Delta y_{nit} - g)}{\sum_1^n x} = \frac{w}{n \cdot x} \sum_1^x (\Delta y_{nit} - g) = \frac{w}{n} (\Delta y_{nt} - g) = \frac{w}{n} \varepsilon_t = \epsilon_t \qquad (6.1)$$

Note that the innovation in total per capita income in our framework is the same as the total per capita income innovation in Clarida's paper (see pp. 855). Also note that the per capita drift component is the same as Clarida's which is equal to $\lambda = \frac{wg}{n}$. We feel that the way in which we aggregate the labour income components, such that they vanish at the generational level instead than at the total economy level, is closer in spirit to Friedman's permanent income hypothesis (and also to both Pischke and Clarida's models)[2]. It is also possible to introduce aggregation issues at the total economy level such that the labour income process becomes more persistent with each generation[3] but we do not purse these issues here[4].

When imperfect information exists we assume that the labour income process for the individual is $\Delta y_{nit} = g + \eta_{nit} - \theta\eta_{nit-1.}$ However, labour income received by each individual does not follow an $ARIMA\,(0,1,1)$ since by assumption labour income is zero with probability one during retirement. This means that we have the following,

$$E_{n=1;it}y_{n=1;it} = g + y_{n=1;it-1} + \eta_{n=1;it} - \theta\eta_{n=1;it-1}$$

$$E_{n=1;it}y_{n=1;it+\tau} = \tau g + y_{n=1;it-1} + (1-\theta)\,\eta_{n=1;it} - \theta\eta_{n=1;it-1}\ for\ \tau < w$$

$$E_{n=1;it}y_{n=1;it+\tau} = 0\ for\ \tau \geq w$$

$$E_{n=2;it}y_{n=2;it} = g + y_{n=2;it-1} + \eta_{n=2;it} - \theta\eta_{n=2;it-1}$$

$$E_{n=2;it}y_{n=2;it+\tau} = \tau g + y_{n=2;it-1} + (1-\theta)\,\eta_{n=2;it} - \theta\eta_{n=2;it-1}\ for\ \tau < w$$

$$E_{n=2;it}y_{n=2;it+\tau} = 0\ for\ \tau \geq w$$

and so on. The notation $n = \alpha$ indicates that the individual i is a member of the αth generation

[2]See the section on the permanent income hypothesis in chapter 2.

[3]This could, for example, represent technological improvements that lead to increased labour income with each generation.

[4]Note that this framework assumes that labour income innovations *at a given point in time* are age cohort specific. A possible justification for this point is that human capital has different vintages. (I am grateful to David Demery for making me aware of this point).

at time t. Workers are assumed to be rational and therefore to 'smooth completely current and expected future consumption subject to an expected present value budget constraint' (Clarida, pp. 855). This means that defining the present value of human and non-human wealth for a person in its jth period of life at time t as

$$\rho\left[A_{n=j;it}+\sum_{0}^{w-j}\rho^{\tau}E_{n=j;it}y_{n=j;it+\tau}\right]\quad for\ w\geq j$$

where $\rho=\frac{1}{1+r}$, and A denotes assets, then the consumption of someone in the jth period of life at time t is $\bar{c}_{n=j,i}$ where

$$\rho\sum_{0}^{n-j}\rho^{\tau}\bar{c}_{n=j;i}=\rho\left[A_{n=j;it}+\sum_{0}^{w-j}\rho^{\tau}E_{n=j;it}y_{n=j;it+\tau}\right]$$

$$\therefore\ c_{n=j;it}=\bar{c}_{n=j;i}=\frac{A_{n=j;it}+\sum_{0}^{w-j}\rho^{\tau}E_{n=j;it}y_{n=j;it+\tau}}{\varphi_1}$$

where $\varphi_1=\sum_{0}^{n-j}\rho^{\tau}$. From this it follows that the consumption of this same individual in the previous period when he or she was in his or her $(j-1)th$ period of life was

$$\begin{aligned}c_{n=j;it-1} &= \bar{c}_{n=j;i}\text{where}\\ \rho\sum_{0}^{n-(j-1)}\rho^{\tau}\bar{c}_{n=j;i} &= \rho\left[A_{n=j;it-1}+\sum_{0}^{w-(j-1)}\rho^{\tau}E_{n=j;it-1}y_{n=j;it+\tau-1}\right]\end{aligned}$$

$$\therefore\ c_{n=j,i;t-1}=\bar{c}_{n=j,i}=\frac{A_{n=j;it-1}+\sum_{0}^{w-(j-1)}\rho^{\tau}E_{n=j;it-1}y_{n=j;it+\tau-1}}{\varphi_2}$$

where $\varphi_2=\sum_{0}^{n-(j-1)}\rho^{\tau}$. So for $j\leq w$ we have

$$c_{n=j;it}=\frac{A_{n=j;it}+\sum_{0}^{w-j}\rho^{\tau}E_{n=j;it}y_{n=j;it+\tau}}{\varphi_1}$$

$$\frac{\varphi_2}{\rho \cdot \varphi_1} c_{n=j;it-1} = \frac{\frac{1}{\rho} A_{n=j;it-1} + \frac{1}{\rho} y_{n=j;it-1} + \sum_1^{w-(j-1)} \rho^{\tau+1} E_{n=j;it-1} y_{n=j;it+\tau-1}}{\varphi_1}$$

$$\therefore c_{n=j;it} - \frac{\varphi_2}{\rho \cdot \varphi_1} c_{n=j;it-1} = \frac{A_{n=j;it} - \left(\frac{1}{\rho} A_{n=j;it-1} + \frac{1}{\rho} y_{n=j;it-1}\right)}{\varphi_1} + \frac{\sum_0^{w-j} \rho^{\tau+1} \Delta E_{n=j;it} y_{n=j;it+\tau}}{\varphi_1}$$

Given the asset accumulation process, $A_t = \frac{1}{\rho}\left[A_{t-1} + y_{t-1} - c_{t-1}\right]$ we can write the last expression as

$$c_{n=j;it} - \left[\frac{\varphi_2}{\rho \cdot \varphi_1} - \frac{1}{\rho \cdot \varphi_1}\right] c_{n=j;it-1} =$$

$$c_{n=j;it} - c_{n=j;it-1} = \frac{\sum_0^{w-j} \rho^{\tau} \Delta E_{n=j;it} y_{n=j;it+\tau}}{\varphi_1}$$

since $\frac{\varphi_2}{\rho \cdot \varphi_1} - \frac{1}{\rho \varphi_1} = 1$. From the income process we have $\Delta E_{n=j;it} y_{it} = \eta_{n=j;it}$, $\Delta E_{n=j;it} y_{n=j+1;it+1} = \Delta E_{n=j;it}\left[y_{n=j;it} + \eta_{n=j+1;it+1} - \theta \eta_{n=j;it}\right] = (1-\theta)\, \eta_{n=j;it}$ and in general

$$\Delta E_{n=j,it} y_{n=j+\tau;it+\tau} = (1-\theta)\, \eta_{n=j;it}$$

and so we can write $\sum_0^{w-j} \rho^{\tau} \Delta E_{n=j;it} y_{n=j+\tau;it+\tau} = (1-\theta) \sum_0^{w-j} \rho^{\tau} \eta_{n=j;it} + \theta \eta_{n=j;it}$. Hence, the change in consumption is

$$c_{n=j;it} - c_{n=j;it-1} = (1-\theta)\, \mu(j)\, \eta_{n=j;it} + \frac{\theta}{\varphi_1} \eta_{n=j;it}$$

where $\mu(j) = \frac{1+\rho+\rho^2+\dots+\rho^{(w-j)}}{1+\rho+\rho^2+\dots+\rho^{(n-j)}} < 1$. This equation represents the consumption response associated to a labour income innovation at t for an individual i who is j years old at time t. Two special cases are worth mentioning [compare these to Clarida, pp. 856]; first, with infinite lifetimes and no retirement, i.e. $w = n$, then $\mu(j) = 1$ and the consumption innovation is the same as the one which is predicted by the imperfect information REPI. Second in Clarida's work, 'as $r \to 0$, $\mu(j) \to (w-j+1)/(n-j+1)$ which is the fraction of a *constant* labour income flow consumed in each period predicted by the textbook life cycle model [Dornbusch and Fischer, 1990, Ch. 8].' [Clarida, pp. 856]. In our case, as $r \to 0$, $c_{n=j;it} - c_{n=j;it-1} \to (1-\theta)\left(\frac{(w-j+1)}{(n-j+1)}\right) + \frac{\theta}{(n-j+1)} = \left(\frac{(w-j+1)}{(n-j+1)}\right) - \theta\left(\frac{(w-j)}{(n-j+1)}\right)$ which is less than the stated fraction of a constant labour income flow consumed in each period. With imperfect information, individuals smooth their consumption more during their working lifetimes than the normal LCH and REPI. We can re-write the last equation as

$$c_{n=j;it} - c_{n=j;it-1} = \left[(1-\theta)\,\mu(j) + \frac{\theta}{\varphi_1}\right]\eta_{n=j;it} = \left[(1-\theta)\,\mu(j) + \frac{\theta}{\varphi_1}\right]\frac{\Delta y_{n=j;it} - g}{1-\theta L}$$

provided the income process is invertible[5]. Since there are x individuals of this type in the economy, the average consumption change for this generation is

$$c_{n=j;t} - c_{n=j;t-1} = \left[(1-\theta)\,\mu(j) + \frac{\theta}{\varphi_1}\right]\frac{\Delta y_t - g}{1-\theta L} = \left[(1-\theta)\,\mu(j) + \frac{\theta}{\varphi_1}\right]\frac{\varepsilon_t}{1-\theta L}$$

where we have dropped the i subscript after aggregation and the n subscript as it is assumed that the labour income received by each generation is the same (Clarida, pp. 855, makes the same assumption). Hence the generational change in consumption *is different* from that which Clarida obtained since this change is autoregressive of order one provided $(1-\theta L)$ is invertible. As in Clarida, the generational change in consumption does not exhibit any drift components related to the labour income process. However, as our aggregation occurs at the generational level it is here where we encounter the characteristics of imperfect information and hence the autoregressive terms in the changes in consumption.

Letting C_t denote the economy's total consumption at date t, the change in the economy's total consumption is given by the following expression

$$\Delta C_t = \sum_{j=2}^{w}\left(c_{n=j;t} - c_{n=j;t-1}\right) + \left(c_{n=1;t} - c_{n=n;t-1}\right) \tag{6.2}$$

The first bracket represents the change in consumption for those individuals who are still working. We have already obtained an expression for this difference. The first term in the last bracket represents the consumption of those born at time t or those joining the labour force

[5] See chapter 4 for more details on the aggregation procedure.

then. The second term in the last bracket represents the consumption of those individuals that die at time t but who had consumed in the previous time period. We now seek an expression for these last two terms to derive an expression for the change in total consumption in the economy.

Before proceeding, note that we can write the income process in the following manner

$$\begin{aligned} y_{n=j;it+k} &= g[n+2] + kg + y_{n=j;it-(n+2)} + \sum_{1}^{n+1} \eta_{n=j;it-\tau} - \theta \sum_{1}^{n+1} \eta_{n=j;it-\tau-1} \\ &\quad + \eta_{n=j;it} - \theta \eta_{n=j;it-1} + \sum_{1}^{k} \eta_{n=j;it+\tau} - \theta \sum_{1}^{k} \eta_{n=j;it+\tau-1} \end{aligned}$$

Thus

$$\begin{aligned} E_{n=j;it} y_{n=j;it+k} &= g[n+2] + kg + y_{n=j;it-(n+2)} + \sum_{1}^{n+1} \eta_{n=j;it-\tau} - \theta \sum_{1}^{n+1} \eta_{n=j;it-\tau-1} \\ &\quad + (1-\theta)\, \eta_{n=j;it} - \theta \eta_{n=j;it-1} \end{aligned}$$

To derive the expression $c_{n=1;t}$ note that the generation that is 1 year old in period t, lives from period t to period $t+n-1$ (i.e. for n periods). It works from period t to period $t+w-1$ inclusive. At time t, the wealth of an individual in generation 1 ($n=1$) in present value terms is $\rho \sum_0^{w-1} \rho^\tau E_{n=1;it} y_{n=1+\tau;it+\tau}$ since it is assumed that no bequests exist in this economy. The planned consumption of this individual is $\bar{c}$ such that, $\rho \sum_0^{n-1} \rho^\tau \bar{c} = \rho \sum_0^{w-1} \rho^\tau E_{n=1;it} y_{n=1;it+\tau}$.

$$\therefore \bar{c} = c_{n=1;it} = \frac{\sum_0^{w-1} \rho^\tau E_{n=1;it} y_{n=1;it+\tau}}{\sum_0^{n-1} \rho^\tau}$$

Given the nature of the income process we can write

$$\begin{aligned} c_{n=1;it} &= \frac{\sum_0^{w-1} \rho^\tau}{\sum_0^{n-1} \rho^\tau} \left[g[n+2] + \tau g + y_{n=1;it-(n+2)} + \sum_{1}^{n+1} \eta_{n=1;it-\tau} - \theta \sum_{1}^{n+1} \eta_{n=1;it-\tau-1} \right] \\ &\quad + \frac{\sum_0^{w-1} \rho^\tau}{\sum_0^{n-1} \rho^\tau} \left[(1-\theta)\, \eta_{n=1;it} - \theta \eta_{n=1;it-1} \right] + \frac{1}{\sum_0^{n-1} \rho^\tau} \theta \eta_{n=1;it} \end{aligned}$$

Hence the level of consumption of this generation is equal to

$$\begin{aligned}c_{n=1;t} &= \frac{\sum_0^{w-1}\rho^\tau}{\sum_0^{n-1}\rho^\tau}\left[g\,[n+2]+\tau g+y_{n=1;t-(n+2)}+\sum_1^{n+1}\frac{\Delta y_{n=1;t-\tau}-g}{1-\theta L}-\theta\sum_1^{n+1}\frac{\Delta y_{n=1;t-\tau-1}-g}{1-\theta L}\right]\\&+\frac{\sum_0^{w-1}\rho^\tau}{\sum_0^{n-1}\rho^\tau}\left[(1-\theta)\frac{\Delta y_{n=1;t}-g}{1-\theta L}-\theta\frac{\Delta y_{n=1;t-1}-g}{1-\theta L}\right]+\frac{1}{\sum_0^{n-1}\rho^\tau}\frac{\theta\Delta y_{n=1;t}-g}{1-\theta L}\end{aligned}$$

provided the income process is invertible. Since the income processes are assumed equal for each generation we can drop the subscript $n=1$ without loss of generality.

The consumption of the representative agent in the generation that dies at the end of period $t-1$ (such representative agent would have been $n+1$ at t), $c_{n=n+1;it-1}$, will not have changed its level of consumption since the period it last worked. This generation lived from period $t-n$ to period $t-1$ inclusive. It worked from period $t-n$ to period $t-(n+1)+w$ inclusive. The consumption of an individual in this generation is equal to its consumption in the last period it worked which in turn is equal to its consumption in its first period alive, period $t-n$, plus[6] $\sum_2^w\left[\mu(j)(1-\theta)+\frac{\theta}{\varphi_1}\right]\eta_{n=n+1;it-(n+1)+j}$. Consumption in its first period alive, $t-n$, can be deduced as follows; at period $t-n$ the wealth of this individual in present value terms is $\rho\sum_0^{w-1}\rho^\tau E_{n=n+1;it-n}y_{n=n+1;it-n+\tau}$ and the planned constant level of consumption, $\bar{c}$, is such that $\rho\sum_0^{n-1}\rho^\tau\bar{c}=\rho\sum_0^{w-1}\rho^\tau E_{n=n+1;it-n}y_{n=n+1;it-n+\tau}$. Given the income process, we know that $E_{n=n+1;it-n}y_{n=n+1;it-n+\tau}=g\,[n+2]-ng+\tau g+y_{n=n+1;it-(n+2)}+(1-\theta)\,\eta_{n=n+1;it-n}+(1-\theta)\,\eta_{n=n+1;it-(n+1)}-\theta\eta_{n=n+1;it-(n+2)}$.

$$\begin{aligned}\therefore\bar{c} &= \frac{\sum_0^{w-1}\rho^\tau}{\sum_0^{n-1}\rho^\tau}\left[g\,[n+2]-ng+\tau g+y_{n=n+1;it-(n+2)}\right]\\&+\frac{\sum_0^{w-1}\rho^\tau}{\sum_0^{n-1}\rho^\tau}\left[(1-\theta)\,\eta_{n=n+1;it-n}+(1-\theta)\,\eta_{n=n+1;it-(n+1)}-\theta\eta_{n=n+1;it-(n+2)}\right]\\&+\frac{\theta\eta_{n=n+1;it-n}}{\sum_0^{n-1}\rho^\tau}\end{aligned}$$

[6] These represent the changes in consumption associated with innovations in the labour income process when this individual worked.

$$\begin{aligned}
\therefore c_{n=n+1;it-1} &= \bar{c}+\sum_{2}^{w}\left[\mu(j)(1-\theta)+\frac{\theta}{\varphi_1}\right]\eta_{n=n+1;it-(n+1)+j} \\
&= \mu(1)\left[y_{n=n+1;it-(n+2)}+g[n+2]-ng+\tau g\right] \\
&\quad +\mu(1)\left[(1-\theta)\eta_{n=n+1;it-n}+(1-\theta)\eta_{n=n+1;it-(n+1)}-\theta\eta_{n=n+1;it-(n+2)}\right] \\
&\quad +\frac{\theta\eta_{n=n+1;it-n}}{\sum_0^{n-1}\rho^\tau}+\sum_{2}^{w}\left[\mu(j)(1-\theta)+\frac{\theta}{\varphi_1}\right]\eta_{n=n+1;it-(n+1)+j}
\end{aligned}$$

Aggregating across the generation

$$\begin{aligned}
\therefore c_{n=n+1;t-1} &= \mu(1)\left[y_{n=n+1;t-(n+2)}+g[n+2]-ng+\tau g\right] \\
&\quad +\mu(1)\left[(1-\theta)\frac{\Delta y_{n=n+1;t-n}-g}{1-\theta L}+(1-\theta)\frac{\Delta y_{n=n+1;t-(n+1)}-g}{1-\theta L}\right] \\
&\quad -\mu(1)\theta\frac{\Delta y_{n=n+1;t-(n+2)}-g}{1-\theta L}+\frac{\frac{\theta\Delta y_{n=n+1;t-n}-g}{1-\theta L}}{\sum_0^{n-1}\rho^\tau} \\
&\quad +\sum_{2}^{w}\left[\mu(j)(1-\theta)+\frac{\theta}{\varphi_1}\right]\frac{\Delta y_{n=n+1;t-(n+1)+j}-g}{1-\theta L}
\end{aligned}$$

Because the income processes are equal for each generation we can again drop the subscript $n = n + 1$ without loss. The change in total consumption is therefore given by the following expression

$$\begin{aligned}
\Delta C_t &= \sum_{j=2}^{w}\left(\left[(1-\theta)\mu(j)+\frac{\theta}{\varphi_1}\right]\frac{\Delta y_t-g}{1-\theta L}\right)+\mu(1)\left[y_{t-(n+2)}+g[n+2]+\tau g\right] \\
&\quad +\mu(1)\left[\sum_{1}^{n+1}\frac{\Delta y_{t-\tau}-g}{1-\theta L}-\theta\sum_{1}^{n+1}\frac{\Delta y_{t-\tau-1}-g}{1-\theta L}+(1-\theta)\frac{\Delta y_t-g}{1-\theta L}-\theta\frac{\Delta y_{t-1}-g}{1-\theta L}\right] \\
&\quad +\frac{\frac{\theta\Delta y_t-g}{1-\theta L}}{\sum_0^{n-1}\rho^\tau}-\mu(1)\left[y_{t-(n+2)}+g[n+2]-ng+\tau g\right] \\
&\quad -\mu(1)\left[(1-\theta)\frac{\Delta y_{t-n}-g}{1-\theta L}+(1-\theta)\frac{\Delta y_{t-(n+1)}-g}{1-\theta L}-\theta\frac{\Delta y_{t-(n+2)}-g}{1-\theta L}\right] \\
&\quad -\frac{\frac{\theta\Delta y_{t-n}-g}{1-\theta L}}{\sum_0^{n-1}\rho^\tau}-\sum_{j=2}^{w}\left[(1-\theta)\mu(j)+\frac{\theta}{\varphi_1}\right]\frac{\Delta y_{t-(n+1)+j}-g}{1-\theta L}
\end{aligned}$$

Total consumption can be re-expressed as

$$\begin{aligned}(1-\theta L)\,\Delta C_t &= \mu(1)\,nh+\sum_{j=1}^{w}\left[(1-\theta)\,\mu(j)+\frac{\theta}{\varphi_1}\right]\varepsilon_t \\ &\quad -\sum_{j=1}^{w}\left[(1-\theta)\,\mu(j)+\frac{\theta}{\varphi_1}\right]\varepsilon_{t-(n+1)+j} \\ &\quad +\mu(1)\sum_{1}^{n+1}\varepsilon_{t-\tau}-\mu(1)\,\theta\sum_{0}^{n+1}\varepsilon_{t-\tau-1}+\mu(1)\,\theta\varepsilon_{t-(n+2)}-\mu(1)\,(1-\theta)\,\varepsilon_{t-(n+1)}\end{aligned}$$

where $h=(1-\theta)\,g$. Therefore per capita consumption, $\Delta c_t=\frac{\Delta C_t}{n}$, is equal to

$$\begin{aligned}(1-\theta L)\,\Delta c_t &= \frac{\mu(1)\,n\lambda}{w}+\sum_{j=1}^{w}\left[(1-\theta)\,\mu(j)+\frac{\theta}{\varphi_1}\right]\frac{\epsilon_t}{w} \\ &\quad -\sum_{j=1}^{w}\left[(1-\theta)\,\mu(j)+\frac{\theta}{\varphi_1}\right]\frac{\epsilon_{t-(n+1)}+j}{w}+\mu(1)\,(1-\theta)\sum_{1}^{n}\frac{\epsilon_{t-\tau}}{w}\end{aligned}$$

where, $\lambda=\frac{wh}{n}$, $\epsilon_t=\frac{w\varepsilon_t}{n}$. We write this simply as,

$$(1-\theta L)\,\Delta c_t=\psi\lambda+\pi\epsilon_t+\psi\xi_{t-1} \tag{6.3}$$

where

$$\pi=\left(\frac{\mu+\sum_{j=1}^{w}\frac{\theta}{\varphi_1}}{w}\right) \tag{6.4}$$

$$\mu=(1-\theta)\,[\mu(1)+\mu(2)+\cdots+\mu(w)] \tag{6.5}$$

$$\psi=\frac{\mu(1)\cdot n}{w} \tag{6.6}$$

and

$$\xi_{t-1}=\frac{(1-\theta)}{n}\sum_{1}^{n}\epsilon_{t-\tau}-\sum_{j=1}^{w}\left[(1-\theta)\,\mu(j)+\frac{\theta}{\varphi_1}\right]\epsilon_{t-(n+1)+j} \tag{6.7}$$

Two implications of the change in per capita consumption are worth noting; first, like in the models of imperfect information, we have lagged terms in per capita consumption[7]. Second, like in Clarida's paper, changes in per capita consumption are correlated with lagged innovations in labour income and the drift component in per capita consumption is a function of the drift component of the individual labour income process. As in Clarida's paper, it is aggregate per capita consumption changes which exhibit the drift component not the individual or generational changes. Compared to the standard permanent income case where $\pi = 1$ and $\psi = 0$, we note as Clarida did before us (Clarida, pp. 865), that the effect of a per capita labour income shock that is permanent from the point of the econometrician is not permanent from any household's perspective. In fact, the aggregate marginal propensity to consume out of these permanent shocks is less than unity.

Assuming as in Clarida's paper that $\xi_{t-1} = \frac{1}{n}\sum_1^n \epsilon_{t-\tau}$ is the maximum possible value for ξ_{t-1}, an upper variance on the unconditional variance of ξ_{t-1} is given by $\left(\frac{1}{n}\right)\sigma_\epsilon^2$ so that the upper bound on the unconditional variance of per capita consumption can be given by the following expression[8]

$$\sigma_{\Delta c}^2 = \left(\frac{A + \theta B}{1 - \theta^2}\right)\sigma_\epsilon^2 \tag{6.8}$$

The covariance between Δc_t and Δc_{t-1}, $cov\left(\Delta c_t, \Delta c_{t-1}\right)$ is given by,

$$cov\left(\Delta c_t, \Delta c_{t-1}\right) = \left(\frac{\theta A + B}{1 - \theta^2}\right)\sigma_\epsilon^2 \tag{6.9}$$

[7]The interpretation of this result is the standard one given to models of imperfect information where the aggregate shock is more persistent than the individual shock; following a (positive) aggregate shock to the economy, agents see their income change (an increase) but they cannot make up their minds as to whether the shock was a permanent (aggregate) one or a transitory (individual) one. Consumers will attribute part of this (positive) shock to the transitory component and some to the permanent part. They will therefore change (increase) their consumption but not as much as warranted by the life-cycle hypothesis. Since the shock is persistent, in the following period agents will be surprised again to see their income change (and be higher than expected) and so they will change (increase) their consumption again. This explains the autoregressive part.

[8]Derivation of the variance and the correlation terms are shown in the appendix.

where

$$\begin{aligned} A &= \pi^2 + \frac{1}{n}\psi\left[\frac{1}{n}\psi + \theta\pi\right] + \frac{1}{n}\psi\left[\frac{1}{n}\psi + \theta\frac{1}{n}\psi + \theta^2\pi\right] \\ &\quad + \frac{1}{n}\psi\left[\frac{1}{n}\psi + \theta\frac{1}{n}\psi + \theta^2\frac{1}{n}\psi + \theta^3\pi\right] + \cdots \\ &\quad + \frac{1}{n}\psi\left[\frac{1}{n}\psi\sum_{1}^{n-1}\theta^j + \theta^n\pi\right] \end{aligned}$$

and

$$\begin{aligned} B &= \frac{1}{n}\psi\pi + \frac{1}{n}\psi\left[\frac{1}{n}\psi + \theta\pi\right] + \frac{1}{n}\psi\left[\frac{1}{n}\psi + \theta\frac{1}{n}\psi + \theta^2\pi\right] \\ &\quad + \frac{1}{n}\psi\left[\frac{1}{n}\psi + \theta\frac{1}{n}\psi + \theta^2\frac{1}{n}\psi + \theta^3\pi\right] + \cdots \\ &\quad + \frac{1}{n}\psi\left[\frac{1}{n}\psi\sum_{1}^{n-2}\theta^j + \theta^{n-1}\pi\right] \end{aligned}$$

Note that if $\theta = 0$ is imposed the standard results obtained by Clarida (apart from the drift) result. An upper bound on the fraction of unconditional consumption variance accounted for by lagged income innovations (or $R^2\left(\Delta c, \xi_{t-1}\right)$ as Clarida terms it) is

$$R^2\left(\Delta c, \xi_{t-1}\right) = \frac{\psi \cdot cov(\Delta c, \xi_{t-1})}{var(\Delta c)} = \frac{\psi \cdot (C + D)}{var(\Delta c)} \tag{6.10}$$

where

$$C = \pi\left(\frac{\theta}{n} + \frac{\theta^2}{n} + \cdots + \frac{\theta^n}{n}\right)\sigma_\epsilon^2$$

and

$$D = \frac{\psi}{n^2}\left(\sum_{i=1}^{n}\sigma_\epsilon^2 + \theta\sum_{i=1}^{n-1}\sigma_\epsilon^2 + \theta^2\sum_{i=1}^{n-2}\sigma_\epsilon^2 + \cdots + \theta^{n-(n-(n+1))}\sum_{i=1}^{n-(n+1)}\sigma_\epsilon^2\right)$$

We are interested in knowing the implications of imperfect information on the life cycle model and so to compare this equation with Clarida's we select the same parameter values for

n, w, r and choose different values of θ. These results are shown in Table 1[9,10].

Comparing Table 1 with Clarida's Table I [$\theta = 0$], we see that introducing imperfect information into the model results in a lower marginal propensity to consume out of current income and a lower ratio of the standard deviation of changes in per capita consumption to the standard deviation of permanent shocks to labour income. On the other hand, the $R^2 \left(\Delta c, \xi_{t-1}\right)$ term is higher in the imperfect information case compared to Clarida's. This suggests that our framework generates more excess sensitivity and smoothness than Clarida's[11]. Since ψ is not a function of the imperfect information parameter, the marginal propensity to consume out of past labour income innovations is the same in our framework as in Clarida's. However, the term ξ_{t-1} is a function of this parameter and it decreases as θ increases. Thus it appears that the importance of the lagged income terms would diminish as θ increases.

From Table 1 we see that as the degree of imperfect information increases (i.e. as θ increases), the marginal propensity to consume out of *current* income innovations (π) decreases. Furthermore, the rate of decrease in this marginal propensity to consume increases as θ increases. This is a result that is consistent with models of imperfect information[12]; as θ increases (from 0 to 0.5), the proportion of consumption changes explained by current income innovations decreases and appears to be replaced by 'lagged income innovations' ($R^2 \left(\Delta c, \xi_{t-1}\right)$ increases)[13]. Note however, that as θ continues to increase (from 0.5 to 0.9) the proportion of consumption changes explained by past income innovations begins to decrease ($R^2 \left(\Delta c, \xi_{t-1}\right)$decreases). We can understand this more clearly if we take a closer look at the covariance between consumption changes and past income innovations (ξ_{t-1})[14]. This covariance increases as θ increases, thus the correlation between innovations in consumption and lagged income shocks increases as θ increases. This appears to increase the explanatory power of the 'lagged income terms' in

[9] Note that ψ is not a function of θ.

[10] The estimates in this table and the next were obtained using Mathematica.

[11] The exception is $\theta = 0.9$.

[12] Take Pischke's equation (9), pp. 812:

$$\Delta c_t = \theta \Delta c_{t-1} + A\varepsilon_t$$

where $A = \frac{1+r-\theta}{1+r}$. As θ increases, the proportion of consumption changes accounted by current income innovations decreases.

[13] For more details see table 3 in the appendix which shows the values of A, B, C and D which make up equations (6.8), (6.9) and (6.10).

[14] This covariance is reported in table 3 in the appendix.

Table 6.1: Replications for Finite Lives and Imperfect Information (No Persistence)
$\Delta y_t^{pc} = \lambda + \epsilon_t$; from $\Delta y_{nit} = g + \varepsilon_t + \Delta u_{nit}$.
$n = 50, w = 40, r = \delta$

	r=0.01(%change)	r=0.03(%change)	r=0.05(%change)
$\theta = 0$ (*Clarida*)			
π	0.64	0.69	0.73
$\sigma_{\Delta c}/\sigma_\epsilon$	0.65	0.71	0.75
ψ	1.05	1.12	1.17
$R^2(\Delta c, \xi_{t-1})$	0.05	0.05	0.05
$\theta = 0.1$			
π	0.58(-10%)	0.63(-9%)	0.67(-8%)
$\sigma_{\Delta c}/\sigma_\epsilon$	0.60(-8%)	0.65(-9%)	0.698(-7%)
$R^2(\Delta c, \xi_{t-1})$	0.072(44%)	0.069(38%)	0.066(32%)
$\theta = 0.2$			
π	0.52(-11%)	0.56(-12%)	0.6(-11%)
$\sigma_{\Delta c}/\sigma_\epsilon$	0.56(-7%)	0.61(-7%)	0.65(-7%)
$R^2(\Delta c, \xi_{t-1})$	0.096(33%)	0.092(33%)	0.088(33%)
$\theta = 0.3$			
π	0.46(-12%)	0.49(-13%)	0.53(-12%)
$\sigma_{\Delta c}/\sigma_\epsilon$	0.53(-6%)	0.58(-5%)	0.617(-5%)
$R^2(\Delta c, \xi_{t-1})$	0.126(31%)	0.115(25%)	0.115(30%)
$\theta = 0.5$			
π	0.34(-27%)	0.37(-25%)	0.4(-25%)
$\sigma_{\Delta c}/\sigma_\epsilon$	0.44(-17%)	0.55(-5%)	0.589(-5%)
$R^2(\Delta c, \xi_{t-1})$	0.26(106%)	0.187(62%)	0.181(57%)
$\theta = 0.7$			
π	0.22(-36%)	0.24(-35%)	0.27(-32%)
$\sigma_{\Delta c}/\sigma_\epsilon$	0.60(36%)	0.658(20%)	0.698(18%)
$R^2(\Delta c, \xi_{t-1})$	0.22(-16%)	0.21(12%)	0.208(15%)
$\theta = 0.9$			
π	0.1(-53%)	0.12(-50%)	0.13(-51%)
$\sigma_{\Delta c}/\sigma_\epsilon$	1.42(236%)	1.54(234%)	1.62(232%)
$R^2(\Delta c, \xi_{t-1})$	0.099(-55%)	0.097(-116%)	0.095(-55%)

explaining the variability of consumption changes, we see this in the $R^2(\Delta c, \xi_{t-1})$ term as it increases when θ increases from 0 to 0.5. However, as θ continues to increase, the importance of the lagged income terms starts to decrease. Whilst the covariance term continues to increase with the degree of imperfect information, the variability of such consumption changes increases even more and so these lagged income terms are not able to explain the behaviour of consumption innovations. In this framework, the lagged income innovations are not only present in the ξ_{t-1} term which is smaller in our case relative to Clarida, but *also* in the lagged consumption terms. Hence, we must take into account that as θ increases, the importance of the lagged consumption term increases, thus 'making' the lagged income terms 'more important'[15].

The ratio of the standard deviations of consumption changes and income innovations shows that excess smoothness occurs in all cases. The degree of smoothness is a function of the level of imperfect information (θ), the rate of interest, the working and lifetime periods (these are present in ψ) and the marginal propensity to consume (π which is itself a function of θ). As we move from a world of perfect information to a world of imperfect information (i.e. when θ is small but increasing) the ratio decreases, thus making consumption smoother compared to income. From the formula for the variance, we can begin to understand these results; as θ increases, the marginal propensity to consume out of current income decreases. This term appears to dominate the behaviour of A (see appendix) and as a result the numerator of this ratio decreases[16]. At the same time, the denominator decreases as θ increases, thus preventing the variability from decreasing even further. Economically speaking, the variability of consumption has decreased as consumers decide to 'reduce' their consumption out of current income innovations. At the same time, the importance of the lagged income innovations[17], either through lagged consumption or through the ξ_{t-1} term[18], although increasing is not high. But then the variability increases once more as the current and lagged income terms become less important whilst the consumption term becomes more and more important as θ approaches 1[19]. In this case, the 'numerator effect' (a lower marginal propensity leading to a lower A) is

[15] We see that that $cov(\Delta c_t, \Delta c_{t-1})$ increases as θ increases.

[16] The B term is small and is multiplied by $\theta < 1$.

[17] As represented by $R^2(\Delta c, \xi_{t-1})$.

[18] Above we saw that the importance of ξ_{t-1} decreased as θ increases. Thus, the importance of the lagged income terms is through past innovations in consumption.

[19] We should dismiss the posibility that $\theta = 1$ since that would imply that consumption changes are $I(1)$.

outweighed by the 'denominator effect' which now drives the results[20].

6.3 Introducing Imperfect Information into LCH: A Simple Difference Stationary Income Process with Drift and a Lag in Income

Despite Clarida's evidence from the Blinder and Deaton data set that income can be represented well by an $ARIMA(1,1,0)$ for annual data - Clarida's footnote 7, pp. 862 states that in an $ARIMA(4,1,1)$ for labour income none of the autoregressive coefficients are statisitically significant - Campbell and Deaton (1989) argued that changes in labour income are autocorrelated in quarterly data. In this section and following Clarida himself, we examine the implications for consumption changes within the life cycle model of an $ARIMA(1,1,1)$ process at the individual level. We do this to examine the implications for consumption of *introducing more persistence to the labour income process.* This is an interesting exercise because when Clarida introduced persistence to the labour income process, he found that the life cycle model suffered from excess smoothness (see the last equation in pp. 864). Thus, we want to see if introducing Pishcke's imperfect information when the labour income process is persistent removes the phenomenon of excess smoothness in a life cycle model. We assume the following labour income process is

Note that the closer this coefficient is to one, the more important the lagged innovations in income become in the labour income process. This suggests that consumers do not have any information about the labour income process since they would be observing a labour income process that is $I(0)$:

$$\begin{aligned}(1-L)\,y_{it} &= (1-L)\,\eta_{it}\\ y_{it} &= \eta_{it}\end{aligned}$$

whereas the original 'true' income representation is an $I(1)$.

[20]In Pischke's paper, an extension of REPI and where no lagged income terms affect consumption innovations, the standard deviation of consumption changes is

$$\frac{\sigma_{\Delta c}}{\sigma_{\epsilon}} = \frac{A}{\sqrt{1-\theta^2}}$$

In the case where the rate of interest is small,

$$\lim_{r\to 0}\frac{\sigma_{\Delta c}}{\sigma_{\epsilon}} = \sqrt{\frac{1-\theta}{1+\theta}}$$

Hence, as θ increases, the standard deviation ratio decreases.

the true one at the individual level for each working generation

$$y_{it} = g + (1+\phi)\, y_{it-1} - \phi y_{it-2} + \varepsilon_t + u_{it} - u_{it-1}$$

Given imperfect information the income process now looks like

$$\Delta y_{it} = g + \phi \Delta y_{it-1} + \eta_{it} - \theta \eta_{it-1}$$

In this case

$$\Delta E_{n=j;it} y_{n=j;it} = \eta_{n=j;it}$$

$$\Delta E_{n=j;it} y_{n=j;it+1} = (1+\phi)\, \eta_{n=j;it} - \theta \eta_{n=j;it}$$

$$\Delta E_{n=j;it} y_{n=j;it+2} = \left(1+\phi+\phi^2\right) \eta_{n=j;it} - (1+\phi)\, \theta \eta_{n=j;it}$$

$$\begin{aligned} \Delta E_{n=j;it} y_{n=j;it+\tau} &= \left(1+\phi+\phi^2+\cdots+\phi^{\tau}\right) \eta_{n=j;it} \\ -\left(1+\phi+\phi^2+\cdots+\phi^{\tau-1}\right) \theta \eta_{n=j;it-1}\ for\ \tau &< w \end{aligned}$$

$$\Delta E_{n=1;it} y_{n=1;it+\tau} = 0\ for\ \tau \geq w$$

Therefore

$$\begin{aligned} \sum_{0}^{w-j} \rho^{\tau} \Delta E_{n=j;it} y_{n=j;it+\tau} &= \eta_{n=j;it} + \rho\,[1+\phi-\theta]\, \eta_{n=j;it} + \rho^2 \left[\left(1+\phi+\phi^2\right) - \theta\,(1+\phi)\right] \eta_{n=j;it} \\ &+ \rho^3 \left[\left(1+\phi+\phi^2+\phi^3\right) - \theta\left(1+\phi+\phi^2\right)\right] \eta_{n=j;it} + \cdots \\ &+ \rho^{w-j} \left[\left(1+\phi+\cdots+\phi^{w-j}\right) - \theta\left(1+\phi+\cdots+\phi^{w-j-1}\right)\right] \eta_{n=j;it} \end{aligned}$$

The change in individual consumption is therefore given by

$$\begin{aligned} c_{n=j,i;t} - c_{n=j-1,i;t} &= \frac{\sum_0^{w-j} \rho^\tau \Delta E_{n=j;it} y_{n=j;it+\tau}}{\varphi_1} \\ &= \left[(1-\theta\rho)\,\mu(j;\phi) + \frac{\theta\rho\left(1+\phi+\cdots+\phi^{w-j}\right)\rho^{w-j}}{\varphi_1} \right] \eta_{n=j;it} \end{aligned}$$

where

$$\mu(j;\phi) = \frac{1+(1+\phi)\rho+\left(1+\phi+\phi^2\right)\rho^2+\cdots+\left(1+\phi+\phi^2+\cdots+\phi^{w-j}\right)\rho^{w-j}}{1+\rho+\rho^2+\cdots+\rho^{n-j}}$$

Again, the individual change in consumption is similar to Clarida's. The average consumption change for this generation is

$$\begin{aligned} c_{n=j;t} - c_{n=j-1;t} &= \left[(1-\theta\rho)\,\mu(j;\phi) + \frac{\theta\rho\left(1+\phi+\cdots+\phi^{w-j}\right)\rho^{w-j}}{\varphi_1} \right] \frac{\Delta y_t - \phi\Delta y_{t-1} - g}{1-\theta L} \\ &= \left[(1-\theta\rho)\,\mu(j;\phi) + \frac{\theta\rho\left(1+\phi+\cdots+\phi^{w-j}\right)\rho^{w-j}}{\varphi_1} \right] \frac{\varepsilon_t}{1-\theta L} \\ &= \pi(j;\phi)\frac{\varepsilon_t}{1-\theta L} \end{aligned}$$

where we have dropped the i subscript after aggregation and the n subscript as it is assumed that labour income received by each generation is the same. The generational consumption change is different from Clarida's as we have an autoregressive process for Δc.

Define $\pi(\phi) = \frac{1}{w}\sum_{j=1}^{j=w} \pi(j;\phi)$. Ignoring the impact of lagged income innovations[21,22] on

[21] i.e. the equivalent ξ_{t-1} term is ignored. From table 1, we saw that lagged income innovations should not be ignored specially for large values of θ. However, the derivation of the equivalent term ξ_{t-1} is not very easy and it becomes quite intractable.

[22] Note that we are not ignoring the lagged income terms per se, as we still have lagged consumption terms. For instance, take Pischke's example (pp. 812): Assume an econometrician estimates the following model

$$\Delta c_t = \alpha + \beta\Delta y_{t-1} + e_t$$

when the income process is given by

$$\Delta y_t = \phi\Delta y_{t-1} + \varepsilon_t$$

If the true data generation process is given by

$$\Delta c_t = \theta\Delta c_{t-1} + \pi(\phi)\,\varepsilon_t$$

the forecastability of consumption changes (Clarida, pp. 864), then

$$\left(\sigma_{\Delta c}^2 \mid \phi > 0\right) \cong \frac{\left[\pi\left(\phi\right)\right]^2}{1-\theta^2}\sigma_\epsilon^2$$

is an expression for the variance of per capita consumption changes. In Table 2 we examine the implications of this income process on the changes in consumption[23].

From Table 2, the first thing we note is that like in Clarida's paper, persistence in the labour income process continues to play an important role[24]. As in Table 1, the marginal propensity to consume out of current innovations in income (π), decreases as θ increases. The level of persistence (a higher ϕ) continues to increase this marginal propensity to consume, although imperfect information does seem to prevent the marginal propensity from being greater than one in most cases. Even without considering the effects of lagged income changes upon consumption, all the 'standard deviations' for low values of θ i.e. $\theta \leq 0.5$ are larger than the standard deviations encountered in Table 1. It would appear that persistence does in fact increase the standard deviation of consumption changes.

---

then the expected value of $\hat{\beta}$ would be

$$\begin{aligned}\hat{\beta} &= \frac{cov\left(\Delta c_t, \Delta y_{t-1}\right)}{var\left(\Delta y_{t-1}\right)} = \frac{E\left\{\pi\left(\phi\right)\left(\frac{\varepsilon_t}{1-\theta L}\right)\left(\frac{\varepsilon_{t-1}}{1-\phi L}\right)\right\}}{\left(1-\phi^2\right)^{-1}\sigma_\epsilon} \\ &= \frac{\frac{\theta}{1-\theta\phi}\pi\left(\phi\right)\sigma_\epsilon}{\left(1-\phi^2\right)^{-1}\sigma_\epsilon} = \frac{\left(1-\phi^2\right)\theta\pi\left(\phi\right)}{1-\theta\phi}\end{aligned}$$

which is non-zero, thus suggesting that lagged income terms are still present in this model.

[23]The values of $\phi = 0.2$ and $\phi = 0.4$ were the cases that Clarida examined. We do this to compare results with Clarida.

[24]If we had assumed the same income process in Pischke's paper; $y_{it} = g + \left(1+\phi\right)y_{it-1} - \phi y_{it-2} + \varepsilon_t + u_{it} - u_{it-1}$then,

$$\Delta c_t = \theta \Delta c_{t-1} + B\varepsilon_t$$

where $B = \frac{1+r-\theta}{1+r-\phi}$. The standard deviation of consumption changes would be in the case where $r \rightarrow 0$,

$$\lim_{r\rightarrow 0}\sigma_{\Delta c} = \frac{1}{1-\phi}\sqrt{\frac{1-\theta}{1+\theta}}\sigma_\epsilon$$

Hence, as the degree of persistence in the income process increases, i.e. as ϕ increases, the standard deviation increases. However, as in the no persistence case, as θ increases then the standard deviation decreases.

Table 6.2: Replications for Finite Lives and Imperfect Information (Persistence) $(1-\phi L)\,\Delta y_t^{pc} = \lambda + \epsilon_t$; from $(1-\phi L)\,\Delta y_{nit} = g + \varepsilon_t + \Delta u_{nit}$. $n = 50, w = 40, r = \delta$.

	$r = 0.01$ ϕ=0.2	$r = 0.03$ $\phi = 0.2$	$r = 0.05$ $\phi = 0.2$	$r = 0.01$ ϕ=0.4	$r = 0.03$ $\phi = 0.4$	$r = 0.05$ $\phi = 0.4$
$\theta = 0$						
π	0.78	0.85	0.89	1.01	1.09	1.15
$\frac{[\pi(\phi)]}{\sqrt{1-\theta^2}}$	0.78	0.85	0.89	1.01	1.09	1.15
$\theta = 0.1$						
π	0.70	0.76	0.81	0.91	0.98	1.04
$\frac{[\pi(\phi)]}{\sqrt{1-\theta^2}}$	0.7	0.76	0.81	0.91	0.98	1.04
$\theta = 0.3$						
π	0.55	0.6	0.64	0.72	0.78	0.83
$\frac{[\pi(\phi)]}{\sqrt{1-\theta^2}}$	0.58	0.63	0.67	0.75	0.82	0.87
$\theta = 0.5$						
π	0.41	0.44	0.48	0.53	0.57	0.61
$\frac{[\pi(\phi)]}{\sqrt{1-\theta^2}}$	0.47	0.51	0.55	0.61	0.66	0.7
$\theta = 0.7$						
π	0.26	0.29	0.32	0.34	0.37	0.41
$\frac{[\pi(\phi)]}{\sqrt{1-\theta^2}}$	0.36	0.41	0.44	0.48	0.52	0.57
$\theta = 0.9$						
π	0.12	0.14	0.16	0.16	0.18	0.2
$\frac{[\pi(\phi)]}{\sqrt{1-\theta^2}}$	0.27	0.32	0.36	0.37	0.41	0.46

6.4 Conclusion

This chapter has introduced Pischke's (1995) concepts of imperfect information into a model of the life-cycle hypothesis which has a quadratic utility function. The focus of our paper has been on the first and second moment properties of changes in per capita consumption. As in Clarida (1991) and Pischke before us, we find that smooth per capita consumption in the presence of permanent shocks to per capita labour income is the outcome that one should expect. The story goes something like this; in the model, saving for retirement as well as for consumption smoothing is a motive for asset accumulation thus creating a marginal propensity to consume out of current innovations in labour income that is less than one. The 'direct effect' played by lagged income innovations in per capita consumption in this model is the result of the generation that has retired and is using those accumulated assets (built from previous innovations in labour income) to finance retirement. The imperfect information that is introduced to this model appears to *reinforce the effects of a lifetime span.* The marginal propensity to consume out of current labour income innovations is even less in this framework than in Clarida's and the importance of the lagged innovations in labour income is thus increased, not only through ξ_{t-1} (which is the channel examined by Clarida) but also through the effects of lagged consumption innovations. Finally, the impact of persistence in the labour income process was briefly examined in a simpler version of our model. We found that persistence still exerts an important role in the so-called 'Deaton's Paradox' although the strength of this Paradox is greatly diminished by the introduction of imperfect information.

6.5 Appendix 1: Derivation of $\sigma^2_{\Delta c}$, $cov\left(\Delta c_t, \Delta c_{t-1}\right)$, $R^2\left(\Delta c, \xi_{t-1}\right)$ and values for A, B, C, D, and $cov\left(\Delta c_t, \Delta c_{t-1}\right)$, $cov(\Delta c, \xi_{t-1})$

In the text we have $(1-\theta L)\,\Delta c_t = \psi\lambda + \pi\epsilon_t + \psi\xi_{t-1}$. Express it more conveniently as $x_t - \theta x_{t-1} = \pi\epsilon_t + \psi\xi_{t-1}$ and take $\xi_{t-1} = \frac{1}{n}\sum_1^n \epsilon_{t-j}$ as the maximum possible value for this expression. Thus we have

$$x_t - \theta x_{t-1} \approx \pi\epsilon_t + \frac{1}{n}\psi\sum_1^n \epsilon_{t-j} \tag{6.11}$$

Multiply this expression by x_{t-k}, $k \geq 0$ and take expectations

$$\gamma_k - \theta\gamma_{k-1} = E\left[\pi\epsilon_t x_{t-k}\right] + E\left[\frac{1}{n}\psi\epsilon_{t-1}x_{t-k}\right] + E\left[\frac{1}{n}\psi\epsilon_{t-2}x_{t-k}\right] + \cdots + E\left[\frac{1}{n}\psi\epsilon_{t-n}x_{t-k}\right]$$

where $\gamma_k = E\left[x_t x_{t-k}\right]$ denotes the covariance between x_t and x_{t-k}. If $k=0$,

$$\gamma_0 - \theta\gamma_{-1} = E\left[\pi\epsilon_t x_t\right] + E\left[\frac{1}{n}\psi\epsilon_{t-1}x_t\right] + E\left[\frac{1}{n}\psi\epsilon_{t-2}x_t\right] + \cdots + E\left[\frac{1}{n}\psi\epsilon_{t-n}x_t\right] = A \tag{6.12}$$

and $k=1$,

$$\gamma_1 - \theta\gamma_0 = E\left[\pi\epsilon_t x_{t-1}\right] + E\left[\frac{1}{n}\psi\epsilon_{t-1}x_{t-1}\right] + E\left[\frac{1}{n}\psi\epsilon_{t-2}x_{t-1}\right] + \cdots E\left[\frac{1}{n}\psi\epsilon_{t-n}x_{t-1}\right] = B \tag{6.13}$$

where γ_0 is the variance for the autoregressive process (6.11) and $\gamma_{-1} = \gamma_1$. Substituting (6.13) into (6.12) one obtains the variance

$$var(\Delta c) = \gamma_0 = \frac{A + B\theta}{1-\theta^2}$$

and $cov\left(\Delta c_t, \Delta c_{t-1}\right)$,

$$cov\left(\Delta c_t, \Delta c_{t-1}\right) = \gamma_1 = \frac{\theta A + B}{1-\theta^2}$$

where A and B are as defined in the text.

For the derivation of the R^2 expression, recall that

$$R^2\left(\Delta c, \xi_{t-1}\right) = \frac{\psi \cdot cov(\Delta c, \xi_{t-1})}{var(\Delta c)}$$

and we have defined $var(\Delta c)$. We need to obtain an expression for $cov(\Delta c, \xi_{t-1})$. This is given by

$$cov(\Delta c, \xi_{t-1}) = E\left\{\left(\frac{\pi\epsilon_t + \psi\xi_{t-1}}{1-\theta L}\right)\cdot \xi_{t-1}\right\}$$

We have

$$\begin{aligned}\frac{\pi\epsilon_t + \psi\xi_{t-1}}{1-\theta L} &= \lim_{j\to\infty}\left[1+\theta L + (\theta L)^2 + \cdots (\theta L)^j\right]\left[\pi\epsilon_t + \psi\xi_{t-1}\right] \\ &= \lim_{j\to\infty}\left[1+\theta L + (\theta L)^2 + \cdots (\theta L)^j\right]\left[\pi\epsilon_t + \frac{1}{n}\psi\sum_1^n \epsilon_{t-j}\right]\end{aligned}$$

Now,

$$E\left\{\lim_{j\to\infty}\left[1+\theta L + (\theta L)^2 + \cdots (\theta L)^j\right]\left[\pi\epsilon_t\right]\cdot \frac{1}{n}\sum_1^n \epsilon_{t-j}\right\} = C$$

and

$$E\left\{\lim_{j\to\infty}\left[1+\theta L + (\theta L)^2 + \cdots (\theta L)^j\right]\left[\psi\xi_{t-1}\right]\cdot \frac{1}{n}\sum_1^n \epsilon_{t-j}\right\} = D$$

where C and D are as defined in the text.

Table 6.3: Some Results from the Main Text
Values of A, B, C, D, $cov(\Delta c, \xi_{t-1})$, $n = 50, w = 40, r = \delta$

	r=0.01	r=0.03	r=0.05
$\theta = 0.1$			
A	0.359	0.422	0.478
B	0.037	0.043	0.047
C	0.001	0.001	0.001
D	0.023	0.024	0.026
$cov(\Delta c, \xi_{t-1})$	0.025	0.029	0.032
$cov(\Delta c_t, \Delta c_{t-1})$	0.073	0.086	0.957
$\theta = 0.2$			
A	0.298	0.351	0.398
B	0.040	0.046	0.051
C	0.002	0.003	0.003
D	0.026	0.027	0.029
$cov(\Delta c, \xi_{t-1})$	0.030	0.034	0.037
$cov(\Delta c_t, \Delta c_{t-1})$	0.103	0.121	0.136
$\theta = 0.3$			
A	0.246	0.29	0.329
B	0.044	0.051	0.056
C	0.004	0.003	0.004
D	0.029	0.032	0.033
$cov(\Delta c, \xi_{t-1})$	0.035	0.038	0.044
$cov(\Delta c_t, \Delta c_{t-1})$	0.129	0.152	0.17
$\theta = 0.5$			
A	0.166	0.196	0.224
B	0.05	0.065	0.072
C	0.007	0.007	0.008
D	0.041	0.044	0.046
$cov(\Delta c, \xi_{t-1})$	0.05	0.057	0.063
$cov(\Delta c_t, \Delta c_{t-1})$	0.177	0.217	0.245
$\theta = 0.7$			
A	0.129	0.153	0.174
B	0.08	0.097	0.107
C	0.010	0.011	0.012
D	0.066	0.071	0.074
$cov(\Delta c, \xi_{t-1})$	0.08	0.092	0.107
$cov(\Delta c_t, \Delta c_{t-1})$	0.334	0.4	0.449
$\theta = 0.9$			
A	0.21	0.24	0.273
B	0.197	0.228	0.252
C	0.017	0.021	0.023
D	0.172	0.184	0.192
$cov(\Delta c, \xi_{t-1})$	0.199	0.23	0.252
$cov(\Delta c_t, \Delta c_{t-1})$	2.0316	2.337	2.619

Chapter 7

A Model of Consumption and Asset Management

7.1 Introduction

The main premise of Friedman's permanent income hypothesis is that agents will smooth their consumption to maximise their lifetime utility. Saving occurs when a consumer's current income is above its average (when current income is above permanent income) and borrowing when current income is below its average (when current income is below permanent income). These results extend to the rational expectations permanent income hypothesis (REPI); the smoothing behaviour of rational agents means that savings can provide sufficient information about the expected future levels of (labour) income that maybe unobservable to a researcher (Campbell (1987)). However, these results are dependent upon a number of assumptions. One of the assumptions of REPI is that capital markets are perfect and that any amount of lending or borrowing can take place at the constant market rate of interest. Among the most popular explanations given for the observed failure of REPI is the presumption that agents may be restricted in their ability to borrow. The consumption literature has examined the issue of liquidity constraints under different variants; the most popular being the solved out consumption problem (Campbell and Mankiw, 1989, 1991) and the Euler equation approaches under different assumptions but always under the premise that assets can never be negative (Zeldes, 1989a, Deaton, 1991, Seater, 1997). In this chapter we make it difficult for individuals to borrow in an intertemporal maximisation framework.

Zeldes examined the consumption problem for consumers with a CRRA utility function and under the assumption that assets cannot be negative (i.e. $A_t \geqslant 0$). Zeldes tested whether the Euler equation derived from the unconstrained problem is rejected when liquidity constraints exist, and using panel data he found that liquidity constraints have 'important influences on consumption' [pp. 307].

Deaton examined the consumption problem with a utility function that is increasing, strictly concave and convex in its first difference[1] and where assets cannot be negative. He also assumed that consumers are 'impatient' in the sense that the rate of time preference is greater than the rate of interest (this assumption is needed to solve the dynamic programming problem). Deaton found some evidence for a precautionary motive for holding assets regardless of the assumptions about the persistence of the labour income process.

Seater examined liquidity constraints (assets must be nonnegative) using optimal control techniques and assuming that income uncertainty is not present[2]. His solution to the problem does not have a 'closed form' (pp. 130), but his Euler equation for constrained households is equivalent to Zeldes's Euler equation (pp. 133). Seater claims that the properties of the solution depend on the nature of the paths of the rate of interest and income, although in his discussions he maintains a constant rate of interest. His main conclusion is similar to that of Mariger (1987) and Zeldes: *constrained agents need not have zero assets.* For liquidity constrained agents, consumption will tend to follow income when assets are zero and income is expected to increase, assets will be accumulated if income is expected to decrease.

The aim of this chapter is to retain most of the assumptions of REPI but to modify the problem to account for the fact that individuals may find it costly or difficult to borrow.

This chapter is divided into two sections. The first section introduces the problem and obtains a solution for consumption in levels and also its innovation. The second section introduces a number of tests on the equations derived in the first section and extends the results from chapter 5.

---

[1] A convex marginal utility function implies prudent behaviour in the face of uncertainty (in future earnings). See Leland (1968), Rothschild and Stiglitz (1971), Kimball (1990) for more discussion on this point.

[2] In Seater's paper there is no evidence that the marginal utility function is convex, thus there is no 'prudence' effect per se.

7.2 A Simple Problem of Consumption and Asset Management

Consider the following problem,

$$\min\ L = E_t \sum_{i=0}^{\infty} \rho^i \left\{ \left(\frac{1}{2}\right)(b - c_{t+i})^2 + \frac{\alpha}{2}(A_{t+i} - A_{t+i-1})^2 - \beta A_{t+i} - \phi(A_{t+i} - A_{t+i-1}) \right\} \tag{7.1}$$

subject to

$$A_{t+1} = R(A_t + y_t - c_t) \tag{7.2}$$

and

$$C > A_t > B\ \forall\, t \geq 0 \tag{7.3}$$

$$A_t \ given \tag{7.4}$$

$$\alpha,\ \beta,\ \phi > 0$$

where b denotes the bliss point, y labour income, A assets, c consumption, $R = 1 + r$ where r is the rate of interest, $\rho = 1/(1+\delta)$ where δ is the subjective rate of discount. The consumer faces an unknown sequence of labour income terms, $\{y_{t+i}\}_{i=0}^{\infty}$, and it is assumed[3] that for all t there is some $K > 0$, such that $|y_t| < Kx^t$, where $1 \leq x \leq 1/\sqrt{\rho}$ (i.e. the sequence is of exponential order less than $1/\sqrt{\rho}$). We assume further that $\infty > C > 0$ and $0 > B > -\infty$, so that (7.3) requires that assets be bounded from above and below. Thus, assets cannot be either positively or negatively infinity. It is assumed as in Hall (1978), Flavin (1981) and others that $R\rho = 1$.

[3]The arguments here follow those of Sargent's (1987) pp. 200 closely. Series that satisfity this condition are not allowed to grow very fast.

7.2.1 Economic Interpretations of the Problem

The first term in the loss function (7.1), $(b - c_{t+i})^2$, is a standard quadratic utility function. It represents the utility loss of being away from the bliss point b. The remaining terms make it difficult for the consumer to hold negative assets for prolonged periods of time. The second term, $\frac{\alpha}{2}(A_{t+i} - A_{t+i-1})^2$, does not allow assets to be changed rapidly; the third term, βA_{t+i}, makes it costly for the consumer to hold negative assets but rewards the individual if he or she holds positive assets and the fourth term, $\phi(A_{t+i} - A_{t+i-1})$, tells us that running down assets becomes more and more expensive[4].

We do not necessarily view the inclusion of the asset terms as representing a monetary cost or benefit to the agent (e.g. a cost if assets are negative). Ball and Drake (1964) have argued that to introduce wealth in the utility function in an intertemporal utility maximization problem is not unreasonable: 'If an individual adjusts his future consumption to his initial asset holdings and present value of expected income from human wealth, we are essentially deprived of the notion of an excess or undesired holding of assets. Given any set of independent parameters, variations in the initial stock of assets will simply result in an alteration in future consumption rather than any explicit decision to readjust asset holdings, which simply amounts to saying that given the set of parameters there is no unique equilibrium value of the asset stock. [...] The result of this is to rule out the application of short period stock adjustment analysis to asset holdings by consumers, and this makes it virtually impossible to make any sense of the concept of excess demand as applied to a single individual.' [pp. 64]. Introducing asset terms in the loss function is desirable as it does help get around this problem without having to model wealth specifically.

The second term may have another interpretation as the loss in utility associated with having to monitor the level of assets whether these assets are negative or positive. Given habits, etc. agents will not want to change or monitor their asset level[5]. The other two terms could also represent a 'psychic' cost argument associated with holding negative assets

[4] Individuals may have to pay more to finance more borrowing.

[5] One can view this term as having a similar interpretation to the term $\frac{\alpha}{2}(c_{t+i} - c_{t+i-1})^2$ in the paper on adjustment costs for consumption of Attfield et al. (1992). Both quadratic terms do not permit rapid adjustment, but assume slow progressive adjustments in both consumption and asset holdings.

or debt[6]. The Journal of Economic Psychology (1996) has a special edition on these issues. Walker (1996) for example, extends the literature on financial management to explore the inter-relationships between feelings of coping (with debt), debt measured in different scales and financial management (to get out of debt). Walker shows that good financial management on behalf of consumers contributes to getting and staying out of debt. More significantly, she shows that being in debt can lead to improvements in financial management as people seek to get out of debt. We may interpret this to mean that being in debt is an undesirable state which consumers attempt to avoid. This could imply that consumers that hold positive assets obtain some utility from holding them because they can use assets to buffer themselves against future contingencies[7].

Note that our framework allows for negative asset holdings since it may pay the agent to hold negative assets for a while to get close to the bliss point. However, it makes it increasingly difficult to allow individuals to sustain that position for a long time (the third term) or increase the level of debt (the second and fourth terms). This framework therefore penalises the reduction in the level of assets, specially when these are already negative. When agents seek to increase their level of assets, the problem above has two arguments that work in the opposite direction; the second term does penalise the increase in asset holdings since it is assumed that a utility cost is incurred when doing so as individuals are lazy, have habits, etc. However, the third and the fourth terms reward the individual's positive asset holdings.

Finally, the choice of the functional form (7.1) is determined by a number of other factors. First, in choosing quadratic utility, we ignore issues of prudence directly. This enables us to concentrate our efforts on the effects that borrowing may have on consumption and asset decisions. Carroll and Kimball (1999) have shown using theory that in a model where the utility function is quadratic and where hard borrowing constraints exist (i.e. agents are not allowed to borrow at all), consumption behaviour can resemble that which is encountered when precautionary savings exist. Moreover, they also show that it is difficult to distinguish between the effects of hard liquidity constraints and prudence on consumption when utility has a positive

[6]I would like to thank Edmund Cannon for helping me with the idea of a 'psychic cost' and thus pointing me in the direction of the psychology literature.

[7]Because agents know that it is expensive to run their assets down, they can build their assets against uncertainty, etc. This may be equivalent to 'prudent' behaviour or a 'precautionary motive' for holding assets.

third derivative. Second, we want to introduce soft-type liquidity constraints to the consumption problem to examine their effect on consumption and asset decisions. We do this by allowing individuals to be able to borrow but in doing so paying a higher premium. This is desirable on at least two counts: one is that this author does not believe that all consumers are denied credit and also because it enables us to make the consumption Euler equation continuous which allows us to find close form solutions for the consumption function. Studies that impose hard constraints have discontinuous Euler equations that cannot be solved analytically (see Zeldes, Deaton). Analytical solutions cannot also be obtained when utility functions are not quadratic or CARA. We believe that the inability to find a close form solution is a very high price to pay when hard constraints and/or CRRA utility functions are imposed and this led us to choose the quadratic forms in (7.1)[8].

7.2.2 A Technical Note

Before we proceed, note that if we seek to differentiate the following expression with respect to x_t

$$V = \sum_{t=0}^{\infty} b^t \left[a(L) z_t\right]\left[d(L) x_t\right] \tag{7.5}$$

where $a(L) = \sum_{j=-\infty}^{\infty} a_j L^j$, $d(L) = \sum_{j=-\infty}^{\infty} d_j L^j$, then

$$\frac{\partial V}{\partial x_t} = b^t d\left(bL^{-1}\right) a(L) z_t \tag{7.6}$$

We shall apply this formula to the problem above (see Sargent, 1987, pp. 213).

[8]When considering problems of this type, a researcher has the following choices: to choose between certainty equivalence and a convex marginal utility function and/or to choose between hard (discontinuous) or soft (continuous) liquidity constraints. Because this author is interested in finding an analytical solution to the consumption problem, quadratic utility and soft constraints are chosen. Any deviation from this choice (apart from CARA preferences) results in the impossibility of finding an analytical solution to the consumption problem.

7.2.3 The Problem Without Expectations

Before solving the problem substitute the constraint (7.2) into the objective function (7.1) and expand all the terms to give

$$\begin{aligned}
\min L &= \sum_{i=0}^{\infty}\rho^i\left\{\frac{1}{2}\left[(b-y_{t+i})^2+2(b-y_{t+i})\left(R^{-1}L^{-1}-1\right)A_{t+i}\right]\right\} \qquad (7.7)\\
&+\sum_{i=0}^{\infty}\rho^i\left\{\frac{1}{2}\left(\left(R^{-1}L^{-1}-1\right)A_{t+i}\right)^2+\frac{\alpha}{2}\left((1-L)A_{t+i}\right)^2-\beta A_{t+i}-\phi\left((1-L)A_{t+i}\right)\right\}
\end{aligned}$$

Applying (7.6) to the equation above and noting that if, $d(L)=\left(R^{-1}L^{-1}-1\right)$, where $d_{-1}=R^{-1}$, $d_0=1$, then $d\left(\rho L^{-1}\right)=\left(R^{-1}\rho^{-1}L-1\right)$, and if $d(L)=1-L$, then $d\left(\rho L^{-1}\right)=\left(1-\rho L^{-1}\right)$, one obtains the Euler equation[9]

$$\begin{aligned}
0 &= (b-y_t)\left(R^{-1}\rho^{-1}L-1\right)+\left(R^{-1}L^{-1}-1\right)\left(R^{-1}\rho^{-1}L-1\right)A_t \qquad (7.8)\\
&+\alpha(1-L)\left(1-\rho L^{-1}\right)A_t-\beta\rho L^{-1}-\phi\left(1-\rho L^{-1}\right)
\end{aligned}$$

or

$$\begin{aligned}
0 &= -y_t(L-1)+\left(R^{-1}L^{-1}-1\right)(L-1)A_t\\
&+\alpha(1-L)\left(1-\rho L^{-1}\right)A_t-\beta\rho-\phi(1-\rho)
\end{aligned}$$

since we assume that $R\rho=1$. The transversality condition for this problem is obtained with the following steps:

1. consider a finite T horizon version of the problem in question,
2. calculate the first order conditions for A_{t+T},
3. multiply the first order condition by A_{t+T},
4. take the limit of this condition as $T\to\infty$ and equate this limit to zero.

[9] To see how to solve this Euler equation without any restrictions see the appendix.

The first order condition with respect to A_{t+T} is

$$\begin{aligned}\frac{\partial L}{\partial A_{t+T}} &= \rho^{T-1}\left\{(b-y_{t+T-1})R^{-1}+(A_{t+T}-A_{t+T-1})R^{-1}\right\}\\ &\quad+\rho^{T}\left\{-(b-y_{t+T})-(A_{t+T+1}-A_{t+T})+\alpha(A_{t+T}-A_{t+T-1})-(\beta+\phi)\right\}\end{aligned}$$

We impose the restriction that $R\rho = 1$ and noting that A_{t+T+1} does not exist,

$$\begin{aligned}\frac{\partial L}{\partial A_{t+T}} &= \rho^{T}\left\{(b-y_{t+T-1})+(A_{t+T}-A_{t+T-1})\right\}\\ &\quad+\rho^{T}\left\{-(b-y_{t+T})+A_{t+T}+\alpha(A_{t+T}-A_{t+T-1})-(\beta+\phi)\right\}\end{aligned}$$

collecting terms gives the transversality condition,

$$\lim_{T\to\infty}\rho^{T}\left\{(y_{t+T}-y_{t+T-1})+(\alpha+2)A_{t+T}-(1-\alpha)A_{t+T-1}-(\beta+\phi)\right\}A_{t+T}=0 \qquad (7.9)$$

Sufficient conditions for the transversality condition[10] (7.9) to hold are, first, that $\{y_{t+i}\}_{i=0}^{\infty}$ is of exponential order less than $1/\sqrt{\rho}$, and second, that the solution for $\{A_{t+i}\}_{i=0}^{\infty}$ be of exponential order less than $1/\sqrt{\rho}$. Then notice that

$$\begin{aligned}&\left|\rho^{T}\left[y_{t+T}A_{t+T}-y_{t+T-1}A_{t+T}+(\alpha+2)A_{t+T}^{2}-(1-\alpha)A_{t+T-1}A_{t+T}-(\beta+\phi)A_{t+T}\right]\right|\\ \le\; &\rho^{T}\left[\left|y_{t+T}A_{t+T}\right|+\left|y_{t+T-1}A_{t+T}\right|+(\alpha+2)\left|A_{t+T}^{2}\right|+(1-\alpha)\left|A_{t+T-1}A_{t+T}\right|+(\beta+\phi)\left|A_{t+T}\right|\right]\\ \le\; &\rho^{T}Kx^{2(t+T)}+\rho^{T}Kx^{2(t+T)-1}+\rho^{T}(\alpha+2)Kx^{2(t+T)}\\ &+\rho^{T}(1-\alpha)Kx^{2(t+T)-1}+\rho^{T}(\beta+\phi)Kx^{t+T}\\ =\; &(\beta+\phi)Kx^{t}\rho^{T}x^{T}+(\alpha+3)Kx^{2t}\rho^{T}x^{2T}+(2-\alpha)Kx^{2t-1}\rho^{T}x^{2T}\end{aligned}$$

From the definition of a sequence of exponential order less than $1/\sqrt{\rho}$, we have that $0 < x\sqrt{\rho} < 1$. This implies that $x\rho < x\sqrt{\rho} < 1$, and since we are taking the limit $T \to \infty$, we see that the last line in the expression above will tend to zero.

[10] See Sargent (1987) pp. 201.

The necessary conditions for optimality for the infinite horizon problem are satisfied when we find a solution to the difference (Euler) equation (7.8). Re-express the Euler equation as

$$(1+\alpha)(1-L)\left(R^{-1}L^{-1}-1\right)A_t=(1-L)\,y_t-\beta\rho-\phi(1-\rho) \tag{7.10}$$

The zeros of the characteristic polynomial in A_t are at $L=R^{-1}<1$ and $L=1$. To prevent assets from ever accumulating[11], the polynomial $\left(R^{-1}L^{-1}-1\right)$ is solved forwards. Divide (7.10) by $\left(R^{-1}L^{-1}-1\right)$ to obtain

$$(1+\alpha)A_t=(1+\alpha)A_{t-1}+\frac{(1-L)}{(R^{-1}L^{-1}-1)}y_t-\left(\frac{\beta\rho+\phi(1-\rho)}{(R^{-1}-1)}\right) \tag{7.11}$$

Substitute the budget constraint (7.2) into (7.11)

$$(1+\alpha)Rc_{t-1}=(1+\alpha)R\left(A_{t-1}+y_{t-1}\right)-(1+\alpha)A_{t-1}-\frac{(1-L)\,y_t}{(R^{-1}L^{-1}-1)}+\left(\frac{\beta\rho+\phi(1-\rho)}{(R^{-1}-1)}\right)$$

We expand the terms related to y_t and shift forward everything by one period to obtain

$$\begin{aligned} c_t &= \left(1-R^{-1}\right)A_t+\frac{R^{-1}-(1+\alpha)}{(R^{-1}L^{-1}-1)(1+\alpha)}y_t+\frac{\alpha R^{-1}}{(R^{-1}L^{-1}-1)(1+\alpha)}y_{t+1} \\ &\quad +\left(\frac{(\beta\rho+\phi(1-\rho))R^{-1}}{(R^{-1}-1)(1+\alpha)}\right) \end{aligned}$$

[11] Strictly speaking, in equation (10) assets will continue to accumulate or decrease forever because the polynomial $(1-L)$ has a unit root. This is is consistent with this variant of the (standard) permanent income case where it is assumed that the rate of interest is equal to the rate of time preference (see Sargent, pp. 214 equation (106) and assume that $Rb=1$ which he assumes later on in the derivation, or footnote 5 in pp. 366 where Sargent explains that this model of consumption does have unsatisfactory qualities for either assets or consumption; also see Deaton, 1992, pp. 196, equation 22). A way to 'overcome' this problem is to substitute the budget constraint into the Euler equation and find the solution for consumption. Once this is done note that there are no situations where assets or any of the other arguments continue to accumulate or decrease in the sense that there are no unit roots. In fact, in Charles H. Whiteman's book about solutions to Sargent's 2nd edition macroeconomic problems, we face a similar situation in problem 5 (the solution to the problem is given in pages 204-210) and problem 7 (the solution in pages pp. 213-8) of chapter XIV. There, the Euler equation for the sequence which we are interested in solving also posseses a unit root. Whiteman gets around this problem by substituting the demand curve and rearranging thus ensuring that the resulting expression does not have a unit root. This is what happens to our problem once we substitute the budget constraint to the Euler equation.

This can be re-expressed as

$$
\begin{aligned}
c_t &= \left(1-R^{-1}\right)\left[A_t+\frac{1}{(1+\alpha)}\sum_{j=0}^{\infty}R^{-j}y_{t+j}\right]+\left(\frac{(\beta\rho+\phi(1-\rho))R^{-1}}{(R^{-1}-1)(1+\alpha)}\right) \\
&+\frac{\alpha}{(1+\alpha)}\sum_{j=0}^{\infty}R^{-j}y_{t+j}-\frac{\alpha R^{-1}}{(1+\alpha)}\sum_{j=0}^{\infty}R^{-j}y_{t+j+1}
\end{aligned}
$$

where we note that

$$
c_t=\mu+\left(1-R^{-1}\right)\left[A_t+\frac{1}{(1+\alpha)}\sum_{j=0}^{\infty}R^{-j}y_{t+j}\right]+\frac{\alpha}{(1+\alpha)}y_t \tag{7.12}
$$

$\mu=\left(\frac{(\beta\rho+\phi(1-\rho))R^{-1}}{(R^{-1}-1)(1+\alpha)}\right)=-\frac{\beta+\phi r}{(1+r)(1+\alpha)r}$ is a negative constant. The terms in the bracket are present in the standard (Flavin) definition of permanent income. However, with a negative constant and given that the terms that represent future earnings, $\sum_{j=1}^{\infty}R^{-j}y_{t+j}$, are multiplied by the fraction $\frac{1}{(1+\alpha)}<1$ which is a function of the quadratic adjustment costs of changing assets we have that the level of consumption in our framework is lower than the standard level of consumption (i.e. the Flavin-type definition of permanent income). Note that the constant represents the 'new' linear components to the liquidity constraint problem.

7.2.4 The Problem with Expectations

First, to accommodate the expectation operator we modify the formula given above. Our problem is

$$
V=\sum_{t=0}^{\infty}E_t b^t\left[a(L)z_t\right]\left[d(L)x_t\right]
$$

The formula from the previous section still applies, hence we have

$$
\frac{\partial V}{\partial(E_t x_t)}=b^t E_t d\left(bL^{-1}\right)a(L)z_t \quad t=0,1,2,\ldots \tag{7.13}
$$

We define the B operator for future purposes as $B^{-j}E_{t-1}x_t = E_{t-1}x_{t+j}$ for integer j[12]. The reason why this operator is used later in our problem instead of the lagged operator is that the latter does change the information set by changing it in time at the same time that the variable whose expectations are being formed is changed in time (i.e. the B operator does not change the nature of the information set but changes the variable through time). Applying (7.13) to the problem yields the following Euler equation

$$\begin{aligned} 0 &= E_{t+i}(b - y_{t+i})\left(R^{-1}\rho^{-1}L - 1\right) - \beta\rho - \phi(1-\rho) \\ &\quad + E_{t+i}\left(R^{-1}L^{-1} - 1\right)\left(R^{-1}\rho^{-1}L - 1\right)A_{t+i} + E_{t+i}\alpha(1-L)\left(1-\rho L^{-1}\right)A_{t+i} \end{aligned} \tag{7.14}$$

The transversality condition is

$$\lim_{T\to\infty} \rho^T E_t\left\{(y_{t+T} - y_{t+T-1}) + (\alpha+2)A_{t+T} - (1-\alpha)A_{t+T-1} - (\beta+\phi)\right\}A_{t+T} = 0$$

which is obtained in a similar manner to the problem without expectations.

The method for solving this equation is somewhat similar to the method that we used in the previous section. We use operator techniques to solve for equation (7.14) (Wallis (1980), Shiller (1978), Sargent (1979, 1987); here we follow Whiteman (1983) pp. 61 and Whiteman (1987) pp. 205). The solution technique is implemented by leading the Euler equation (7.14) and taking the conditional expectations as of time $t+i$ to get

$$\begin{aligned} 0 &= E_{t+i}\left\{E_{t+i}\left(R^{-1}\rho^{-1} - L^{-1}\right)(b - y_{t+i})\right\} \\ &\quad + E_{t+i}\left\{E_{t+i}\left(R^{-1}L^{-1} - 1\right)\left(R^{-1}\rho^{-1} - L^{-1}\right)E_{t+i}A_{t+i}\right\} \\ &\quad + E_{t+i}\left\{E_{t+i}\alpha\left(L^{-1} - 1\right)\left(1 - \rho L^{-1}\right)A_{t+i}\right\} - \beta\rho - \phi(1-\rho) \quad i = 0,1,2,\ldots \end{aligned}$$

[12] 'It is necessary to distinguish two operators B and L.The operator B is defined by

$$B^{-1}\left[Ex_{t+j} \mid \Omega_{t-1}\right] = Ex_{t+j+1} \mid \Omega_{t-1}$$

i.e., application of B^{-1} shifts forward by one period the date on the variables whose conditional forecast is being computed, but leaves the information set unaltered.[...]
[The lag operator is defined in such a way] that

$$L^{-1}\left[Ex_{t+j} \mid \Omega_{t-1}\right] = Ex_{t+j+1} \mid \Omega_t$$

so that application of L^{-1} shifts both the random variable x *and* the information set Ω forward by one period.' Sargent (1987) pp. 395, footnote 2.

By the Law of Iterated Expectations we can re-write the last expression as

$$\begin{aligned} 0 &= \left(R^{-1}\rho^{-1} - B^{-1}\right)\left(b - E_{t+i}y_{t+i}\right) - \beta\rho - \phi\left(1-\rho\right) \\ &\quad + \left(R^{-1}B^{-1} - 1\right)\left(R^{-1}\rho^{-1} - B^{-1}\right)E_{t+i}A_{t+i} + \alpha\left(B^{-1} - 1\right)\left(1 - \rho B^{-1}\right)E_{t+i}A_{t+i} \end{aligned} \tag{7.15}$$

We continue to assume that $R\rho = 1$. Re-express the Euler equation as

$$(1+\alpha)\left(B^{-1}-1\right)\left(R^{-1}B^{-1}-1\right)E_{t+i}A_{t+i} = \left(B^{-1}-1\right)E_{t+i}y_{t+i} - \beta\rho - \phi\left(1-\rho\right) \tag{7.16}$$

The zeros of the characteristic polynomial in A_t are at $B = R^{-1} < 1$ and $B = 1$. To prevent assets from permanently accumulating or decreasing, the polynomial $\left(R^{-1}B^{-1} - 1\right)$ is solved forwards. Divide the last equation by $\left(R^{-1}B^{-1} - 1\right)$[13] and after re-arranging we get

$$(1+\alpha)E_{t+i}A_{t+i+1} = (1+\alpha)E_{t+i}A_{t+i} + \frac{\left(B^{-1}-1\right)}{\left(R^{-1}B^{-1}-1\right)}E_{t+i}y_{t+i} - \frac{\beta\rho + \phi\left(1-\rho^{-1}\right)}{\left(R^{-1}-1\right)}$$

We now substitute the budget constraint into this last expression to obtain

$$\begin{aligned} (1+\alpha)RE_{t+i}c_{t+i} &= (1+\alpha)RE_{t+i}\left(A_{t+i} + y_{t+i}\right) - (1+\alpha)E_{t+i}A_{t+i} \\ &\quad - \frac{\left(B^{-1}-1\right)E_{t+i}y_{t+i}}{\left(R^{-1}B^{-1}-1\right)} + \left(\frac{\beta\rho + \phi\left(1-\rho^{-1}\right)}{\left(R^{-1}-1\right)}\right) \end{aligned}$$

We expand the terms related to y_t

$$\begin{aligned} E_{t+i}c_{t+i} &= \left(1-R^{-1}\right)E_{t+i}A_{t+i} + \frac{\left(R^{-1}B^{-1}-1\right)\alpha - \left(1-R^{-1}\right)}{\left(R^{-1}B^{-1}-1\right)(1+\alpha)}E_{t+i}y_{t+i} \\ &\quad + \left(\frac{\left(\beta\rho + \phi\left(1-\rho^{-1}\right)\right)R^{-1}}{\left(R^{-1}-1\right)(1+\alpha)}\right) \end{aligned}$$

[13] There is an important proviso about the B operator. Sargent warns; 'We must be careful here because the properties of B make the forward inverse of $(R^{-1}B^{-1} - 1)$ the only legitimate one, apart from the reasons of convergence. Operating on both sides of an equation with polynomials in nonpositive powers of B is legitimate. But it is *not legitimate* to operate with polynomials in positive powers of B. For example, $E_t x_{t+1} = E_t y_{t+1}$does *not* imply that $BE_t x_{t+1} = BE_t y_{t+1}$, i.e., $x_t = y_t$. The operation in the text is legitimate because it involves operating only with polynomials in nonpositive powers of B.' Sargent (1987), pp. 395.

and re-express everything as

$$E_{t+i}c_{t+i} = \mu + \left(1 - R^{-1}\right) E_{t+i}\left[A_{t+i} + \frac{1}{(1+\alpha)}\sum_{j=0}^{\infty} R^{-j} y_{t+i+j}\right] + \frac{\alpha}{(1+\alpha)} E_{t+i} y_{t+i} \qquad (7.17)$$

where μ is defined as before. The interpretation is once more equivalent to the one assumed for the nonstochastic equation. Again, it is the quadratic terms that play the biggest role in this framework, the linear liquidity constraints also affecting the level of consumption which is less than the permanent income-Flavin framework.

7.2.5 What Should the Change in Consumption be Equal to?

No Expectations

From the Euler equation the innovation in consumption ought to be

$$\begin{aligned} 0 &= \left(R^{-1}\rho^{-1}L - 1\right)\left\{(b - y_t) + \left(R^{-1}L^{-1} - 1\right)A_t\right\} \\ &\quad + \alpha(1-L)\left(1-\rho L^{-1}\right)A_t - \beta\rho L^{-1} - \phi(1-\rho L) \end{aligned}$$

or

$$c_t - R^{-1}\rho^{-1}c_{t-1} = -\alpha(1-L)\left(1-\rho L^{-1}\right)A_t - \left(R^{-1}\rho^{-1} - 1\right)b + \beta\rho + \phi(1-\rho) \qquad (7.18)$$

and assuming that $R\rho = 1$

$$\Delta c_t = \lambda - \alpha\left(1 - \rho L^{-1}\right)\Delta A_t$$

where $\lambda = \beta\rho + \phi\left(1-\rho^{-1}\right)$, one obtains the change in consumption. All the linear adjustment cost terms play a role in the change in consumption. The quadratic terms continue to exert an influence, this time the level of assets up to two periods can affect the change in consumption. This suggests that accumulation of assets becomes very important when making consumption decisions.

Expectations

To obtain the first difference in consumption, first use the evolution of assets equation to substitute A_t in (7.17)

$$\begin{aligned} c_{t+i} &= \mu + r\left[A_{t+i-1} + y_{t+i-1} - c_{t+i-1}\right] + \frac{1}{(1+\alpha)}\sum_{j=0}^{\infty} E_{t+i}R^{-j+1}y_{t+i+j} \\ &+ \frac{\alpha}{(1+\alpha)}E_{t+i}y_{t+i} \end{aligned} \tag{7.19}$$

Second, multiply (7.17) by $(1+r)$ and lag it one period,

$$\begin{aligned} (1+r)\,c_{t+i-1} &= rA_{t+i-1} + \frac{r}{(1+\alpha)}y_{t+i-1} + \frac{r}{(1+\alpha)}\sum_{j=0}^{\infty} E_{t+i-1}R^{-j+1}y_{t+i+j} \\ & \frac{(1+r)\,\alpha}{(1+\alpha)}E_{t+i-1}y_{t+i-1} + (1+r)\,\mu \end{aligned} \tag{7.20}$$

Subtract (7.20) from (7.19) to obtain,

$$\Delta c_{t+i} = \gamma + \frac{1}{(1+\alpha)}\sum_{j=0}^{\infty} \Delta E_{t+i}R^{-j+1}y_{t+i+j} + \frac{\alpha}{1+\alpha}\Delta y_{t+i} \tag{7.21}$$

or

$$\Delta c_{t+i} = \gamma + \frac{\alpha}{1+\alpha}\Delta y_{t+i} + \frac{1}{(1+\alpha)}\Delta y^p_{t+i}$$

where $\gamma = -r\mu$, and Δy^p_{t+i} denotes the standard definition of the innovation in permanent income given in the literature. The change in consumption is different from the permanent income case. First, we note that excess sensitivity to current income exists (the last term in (7.21)). Also, there is a drift term component. Note the following implication of this model; define $\frac{\alpha}{1+\alpha} = \xi$, where ξ is the excess sensitivity coefficient as defined by Flavin (1993), then

$$\Delta c_{t+i} = \gamma + \xi \Delta y_{t+i} + (1-\xi)\,\Delta y^p_{t+i}$$

which is Flavin's excess sensitivity equation with a drift term. Note that the drift term is a function of the linear and quadratic adjustments to asset changes and not a scaled function of the drift in the labour income process as Clarida (1991) showed under finite lifetimes. Note also that since μ is negative the drift term is positive in this case. We note further that the change in consumption does not follow an autoregressive process as Attfield et al.'s PIH1 model did. Even though their problem and ours appears to be similar as a result of the quadratic adjustment cost, the resulting dynamics are totally different even though for both problems the quadratic term ends up dominating most of the behaviour in consumption.

An alternative form of obtaining the innovation in consumption is to examine the stochastic Euler Equation (7.15) at time t,

$$\begin{aligned} 0 &= \left(R^{-1}\rho^{-1} - B^{-1}\right)\left(b - E_{t+i}y_{t+i}\right) - \beta\rho - \phi\left(1-\rho\right) \\ &\quad + \left(R^{-1}B^{-1} - 1\right)\left(R^{-1}\rho^{-1} - B^{-1}\right)E_{t+i}A_{t+i} + \alpha\left(B^{-1} - 1\right)\left(1 - \rho B^{-1}\right)E_{t+i}A_{t+i} \end{aligned}$$

Given the budget constraint, we can write after taking expectations and expanding the asset terms

$$\begin{aligned} E_{t+i}\left(R^{-1}\rho^{-1} - B^{-1}\right)\left(c_{t+i} - b\right) &= E_{t+i}\alpha\left[B^{-1} - \rho B^{-2} - 1 + \rho B^{-1}\right]A_{t+i} \\ &\quad -\beta\rho - \phi\left(1-\rho\right) \end{aligned}$$

Given $A_{t+i+1} - A_{t+i} = (1+r)\,s_{t+i}$, substitute this into the last expression to get

$$\begin{aligned} E_{t+i}\left(R^{-1}\rho^{-1} - B^{-1}\right)\left(c_{t+i} - b\right) &= \alpha E_{t+i}\left(1+r\right)s_{t+i} - \alpha\rho E_{t+i}\left(1+r\right)s_{t+i+1} \\ &\quad -\beta\rho - \phi\left(1-\rho\right) \end{aligned}$$

Assume that consumers are not impatient $R^{-1}\rho^{-1} = 1$

$$E_{t+i}\Delta c_{t+i+1} = \alpha E_{t+i}s_{t+i+1} - \alpha E_{t+i}\left(1+r\right)s_{t+i} + \beta\rho + \phi\left(1-\rho\right)$$

If $\alpha = \beta = \phi = 0$ (there are no constraints to asset management) we would have the standard Euler equation for the permanent income hypothesis. Again, assets play a role in the change in consumption, up to two periods and the linear terms continue to exert some influence on the first difference of consumption. Note that this equation is similar to an 'error correction' specification: define savings as $s_{t+i} = y_{t+i}^d - c_{t+i}$ where y_{t+i}^d is disposable income

$$\begin{aligned} E_{t+i}\Delta c_{t+i+1} &= \alpha E_{t+i}\left(y_{t+i+1}^d - c_{t+i+1}\right) - \alpha E_{t+i}(1+r)\left(y_{t+i}^d - c_{t+i}\right) \\ &\quad +\beta\rho + \phi(1-\rho) \end{aligned}$$

collect all the consumption terms and write

$$E_{t+i}\Delta c_{t+i+1} = \kappa + \frac{\alpha}{1+\alpha}E_{t+i}\left(y_{t+i+1}^d - y_{t+i}^d\right) - \frac{\alpha}{1+\alpha}rE_{t+i}\left(y_{t+i}^d - c_{t+i}\right)$$

which is an expectational error correction specification where $\kappa = \frac{\beta\rho+\phi(1-\rho)}{1+\alpha}$.

7.2.6 Savings and Superior Information

The implications for savings in this model are similar to Flavin's. To see this, substitute equation (7.17) into the definition of savings

$$\begin{aligned} s_t &= \frac{r}{1+r}A_t + y_t - \mu - \left(1 - R^{-1}\right)A_t - \frac{\alpha}{1+\alpha}y_t \\ &\quad - \left(1 - R^{-1}\right)\frac{1}{1+\alpha}\sum_{j=0}^{\infty} R^{-j}E_t y_{t+j} \end{aligned}$$

or

$$\begin{aligned} s_t &= -\mu + \frac{1}{1+\alpha}\left[y_t - r\sum_{j=0}^{\infty} R^{-j-1}E_t y_{t+j}\right] \\ &= -\mu + (1-\xi)\left[y_t - r\sum_{j=0}^{\infty} R^{-j-1}E_t y_{t+j}\right] \end{aligned}$$

where ξ denotes the excess sensitivity coefficient once more. Hence our savings equation is equivalent to Flavin's definition of savings minus the constant term. Since the constant term is negative, savings are therefore higher than in Flavin's case. The constant term may represent buffer stock behaviour; in the face of uncertain income, agents will decide to save a constant amount μ to buffer themselves against uncertainty. The income terms continue to exert a role in savings; if individuals expect their future income to increase they will dissave although the amount of dissaving will be lower than in the standard permanent income case since they are scaled down by $\left(\frac{1}{1+\alpha}\right)$ (the 'excess sensitivity' coefficient). Thus in this case we have managed to introduce a 'precautionary motive' for savings without resorting to marginal utility functions that are convex or assuming impatient consumers[14]. Finally, provided the econometrician knows what μ is (and thus what α is), then savings will provide information about the expected future levels of labour income.

7.3 Do the Linear Cost Terms Matter? A Comparison of the Estimates of the Consumption Equation with Respect to Previous Results

Our equations in the text suggest that we can calculate a variant of the Flavin excess sensitivity specification adding a constant to it. In our first empirical chapter -chapter 3 where we replicated and extended Campbell (1987), Campbell and Deaton (1989) and Flavin's (1993) results - we observed that the excess sensitivity hypothesis failed for a VAR framework for both US and UK data. Such a framework (as were Flavin's original calculations) was an unrestricted version of Flavin's in the sense that it already included a constant in the estimation. Thus we can conclude from those type of tests that the equations developed in this chapter cannot by themselves explain US and UK data on consumption. However, we can extend the equations formulated in the previous section and combine them with the models of imperfect information, lagged information and partial adjustment as we saw in chapter 4. The first equation to be estimated should be the specification that introduces imperfect information to the excess sensitivity hypothesis with a constant (this specification statistically encompasses the next two equations)

$$A(L)\,\Delta c_t = \lambda_1 + \xi A(L)\,\Delta y_t + (1-\xi)A(\delta)\,\phi(L)\,\varepsilon_t \tag{7.22}$$

[14]Thus, the effect of introducing soft constraints on our problem is consistent with Carroll and Kimball's (1999) results. These results are also consistent with those of Mariger and Zeldes.

The equation that introduces lagged information is

$$\begin{aligned}\Delta c_t &= \lambda_2 + \xi \Delta y_t + (1-\xi)\left[\phi(\delta) + \theta(\delta)(1-\omega)\right]\varepsilon_t \\ &\quad + (1+r)(1-\xi)\left[\phi(\delta) - \theta(\delta)(1-\omega)\right]\varepsilon_{t-1}\end{aligned} \tag{7.23}$$

and the last equation that we consider results from introducing partial adjustment to the excess sensitivity hypothesis with a constant,

$$\begin{aligned}\Delta c_t &= \lambda_3 + \gamma_0\psi_0\varepsilon_t - \xi(\psi_1+\psi_2)\varepsilon_t \\ &+ [\gamma_1\zeta_0 - \gamma_0]\psi_0\varepsilon_{t-1} + (1+r\gamma_1)r\xi(\psi_1+\psi_2)\varepsilon_{t-1} - \xi\psi_3\varepsilon_{t-1} \\ &+ [\gamma_2\zeta_1 - \gamma_1]\zeta_0\psi_0\varepsilon_{t-2} + [\gamma_2\zeta_1 - \gamma_1]r\xi(\psi_1+\psi_2)\varepsilon_{t-2} + (1+r\gamma_2)\xi\psi_3\varepsilon_{t-2} \\ &+ [\gamma_3\zeta_2 - \gamma_2]\zeta_1\zeta_0\psi_0\varepsilon_{t-3} + [\gamma_3\zeta_2 - \gamma_2]r\xi(\psi_1+\psi_2)\varepsilon_{t-3} + [\gamma_3\zeta_2 - \gamma_2]r\xi\psi_3\varepsilon_{t-3} \\ &+ \cdots + [\gamma_{n-1}\zeta_{n-2} - \gamma_{n-2}]\left[\prod_0^{n-3}\zeta_i\psi_0 + \prod_1^{n-3}\zeta_i r\xi(\psi_1+\psi_2) + \prod_2^{n-3}\zeta_i r\xi\psi_3\right]\varepsilon_{t-n+1} \\ &+ [\gamma_{n-1} - \zeta_{n-1}]\left[\prod_0^{n-2}\zeta_i\psi_0 + \prod_1^{n-2}\zeta_i r\xi(\psi_1+\psi_2) + \prod_2^{n-2}\zeta_i r\xi\psi_3\right]\varepsilon_{t-n}\end{aligned} \tag{7.24}$$

The restrictions to be applied to equation (7.22) are the same as those mentioned in chapter 5. To estimate these equations we use the same econometric techniques used before and thus refer the reader to chapter 5 for details. The best fitting equations for the US and the UK are reported in table 1[15]. The terminology is the same as in chapter 5 but we add the expression '$\lambda =$ ' which reports the $t-ratio$ for the estimated constant in equation (7.22)[16].

Table 1 shows that the number of lags in the consumption and income variables and, the number of error terms for equation (7.22) are all very similar to our previous calculations which were made without a constant. The constant terms are only significant in 50% of cases. Apart from nondurable consumption in the US - the equation favours the mixed model of partial adjustment and excess sensitivity with a constant[17] - all of the restrictions that can be imposed to (7.22) are rejected. This result differs from the results in chapter 5 because this is the first time that a restriction about the insignificance of the autoregressive terms for consumption

[15] None of these equations are restricted in anyway.

[16] This will tell us whether the linear cost terms are significant.

[17] The constant is statistically significant in this model. The t-ratio is 8.137.

Table 7.1: Regression Results for Hybrid Specifications with a Constant

US	Total	Nondurable OLS
eqn (7.22)	1 lag in Δy_t, 1 lag in Δc_t, 2 lags in v_t, $L_1 = -629.865$ $aic = 1578.729$ $\lambda = 2.292$ Ser Corr $\chi^2(15) = 6.39$ Normality $\chi^2(2) = 22.79$	1 lag in Δy_t, 1 lag in Δc_t, 2 lags in v_t, $L_1 = -583.215$ $aic = 1487.429$ $\lambda = 0.899$ Ser Corr $\chi^2(15) = 7.85$ Normality $\chi^2(2) = 8.76$
eqn (7.23)	0 lags in Δy_t, 1 lag in v_t, $L_2 = -637.295$ Ser Corr $\chi^2(15) = 23.84$ Normality $\chi^2(2) = 12.81$ Restriction: LR Test $\chi^2(3) = 14.86$	0 lags in Δy_t, 1 lag in v_t, $L_2 = -594.994$ Ser Corr $\chi^2(15) = 36.12$ Normality $\chi^2(2) = 10.37$ Restriction: LR Test $\chi^2(3) = 23.558$
eqn (7.24)	$v_{t-j} = 2$, $L_3 = -638.591$ Ser Corr $\chi^2(15) = 24.02$ Normality $\chi^2(2) = 22.34$ Restriction: LR Test $\chi^2(3) = 18.592$	$v_{t-j} = 3$, $L_3 = -585.563$ Ser Corr $\chi^2(15) = 12.79$ Normality $\chi^2(2) = 10.73$ Restriction: LR Test $\chi^2(3) = 4.696$
UK		
eqn (7.22)	1 lag in Δy_t, 1 lag in Δc_t, 3 lags in v_t, $L_1 = -490.516$ $aic = 1240.850$ $\lambda = 0.893$ Ser Corr $\chi^2(15) = 8.24$ Normality $\chi^2(2) = 10.02$	1 lag in Δy_t, 1 lag in Δc_t, 2 lags in v_t, $L_1 = -481.981$ $aic = 1221.780$ $\lambda = 1.932$ Ser Corr $\chi^2(15) = 9.67$ Normality $\chi^2(2) = 1.83$
eqn (7.23)	0 lags in Δy_t, 1 lags in v_t, $L_2 = -500.314$ Ser Corr $\chi^2(15) = 27.15$ Normality $\chi^2(2) = 22.52$ Restriction: LR Test $\chi^2(4) = 19.596$	0 lags in Δy_t, 1 lags in v_t, $L_2 = -489.233$ Ser Corr $\chi^2(15) = 24.82$ Normality $\chi^2(2) = 1.94$ Restriction: LR Test $\chi^2(3) = 14.504$
eqn (7.24)	3 lags in v_t, $L_3 = -498.407$ Ser Corr $\chi^2(15) = 11.2$ Normality $\chi^2(2) = 24.69$ Restriction: LR Test $\chi^2(3) = 15.782$	2 lags in v_t, $L_3 = -487.002$ Ser Corr $\chi^2(15) = 11.66$ Normality $\chi^2(2) = 2.04$ Restriction: LR Test $\chi^2(3) = 10.042$

cannot be rejected. For both measures of consumption in the UK and total consumption in the US, however, none of the restrictions are met and so we can suggest for these three cases that the model of imperfect information is superior to the model of lagged information. As a final test for the validity of equation (7.22) we test the restriction that the coefficients of the income terms are the same as the lagged consumption terms scaled by the excess sensitivity coefficient. It should be remembered, however, that this is a weak test for the equation. The results of this test are reported in table 2 where the size of the excess sensitivity coefficient is also reported.

Table 7.2: Tests onWeak Restrictions implied Equation (7.22)

US	Total Consumption	NonDurable OLS
LR	$\chi^2(1)$ =1.071	$\chi^2(1)$ =2.696
ALR	$\chi^2(1)$ =1.025	$\chi^2(1)$ =2.561
ξ(s.e.)	0.162	0.049
UK		
LR	$\chi^2(1)$ =3.964	$\chi^2(1)$ =2.918
ALR	$\chi^2(1)$ =3.738	$\chi^2(1)$ =2.773
ξ(s.e.)	0.3987	0.3881

The most significant result from table 2 is that all the Barlett adjusted Likelihood ratio tests suggest that the weak implications of equation (7.22) cannot be rejected at the 5% significance level. For the UK, the excess sensitivity coefficients are significant and have the right sign and size. For the US however, these coefficients are still not significant. It therefore appears that our modified excess sensitivity/imperfect information framework is able to capture some of the characteristics of US and specially UK consumption.

7.4 Conclusion

In this chapter we have examined a problem where we have made it difficult for consumers to hold negative assets for a prolonged period. One could interpret this as a liquidity constraint problem in the sense that the consumer is not free to borrow at all times. The problem also introduces some habit behaviour through the introduction of quadratic adjustment costs terms which include lagged terms. We have maintained most of the assumptions of REPI building on a quadratic type utility function and the assumption that the rate of interest is equal to the rate of time preference. We have done this because we want to ignore the implications for prudent behaviour associated with non-linear marginal utility functions and impatient consumers when these face asset constraints.

We find an analytical solution to the problem using operator techniques. The problem is interesting in that we are able to factorise the polynomials in the operators as functions of the adjustment costs parameters, thus enabling us to understand their full impact on the level of consumption, its innovation and the level of savings. Our solution is close to the excess sensitivity hypothesis formulated by Flavin (1993), although the theory advocates the inclusion of a drift term in consumption. Consumption is always less than in the rational expectations permanent income hypothesis case. Savings are higher and so is the level of assets. We are also able to re-formulate the innovation in consumption as a single equation error correction.

Our estimation results continue to suggest that the model of imperfect information appears to be superior to Goodfriend's model of lagged information. We also show that the constant term which represents the linear adjustment terms is significant in most cases. Our results do not vary much from the conclusions of chapter 5 although this time it is more difficult to reject the equation of imperfect information and excess sensitivity for the UK and for total consumption in the US.

7.5 Appendix 1: Solution to the Euler Equation without Restrictions (No expectations case)

To solve the Euler equation (7.8) in the text expand the terms associated with assets (after multiplying everything by L^{-1}) to obtain a second order difference equation,

$$\begin{aligned} &\left[1-\frac{R^{-2}\rho^{-1}+1+\alpha+\alpha\rho}{\alpha\rho+R^{-1}}L+\frac{R^{-1}\rho^{-1}+\alpha}{\alpha\rho+R^{-1}}L^2\right]A_{t+1} \\ =\ &(b-y_t)\left(R^{-1}\rho^{-1}L-1\right)-\beta\rho-\phi(1-\rho) \end{aligned}$$

To obtain a solution factor the polynomial in the square brackets as,

$$\begin{aligned} &(1-\lambda_1 L)(1-\lambda_2 L)A_{t+1} \\ =\ &(b-y_t)\left(R^{-1}\rho^{-1}L-1\right)-\beta\rho-\phi(1-\rho) \end{aligned}$$

where it is assumed that λ_1 will be the smaller root and λ_2 the higher root. We seek the following factorization,

$$\begin{aligned} \left[1+AL+BL^2\right] &= (1-\lambda_1 L)(1-\lambda_2 L) \\ &= 1-(\lambda_1+\lambda_2)L+(\lambda_1\lambda_2)L^2 \end{aligned}$$

where

$$-A=\frac{R^{-2}\rho^{-1}+1+\alpha+\alpha\rho}{\alpha\rho+R^{-1}}$$

and

$$B=\frac{R^{-1}\rho^{-1}+\alpha}{\alpha\rho+R^{-1}}=\frac{\alpha\rho+R^{-1}}{(\alpha\rho+R^{-1})\rho}=\frac{1}{\rho}>1$$

Equating powers in L gives,

$$-A=\lambda_1+\lambda_2,\ B=\lambda_1\lambda_2$$

Thus λ_1 must satisfy,

$$-A = \lambda_1 + \frac{B}{\lambda_1} = Y$$

We now examine the values for the two roots, that will satisfy our factorization. The minimum value of Y is given by $\frac{dY}{d\lambda_1} = 1 - B\lambda_1^{-2} = 0$ or $\lambda_1 = \sqrt{B} = \sqrt{\frac{1}{\rho}} = \sqrt{1+\delta} > 1$. At this value of λ_1, Y is equal to $2\sqrt{B} = 2\sqrt{1+\delta}$. To have two distinct roots that will solve this factorization, we need $-A > 2\sqrt{B}$. Note further that if $\lambda_1 = 1, Y = 1 + B = 2 + \delta$ which is greater than the minimum value of λ_1 since $0 < \delta < 1$. We now focus on $-A$ to see whether two different roots exist. If $-A > 1 + B$ then one of the roots that will solve the factorization will be greater than one and the other will be less than one. What conditions make this possible?

$$\frac{R^{-2}\rho^{-1} + 1 + \alpha + \alpha\rho}{\alpha\rho + R^{-1}} > 1 + \frac{R^{-1}\rho^{-1} + \alpha}{\alpha\rho + R^{-1}}$$

or

$$R^{-2}\rho^{-1} + 1 + \alpha(1+\rho) > R^{-1}\left(1 + \rho^{-1}\right) + \alpha(1+\rho)$$

which means

$$R^{-1}\rho^{-1} + R > \left(1 + \rho^{-1}\right)$$

or

$$\begin{aligned} \frac{1+\delta}{1+r} + (1+r) &> 1 + (1+\delta) \\ (1+\delta)\left[1 - (1+r)\right] &> \left[1 - (1+r)\right](1+r) \end{aligned}$$

Obviously, if $\delta > r$ meaning that consumers are impatient, then this condition will be satisfied; $\lambda_2 > 1$ and $\lambda_1 < 1$. If $\delta = r$ then we will have two different roots, one of them will be equal to 1 and the other one will be less than one. We saw this case in the text. If $r > \delta$ then we cannot say much about the nature of the solution. Note that these conditions are not affected by the extra terms that we added to the problem, the conditions for convergence boil down to whether consumers are impatient or not.

Thus to solve the Euler equation, we operate on both sides of this equation with the 'forward' inverse of $(1-\lambda_2 L)$ (since this is the root that will be greater than one) to get,

$$\begin{aligned}(1-\lambda_1 L)\, A_{t+1} &= \frac{(\lambda_2 L)^{-1}\left(1-R^{-1}\rho^{-1}L\right)}{\left(1-\lambda_2^{-1}L^{-1}\right)} y_t \\ &+ \frac{(\lambda_2 L)^{-1}\left[b\left(R^{-1}\rho^{-1}L-1\right)-\beta\rho-\phi(1-\rho)\right]}{\left(1-\lambda_2^{-1}L^{-1}\right)}\end{aligned}$$

This expression ought to converge provided $\delta > r$. We re-write this equation as,

$$A_{t+1} = \lambda_1 A_t + \frac{1}{\lambda_2}\sum_{j=0}^{\infty}\left(\frac{1}{\lambda_2}\right)^j y_{t+j+1} - \frac{1}{\lambda_2 R\rho}\sum_{j=0}^{\infty}\left(\frac{1}{\lambda_2}\right)^j y_{t+j} + \kappa$$

where $\kappa = \frac{(\lambda_2 L)^{-1}\left[b\left(R^{-1}\rho^{-1}L-1\right)-\beta\rho-\phi(1-\rho)\right]}{\left(1-\lambda_2^{-1}L^{-1}\right)}$ and thus includes the linear and quadratic terms. Note that both roots are functions of the quadratic terms, so these terms continue to affect future levels of human capital. If we substitute this asset equation into the budget constraint and rearrange, we would obtain an expression for consumption in terms of assets and present and future levels of income. In the text we opted to impose the restriction $R\rho = 1$ because this leads to an expression for consumption that is not a function of the two unknown roots, λ_1 and λ_2 but is a function of the parameters of interest in our problem. This leads to unsatisfactory implications for the level of assets.

7.6 Appendix 2: Coefficients for Equations in Table 7.1[18]

7.6.1 Equation (7.22)

US Total Consumption

$$\begin{aligned}\Delta c_t &= \underset{(2.29)}{3.95} + \underset{(10.34)}{0.757}\Delta c_{t-1} + \underset{(1.66)}{0.227}\Delta y_t - \underset{(-1.35)}{0.184}\Delta y_{t-1} \\ &\quad - \underset{(-7.06)}{0.647}\varepsilon_{t-1} - \underset{(-4.19)}{0.378}\varepsilon_{t-2}\end{aligned}$$

US Non-durable Consumption

$$\begin{aligned}\Delta c_t &= \underset{(0.89)}{6.96} + \underset{(1.28)}{0.464}\Delta c_{t-1} + \underset{(1.17)}{0.109}\Delta y_t + \underset{(0.99)}{0.104}\Delta y_{t-1} \\ &\quad - \underset{(-0.52)}{0.125}\varepsilon_{t-1} - \underset{(0.52)}{0.144}\varepsilon_{t-2} + \underset{(1.45)}{0.312}\varepsilon_{t-3}\end{aligned}$$

UK Total Consumption

$$\begin{aligned}\Delta c_t &= \underset{(0.89)}{1.681} + \underset{(1.03)}{0.287}\Delta c_{t-1} + \underset{(2.25)}{0.359}\Delta y_t + \underset{(0.97)}{0.201}\Delta y_{t-1} \\ &\quad - \underset{(-1.45)}{0.391}\varepsilon_{t-1} + \underset{(2.99)}{0.098}\varepsilon_{t-2} + \underset{(2.99)}{0.305}\varepsilon_{t-3}\end{aligned}$$

UK Non-durable Consumption

$$\begin{aligned}\Delta c_t &= \underset{(1.93)}{4.207} - \underset{(-0.57)}{0.105}\Delta c_{t-1} + \underset{(2.37)}{0.377}\Delta y_t + \underset{(1.87)}{0.318}\Delta y_{t-1} \\ &\quad + \underset{(0.54)}{0.107}\varepsilon_{t-1} + \underset{(2.40)}{0.254}\varepsilon_{t-2}\end{aligned}$$

7.6.2 Equation (7.23)

US Total Consumption

$$\Delta c_t = \underset{(2.80)}{14.37} + \underset{(2.01)}{0.284}\Delta y_t + \underset{(1.951)}{0.166}\varepsilon_{t-1}$$

[18] Diagnostics are reported in table 7.1.

US Non-durable Consumption

$$\Delta c_t = \underset{(7.26)}{16.76} + \underset{(1.06)}{0.109}\Delta y_t + \underset{(3.81)}{0.309}\varepsilon_{t-1}$$

UK Total Consumption

$$\Delta c_t = \underset{(2.83)}{4.035} + \underset{(2.53)}{0.436}\Delta y_t - \underset{(-1.18)}{0.108}\varepsilon_{t-1}$$

UK Non-durable Consumption

$$\Delta c_t = \underset{(3.49)}{4.93} + \underset{(2.61)}{0.419}\Delta y_t + \underset{(0.13)}{0.012}\varepsilon_{t-1}$$

7.6.3 Equation (7.24)

US Total Consumption

$$\Delta c_t = \underset{(2.29)}{3.95} + \underset{(2.036)}{0.169}\varepsilon_{t-1} + \underset{(2.19)}{0.183}\varepsilon_{t-2}$$

US Non-durable Consumption

$$\Delta c_t = \underset{(0.89)}{6.96} + \underset{(3.91)}{0.311}\varepsilon_{t-1} - \underset{(-0.71)}{0.059}\varepsilon_{t-2} + \underset{(4.01)}{0.319}\varepsilon_{t-3}$$

UK Total Consumption

$$\Delta c_t = \underset{(0.89)}{1.681} - \underset{(-0.68)}{0.059}\varepsilon_{t-1} + \underset{(2.41)}{0.206}\varepsilon_{t-2} + \underset{(2.96)}{0.258}\varepsilon_{t-3}$$

UK Non-durable Consumption

$$\Delta c_t = \underset{(1.93)}{4.207} + \underset{(0.97)}{0.083}\varepsilon_{t-1} + \underset{(3.39)}{0.293}\varepsilon_{t-2}$$

Chapter 8

Conclusions

8.1 Summary of Results.

Motivation. 'Consumer expenditure accounts for between 50% and 70% of spending in most economies. Not surprisingly, the consumption function has been the most studied of the aggregate expenditure relationships and has been a key element of all the macroeconomic model building efforts since the seminal work of Klein and Goldberger (1955).' Muellbauer and Lattimore (1994) (pp. 292).

Modern consumption research changed in 1978 after the publication of two seminal papers which used two different approaches to explain consumption. Davidson, Hendry, Srba and Yeo (1978) favoured a fundamentally based econometric approach which could, in principle, explain the behaviour of consumption both in the short and the long run. They put forward an error correction model for consumption which they argued could not only achieve such aims but would also be able to produce a general consumption specification capable of encompassing previous research on consumption. This approach had a profound effect on the understanding of time series techniques (i.e. co-integration and stationarity). However, the principal shortcoming of this approach is that it does not pass Lucas's critique.

Hall (1978) based his investigation on the Euler equation approach which results from an intertemporal consumption decision made by a (rational) representative agent. Hall showed that under certain assumptions, consumption would follow a random walk. Hall's approach, termed the rational expectations permanent income hypothesis, has been the favoured one by economic theorists in the last twenty years; its appeal lies in its treatment of economic theory and uncertainty. However, empirical studies have found that consumption does not follow a random walk because it responds too much to current and lagged income and because it is less volatile than permanent income predictions. Theorists have since tried to explain both phenomena (known as the excess sensitivity and smoothness of consumption respectively) and most theoretical explanations for the failure of the random walk result have examined the assumptions made by Hall for solving the Euler equation.

The Euler equation approach and in particular the assumptions required to solve it are not free from criticism. Muellbauer and Lattimore (1994) (pp. 292) provide an example: 'There is a widespread belief that estimating an Euler equation corresponding to the optimization problem of a rational, infinitely lived representative agent operating in efficient financial markets without credit or transaction costs or restrictions is somehow more rigorous than trying to develop more comprehensive, realistic, but necessarily approximate models.' Deviations from Hall's assumptions make it difficult to solve the Euler equation; in fact there has recently been a new literature that has tried to solve the first order Euler equation with less assumptions using complex numerical simulation techniques. A drawback of this literature is that a solution for consumption is not always obtained and approximations are often the only thing we can hope for (see Deaton (1991), Carroll (1992, 1997a)). Nonetheless, this new literature has enhanced our understanding of the consumption problem considerably.

Using the generalized method of moments, Hansen and Singleton (1982, 1983) have shown that we do not have to resort to the assumptions made by Hall to obtain a consumption function that can be used to test for the truth of the Euler equation. However, such techniques are difficult to implement and have been largely disregarded in the literature.

The aim. The aim of this thesis was to enhance our understanding of consumption behaviour (in both the US and the UK) basing our theory on the rational expectations permanent income hypothesis. We paid particular attention to this hypothesis because we want to know whether rational agents attempt to smooth their consumption throughout their lifetimes. We neglected the recent literature on precautionary saving behaviour (Kimball (1990), Deaton (1991), Carroll (1992, 1997a), Kimball and Carroll (1996)) as we sought solutions with quadratic utility functions.

Chapter 2: Review of the literature. The modern literature on consumption was reviewed in our second chapter and in doing so, we identified possible areas for consumption research. We paid particular attention to the empirical studies that have tested the truth of the rational expectations permanent income hypothesis and to those theories that aim to explain why that theory fails at the aggregate level. Two main conclusions may be drawn from this exercise: first, that the main bulk of empirical research on aggregate consumption has been undertaken on *two data sets for the US and the UK economies which end in the mid 1980s.* Second, there have been many different explanations given for the failure of the random walk hypothesis. Most explanations involve different economic theories/concepts which in turn lead to differing consumption specifications. There is therefore *no common consensus about what the consumption function should look like.* We attempt to deal with the first issue in the third chapter of our thesis and because we find that the martingale prediction continues to fail on two more recent data sets, in the remainder of our thesis, we turn our attention to the second issue.

Chapter 3: New evidence about excess sensitivity and smoothness. In their review of the state of consumption research, Blinder and Deaton (1985) fitted a number of popular consumption equations to US time series data. Blinder and Deaton argued that the National Income and Product Accounts (NIPA) series on consumption, labour income and disposable income are not in an appropriate form to test theories of consumption accurately. They recommended a number of revisions to the series and in doing so constructed a data set that runs from 1953:2 to 1984:4. Attfield, Demery and Duck (1990) produced similar recommendations for ONS data in the UK and constructed a data set that runs from 1955:1 to 1987:2. These two data sets have been used extensively to test for the truth of the rational expectations permanent income hypothesis. It has generally been found that consumption in the US and the UK suffers from excess sensitivity and smoothness.

In chapter 3, we follow these recommendations to obtain two data sets for the US and the UK that run from the late 1950s to the mid 1990s. These data sets consist of four variables, total and non-durable consumption, and disposable and labour income. We construct these two data sets for two reasons; first, to understand how the economic conditions of the mid 1980s and the 1990s affected consumption behaviour, and second, to see whether previous empirical results on consumption are sensitive (and thus exclusive) to the data period used. We pay particular attention to the issues of excess sensitivity and smoothness.

In the chapter we also transform the rational expectations permanent income hypothesis framework into logarithmic form. This follows the suggestion made by Campbell and Deaton (1989), that it may be easier for logarithmic data to achieve stationarity after first differencing than differencing the data expressed just in levels. We test the random walk hypothesis for the US and the UK using both data in levels and in logarithms. We do this using the tests developed by Campbell (1987), Campbell and Deaton (1989) and Flavin (1993) which can in turn explain whether consumption suffers from excess sensitivity and smoothness. These tests have been acknowledged as the best way of testing for the truth of Hall's framework because they can overcome the problem of superior information. The problem of superior information arises because it is assumed that consumption decisions are principally influenced by the agent's expectations of his or her future labour income. Since econometricians or researchers cannot have access to the information set an agent uses to predict his or her future labour income, the predictions made by an econometrician are likely to be different to those made by the agent. Hence, it should not be possible to test the random walk hypothesis. The three tests mentioned above are able to overcome this problem because they use the agent's behaviour, represented in his or her savings and/or consumption, to obtain the exact information that is processed by the agent itself. If agents wish to smooth their consumption through time and the permanent income hypothesis is true, savings can tell exactly what the agent's expectations about his or her future labour income are: if at time t the agent receives information that his or her future labour income is to increase, the consumer will increase his or her consumption thereby reducing savings. The econometrician should be able to see this decrease in savings and thus predict that the consumer expects his or her future labour income to increase. Thus savings can help predict labour income if REPI is true.

To perform these tests we first need to establish the order of integration of all the series used in these tests. For this purpose we employ the tests suggested by Dickey and Fuller and Phillips and Perron. We find that these tests suggest that all the four series should be integrated of order one. We also examine the possibility that structural breaks may have occurred during the period using the tests put forward by Barnejee, Lumsdaine and Stock (1992). We find that there is little evidence for the case of structural breaks in the US and the UK and thus we conclude that all the series appear to be $I(1)$. We then test for the stationarity of savings as the residual of a co-integrating relationship between disposable income and consumption and we find that savings are indeed stationary.

Given these results we estimate a VAR equation for savings and the innovation in labour income (both $I(0)$ variables). We discuss the appropriate restrictions that have to be imposed to the VAR to test for the truth of the random walk hypothesis. We find what the appropriate lag length of the VAR ought to be through the Akaike information criteria and the Schwartz Bayesian criterion.

Our first test about the truth of the random walk hypothesis is a weak test and it involves testing for Granger causality from savings to labour income. If savings do not Granger cause innovations in labour income then REPI cannot hold. We find that savings do help to predict future labour income so there is evidence that consumers *seek to smooth their consumption*; although this is not sufficient evidence to conclude that the random walk hypothesis holds. For that purpose we impose the appropriate restrictions to the estimated $VARs$. Using the appropriate tests, we find that the rational expectations permanent income hypothesis *does not hold for both the US and the UK.* There is evidence that consumption suffers from excess sensitivity and smoothness in the US and only excess sensitivity in the UK. For the UK, consumption appears to be more volatile that the predictions made by REPI. For both countries we find that whilst the appropriate restrictions for the truth of the excess sensitivity hypothesis of Flavin do in fact hold, the theory does not make much economic sense because the excess sensitivity coefficient is insignificant and often negative in our data. This is a new result for consumption, Flavin was not able to reject her theory using the Blinder and Deaton data set[1].

Chapter 4: Sensitivity, adjustment and imperfect information; Theory. In our fourth chapter we turn our attention to some of the explanations given for the failure of the random walk hypothesis. We combine three recent extensions to the Hall framework to develop more general specifications for the consumption function which can be tested against the three extensions themselves and REPI. We test all these equations in chapter five. The three extensions can, theoretically speaking, explain the phenomena of excess sensitivity and smoothness and we

[1]In this chapter we also considered the *same periods* for both of our data sets compared to the original data sets of Blinder and Deaton and Attfield et al.. We ran all the tests again and found that none of the results are affected depending on which data set is used.

find that the equations which we develop in this chapter are also capable of explaining both phenomena. We obtain five consumption equations that have rich dynamics and differ slightly in their behaviour. One of those five specifications is a general model of three of the other four equations. Thus, imposing the appropriate restrictions, we can test the general model against the other three equations and we are left with another consumption equation that cannot be tested against any of the other equations. Both general specifications have Pischke's imperfect information as a characteristic of consumption behaviour.

Chapter 5: Sensitivity, adjustment and imperfect information; Empirical Evidence. In our fifth chapter we pick up these issues. We discuss the equations that will be estimated, the restrictions to be imposed and the econometric methods to be used. Since we have consumption equations where labour income features as an independent variable we require the use of instrumental methods. Like Campbell and Mankiw (1989, 1990), Demery and Duck (1997,1998) and others before we use instrumental variables dated at $t-2$ for our estimation. This is because it is assumed that consumption decisions take place continuously whereas our data is quarterly, so it is likely for the first difference in consumption and the changes in labour income lagged one period to be correlated. The instruments used for both the US and the UK are the same instruments that Demery and Duck used to test for the truth of the imperfect information hypothesis of Pischke (1995). To estimate the five equations developed in the fourth chapter, we use the Pagan (1974) estimation method because some of the equations have autocorrelated errors of an order greater than one.

The two most general equations were estimated without imposing any kind of restrictions. We began with the estimation of the most general equations which included to start with, ten lags in all the independent variables and lagged consumption. Those equations that did not exhibit any serial correlation were then chosen as possible explanations for the change in consumption. To obtain the specific order for all the lag terms amongst those remaining equations we performed Likelihood ratio tests for the joint significance of the estimated coefficients and we also looked at an Akaike Information type criteria. Having obtained the two most general specifications, we imposed the appropriate restrictions to the general equation that encompasses the other three.

We find that the *standard martingale specification for consumption can be rejected* in favour of one of the two general models although it is not possible to determine which of the two models is superior. We also find that the model of imperfect information of Pischke is superior to the model of lagged information of Goodfriend (1992). From an economic point of view, we find a common characteristic for all the estimated consumption equations: *agents are not quick to process new information and to adjust to it accordingly.* There is evidence from our data that adjustment to new information is an on-going process that may last forever. The exact reasons why this is so are not known however; whilst imperfect information does appear to be a strong candidate, it is likely that other factors can *also* account for our results. Liquidity constraints or the value an agent attaches to his or her free time above his or her time spent to obtain accurate forecasts about new information are two of the possible explanations our models suggest. It is not possible to determine which of the two is superior however.

It has been argued that the growth rates of many economic variables have decreased from 1973 onwards. We take this point in the second part of chapter 5 to revaluate all the tests performed in chapter 3 and the five equations estimated in the first section by considering only periods that start in 1973 for both the US and the UK.

The post 1973 period produces no significant differences to the results which have either been found in the literature or in earlier parts of our work although the data appears to be a little bit kinder to the martingale result[2]. Our tests continue to suggest that the excess sensitivity hypothesis still passes all the associated restrictions imposed to the VAR but that the excess sensitivity coefficient does not appear to be economically and statistically significant. Moreover, the concept of imperfect information continues to be an important explanation of consumption behaviour in both the US and the UK. As we reported above, we are still not able to statistically discriminate between the concepts of excess sensitivity and partial adjustment although the data appears to be more favourable to the excess sensitivity hypothesis than before.

Chapter 6: Imperfect Information and the Lifecycle Hypothesis. In chapter 6, we deviate from the assumption that agents live forever. We consider their retirement behaviour and how that behaviour affects their consumption decisions when new information about their labour

[2]The $VAR(1)$ system appears to support the permanent income hypothesis for a number of specific cases for the US and the UK.

income becomes available. Clarida (1991) first examined these issues using a framework that was similar to Flavin's (1981). Clarida argued that his model was capable of explaining the excess sensitivity and smoothness phenomena. However, when persistence is introduced to a simple income process, his model is not capable of explaining the excess smoothness puzzle. Armed with the results from the previous chapter (where we found that imperfect information appears to be a characteristic of consumption behaviour for both the US and the UK) we introduce the concept of imperfect information to Clarida's work to see the implications for excess smoothness and sensitivity. The focus of the chapter is on the first and second moment properties of changes in per capita consumption. We find that smooth per capita consumption in the presence of permanent shocks to per capita labour income is the outcome that one should expect. The marginal propensity to consume out of current income innovations is less in our model than in Clarida's and the importance of lagged innovations in labour income is increased. We also find that persistence still exerts an important role in 'Deaton's Paradox' but the strength of this Paradox is greatly diminished by the introduction of imperfect information.

Chapter 7: Borrowing restrictions and habits; theory and evidence. In this chapter we consider the possibility that agents may get into debt but that they cannot remain in that situation for prolonged periods of time. We do this to compare our results to the literature that has tried to examine the effects of liquidity constraints upon consumption decisions. Since that literature assumes that agents are not allowed to get into debt, in chapter 7 we modify the consumption problem to understand the implications for consumption behaviour if agents are allowed to have a certain amount of debt but can only sustain it for long periods if they incur high costs. The problem also introduces some habit behaviour in the sense that we assume that agents are slow to adjust their asset holdings. Slow adjustment arises in this model because we assume that it may be expensive to monitor the optimal level of assets at all times, or because agents may be lazy, or because there are physical costs associated with changing assets (like brokerage costs, etc.). We maintain most of the assumptions of REPI specially a quadratic utility type function and a constant rate of interest equal to the rate of time preference (no 'impatient' consumers). Using lag operator techniques, we find a solution to the consumption problem that is a more general specification than Flavin's excess sensitivity hypothesis. We are also able to re-express the equation as a single equation error correction mechanism. The principal difference with respect to Flavin's model is that our model advocates the inclusion of a drift term for the innovation in consumption. Since the model is similar to Flavin's, it is only a trivial extension to introduce the concepts of imperfect and lagged information and partial adjustment to the equation. This results in three testable equations.

In the second section of the chapter we test these three equations. We find that the general equation of imperfect information is better at explaining the behaviour of consumption although non-durable consumption in the US proved to be the exception to this conclusion. We also find that the model developed in the first section of this chapter which includes imperfect information is able to pass a number of tests and thus appears to be reasonably successful at explaining the behaviour of consumption in the UK (it does not fail any tests at the 5% significance level or better and the excess sensitivity coefficients are still significant) and the US (it does not fail any test at the 10% significance level or better but the excess sensitivity coefficients are not significant). Moreover, the drift term that distinguishes our model from Flavin's appears to be significant in most cases.

8.2 Evaluation and Future Research.

Two recent reviews of the consumption literature have suggested that the Euler approach to consumption modelling at the aggregate level is not necessarily the way forward (Attanasio (1998) and specially Muelbauer and Lattimore (1994)). From our research it seems that such a conclusion is not necessarily the correct one; our calculations suggest that there is evidence that agents do attempt to smooth their consumption. To drop *the Euler equation framework* would therefore be a mistake because our evidence suggested that we can continue to understand consumption behaviour with it. However, we do not suggest that research should only be conducted within the confines of a problem whereby agents allocate their lifetime wealth so as to maximize their lifetime welfare. There is no doubt that we can learn much about consumption with the *estimation of consumption equations using the latest econometric methodology*; Hendry (1992) advocates the use of the data set as the principal mechanism to understand behavioural relationships. This is also one of the recommendations made by Muellbauer and Lattimore. The principal shortcoming to this approach is that it does not necessarily pass the Lucas's critique.

We believe that the main difficulty with the kind of research that we have looked at is that there are perhaps *too many factors that can affect the consumption function* and it is therefore very complicated to isolate what factors are the main determinants of consumption behaviour. This is one of the conclusions that we made from the results obtained in chapters 5 and 7. A final point related to this issue is that all these factors can *change with time* thus making things even more complicated.

Aggregation problems are perhaps one the greatest barriers to obtaining an accurate description of aggregate consumption behaviour. These problems have been well documented in Attanasio's excellent review of the consumption literature and we refer the reader to his article for more details. An extension to this point and a means of understanding consumption behaviour would be to concentrate our efforts to the *micro data*, only then will we be ready to understand aggregation problems in consumption more accurately.

Numerical techniques have enhanced our understanding of the consumption problem. These techniques are extremely complicated and the pay-off is not always worthwhile. In this literature, precautionary behaviour has come to play an important role in explaining consumption behaviour. Work on precautionary savings has made economists aware that savings can be an important influence on consumption decisions and we believe that more research on savings behaviour is needed. Savings have played the supporting role in the consumption problem and have been interpreted as the residual to the difference between disposable income and consumption. Consumption decisions are likely to be determined by the amount of savings an individual has and also by the amount of disposable income that could be potentially earned. It seems therefore a worthwhile exercise to concentrate our efforts to *understanding labour supply and savings decisions* to potentially increase our understanding of consumption. Research should also consider the importance that *different institutional frameworks and financial instruments* may have for consumption, saving and labour supply decisions specially since the increased participation of agents in the stock market appears to be an important and new source for earning income. This is a point that Attanasio makes in his conclusion (pp. 52).

Another field that may be interesting to research is related to Muellbauer and Lattimore's aforementioned suggestion. Recent literature on time series has suggested that we have to test whether a time series is linear or non-linear. There have been recent efforts to determine whether consumption can be modelled more accurately as a linear or non-linear series and there appears to be some evidence that consumption may be best explained as a non-linear equation. Whilst this is mainly an econometric approach to modelling consumption, it can help us understand consumption behaviour; these results may be suggestive of precautionary behaviour, liquidity constraints, slow adjustment, etc.

Bibliography

[1] Arrow, K. J. (1965), 'Aspects of a Theory of Risk Bearing', Yrjo Jahnsson Lectures, Helsinki. Reprinted in *Essays in the Theory of Risk Bearing* (1971). Chicago: Markham Publishing Co.

[2] Attfield, C. L. F. (1995), 'A Barlett Adjustment to the Likelihood Ratio Test for a System of Equations', *Journal of Econometrics*, 66, pp. 207-23.

[3] -, Demery, D. and Duck, N. W. (1990), 'Saving and Rational Expectations: Evidence for the U.K.' *Economic Journal*, 100, pp. 1269-76.

[4] -, -, - (1991), *Rational Expectations in Macroeconomics: An Introduction to Theory and Evidence.* Second Edition, Oxford: Blackwell Publishers.

[5] - (1992), 'Partial Adjustment and the Permanent Income Hypothesis', *European Economic Review*, 36, pp. 1205-22.

[6] Attanasio, O. P. (1998), 'Consumption Demand', *NBER Working Paper No. 6466*, March 1998.

[7] Banerjee, A., Lumsdaine, R. L. and Stock, J. H. (1992), 'Recursive and Sequential Tests of the Unit-Root and Trend-Break Hypotheses: Theory and International Evidence', *Journal of Business & Economic Statistics*, 10, pp. 271-87.

[8] Ball, R. J. and Drake, B. S. (1964), 'The Relationship between Aggregate Consumption and Wealth'. *International Economic Review*, 18, pp. 63-81.

[9] -, Boatwright, B. D., Burns, T., Lobban, P. W. and Miller, G. W. (1975), 'The London Business School Quarterly Econometric Model of the UK Economy' Chapter 1 in Renton, G. A. (ed.) *Modelling the Economy.* London: Heinemann Educational Books.

[10] Bernanke, B. (1985), 'Adjustment Costs, Durables and Aggregate Consumption', *Journal of Monetary Economics*, 15, pp. 41-68.

[11] -, Bohn, H. and Reiss, P. C. (1988), 'Alternative Non-Nested Specification Tests of Time-Series Investment Models', *Journal of Econometrics*, 37, pp. 293-326.

[12] Blanchard, O. J. and Fischer, S. (1989), *Lectures on Macroeconomics.* MIT Press.

[13] Blinder, A. S. and Deaton, A. (1985), 'The Time Series Consumption Function Revisited', *Brookings Papers of Economic Activity*, 2, pp. 465-511.

[14] Brady, D. S. and Friedman, R. (1947), 'Savings and the Income Distribution' *In Studies in Income and Wealth, vol. 1. New York: National Bureau Economic Research*, pp. 247-65.

[15] Branson, W. H. and Klevorick, A. K. (1972), 'Money Illusion and the Aggregrate Consumption Function', *American Economic Review*, 59, pp. 832-50.

[16] Brown, T. M. (1952), 'Habit Persistence and Lags in Consumer Behaviour', *Econometrica*, 20, pp. 355-71.

[17] Browning, M. and Lusardi, A. (1996), 'Household Saving: Micro Theories and Macro Facts', *Journal of Economic Literature*, XXXIV, pp. 1797-1855.

[18] Bureau of Economic Analysis (1996), 'Improved Estimates of the National Income and Product Accounts for 1959-95: Results of the Comprehensive Revision', Survey of Current Business, 76, pp. 1-31.

[19] Caballero, R.J. (1990), 'Consumption Puzzles and Precautionary Saving', *Journal of Monetary Economics*, 25, pp. 113-36.

[20] Campbell, J. Y. (1987), 'Does Saving Anticipate Declining Labour Income? An Alternative Test of the Permanent Income Hypothesis', *Econometrica*, 55, pp. 1249-74.

[21] - and Deaton, A. (1989), 'Why is Consumption so Smooth?', *Review of Economic Studies*, 56, pp. 357-73.

[22] - and Mankiw, N. G. (1989), 'Consumption, Income Interest Rates: Reinterpreting the Time Series Evidence' in Blanchard, O. and Fischer, S. (eds), *NBER Macroeconomic Annual 1989*, Cambridge, Mass.: MIT Press, pp. 185-216.

[23] -, - (1991), 'Permanent Income, Current Income and Consumption', *Journal of Business and Economic Statistics*, 8, pp. 265-79.

[24] Carroll, C. D. (1992), 'The Buffer Stock Theory of Saving: Some Macroeconomic Evidence' *Brookings Papers on Economic Activity*, 2, pp. 61-156.

[25] - (1997a), 'Buffer Stock Saving and the Life Cycle/Permanent Income Hypothesis', *Quarterly Journal of Economics*, 112, pp. 1-55.

[26] - (1997b), 'Death to the Log-Linearized Consumption Euler Equation! (And Very Poor Health to the Second-Order Approximation', *NBER Working Paper* 6298.

[27] - (1997), 'Death to the log-linearised Euler equation', *Mimeo, The Johns Hopkins University.*

[28] - and Kimball, M. (1996), 'On the Concavity of the Consumption Function', *Econometrica*, 64, pp. 981-92.

[29] - (1999), 'When and Why Do Liquidity Constraints Induce Precautionary Saving?', *Mimeo, The Johns Hopkins University.*

[30] Charemza, W. W. and Deadman, D. F. (1992), *New Directions in Econometric Practice: General To Specific Modelling, Cointegration and Vector Autoregression.* Edward Elgar Publishing Limited.

[31] Clarida, R. H.(1991), 'Aggregate Stochastic Implications of the Life Cycle Hypothesis', *Quaterly Journal of Economics*, 106, pp. 851-67.

[32] Cochrane, J. H. (1989), 'The Sensitivity of Tests of the Intertemporal Allocation of Consumption to Near-Rational Alternatives', *American Economic Review*, 79, pp. 319-37.

[33] Darby, M. R. (1987), The Consumption Function. *In The New Palgrave Dictionary Of Economics*, J. Eatwell, M. Milgate and P. Newman (eds.) London: Macmillan.

[34] Davidson, J. E. H., Hendry, D. F., Srba, F. and Yeo, S. (1978), 'Econometric Modelling of the Aggregate Time-Series Relationship between Consumers' Expenditure and Income in the United Kingdom', *Economic Journal*, 88, pp. 661-92.

[35] Davidson, R. and MacKinnon, J. (1981), 'Several Tests for Model Specification in the Presence of Alternative Hypotheses', *Econometrica*, 49, pp. 781-93.

[36] Davis, T. E. (1952), 'The Consumption Function as a Tool for Prediction', *Review of Economic and Statistics*, 34, pp. 270-7.

[37] Davis, E. P. (1984), 'The Consumption Function in Macroeconomic Models: A Comparative Study, *Applied Economics*, 6, pp. 799-838.

[38] Deaton, A. S. (1977), 'Involuntary Saving Through Unanticipated Inflation', *American Economic Review*, 67, pp. 899-910.

[39] - (1987), 'Life-Cycle Models of Consumption: Is the Evidence Consistent with the Theory?', in Truman F. Bewley (ed.), *Advances in Econometrics, Fifth World Congress*, 2, Cambridge and New York. Cambridge University Press, pp. 121-48.

[40] - (1991), 'Saving and Liquidity Constraints', *Econometrica*, 59, pp. 1221-48.

[41] - (1992), *Understanding Consumption.* Oxford University Press.

[42] - and Laroque, G. (1992), 'On the Behaviour of Commodity Prices', *Review Of Economic Studies*, 59, pp. 1-23.

[43] Demery, D. and Duck, N. W. (1999), 'Incomple Information and Consumption in the UK and US', *Economica*, 66, pp. 375-87.

[44] -, - (2000), 'Incomplete Information and the Time Series Behaviour of Consumption', *Journal of Applied Economics,* 15, pp. 355-366.

[45] Dickey, D.A. and Fuller, W.A. (1979), 'Distribution of the Estimates for Autogressive Time Series with Unit Roots', *Journal of the American Statistical Association*, 74, pp. 427-31.

[46] Dornbusch, R. and Fischer, S. (1990), *Macroeconomics.* New York, McGraw-Hill.

[47] Dreze, J. and Modigliani, F. (1972), 'Consumption Decisions Under Uncertainty', *Journal of Economic Theory*, 5, pp. 308-55.

[48] Duesenberry, J. S. (1949), *Income, Saving and the Theory of Consumer Behaviour.* Cambridge, Mass: Harvard University Press.

[49] Enders, W. (1995), *Applied Econometric Time Series*, Wiley Series in Probability and Mathematical Statistics.

[50] Engle, R. F. and Granger, C. W. J. (1987), 'Dynamic Specification with Equilibrium Constraints: Cointegration and Error Correction', *Econometrica*, 55, pp. 251-76.

[51] Ferber, R. (1973), 'Consumer Economics, A Survey', *Journal of Economic Literature*, 11, pp. 1303-42.

[52] Fischer, I. (1930), *The Theory of Interest.* New York: Macmillan.

[53] Flavin, M. (1981), 'The Adjustment of Consumption to Changing Expectations About Future Income', *Journal of Political Economy*, 89, pp. 974-1009.

[54] -, (1985), 'Excess Sensitivity of Consumption to Current Income: Liquidity Constraints or Myopia?', *Canadian Journal of Economics*, 13, pp. 117-37

[55] - (1993), 'The Excess Smoothness of Consumption: Identification and Interpretation', *Review of Economic Studies*, 60, pp. 651-66.

[56] Friedman, M. (1957), *A Theory of the Consumption Function.* Princeton, New Jersey: Princeton University Press.

[57] Galí, J. (1991), 'Budget Constraints and Time Series Evidence on Consumption' *American Economic Review*, 81, pp. 1238-53.

[58] *Gauss Command Reference Manual*, Volume II.

[59] Goodfriend, M. (1992), 'Information-Aggregation Bias' *American Economic Review*, 82, pp. 508-19.

[60] Granger, C. W. J. and Newbold, P. (1974), 'Spurious Regressions in Econometrics', *Journal of Econometrics*, 2, pp. 111-20.

[61] Greene, W. (1993), *Econometric Analysis*, Second Edition, MacMillan, New York.

[62] Grossman, S. J. and Laroque, G. (1990), 'Asset Pricing and Optimal Portfolio Choice in the Presence of Illiquid Durable Consumption of Goods' *Econometrica*, 58, pp. 25-51.

[63] Haavelmo, T. (1947), 'Methods of Measuring the Marginal Propensity to Consume', *Journal of the American Statistical Association*, 42.

[64] Hadjimatheou, G. (1987), *Consumer Economics After Keynes*, Brighton: Wheatsheaf.

[65] Hall, R. E. (1978), 'Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence' *Journal of Political Economy*, 86, pp. 971-87.

[66] - (1989a), 'Intertemporal Substitution in Consumption', *Journal of Political Economy*, 96, pp. 339-57.

[67] - (1989b), 'Consumption', in Robert J. Barro (ed.) *Modern Business Cycle Theory*. Cambridge, Mass. Harvard University Press.

[68] - and Mishkin, S. (1982), 'The Sensitivity of Consumption to Transitoty Income: Estimates From Panel Data on Households', *Econometrica*, 50, pp. 461-81.

[69] Hamilton, J. D. (1994), *Time Series Analysis*, Princeton University Press.

[70] Hansen, L.P. and Sargent, T.J. (1981), 'A Note on Wiener-Kolmogorov Forecasting Formulas for Rational Expectations Models', *Economic Letters*, 8, pp. 253-60.

[71] - and Singleton, K.J. (1982), 'Generalized Instrumental Variables Estimation of Nonlinear rational Expectations Models', Econometrica, 50, pp. 1269-86.

[72] -, - (1983), 'Stochastic Consumption, Risk Aversion and the Temporal Behaviour of Asset Returns', *Journal of Political Economy*, 91, pp. 249-65.

[73] Harris, R. I. D. (1995), *Using Cointegration Analysis in Econometric Modelling*, Prentice Hall/Harvester Wheatsheaf.

[74] Harvey, A. C. (1990), *The Econometric Analysis of Time Series, 2nd Edition*, New York: Phillip Allan.

[75] Heaton, J. (1993), 'The Interaction Between Time-Nonseparable Preferences and Time Aggregation' *Econometrica*, 61, pp. 353-85.

[76] Hendry, D. F. (1974), 'Stochastic Specification in and Aggregate Demand Model of the United Kingdom', *Econometrica*, 42, 559-78.

[77] - (1992), 'The Demand for M1 in the U.S.A., 1960-1988', *Review of Economic Studies*, 59, pp. 25-61.

[78] - and Mizon, G. E. (1978), 'Serial Correlation as a Convenient Simplication not a Nuisance: A Comment on a Study of the Demand for Money by the Bank of England', *Economic Journal*, 88, pp. 549-63.

[79] - and Von ungern-Sternberg, T. (1981), 'Liquidity and Inflation Effects on Consumers' Expenditure'. *In Essays in The Theory and Measurement Of Consumer Behaviour.* A. Deaton (ed.). Cambridge University Press.

[80] Keynes, J. M. (1936), *The General Theory of Employment, Interest and Money.* London: Macmillan Co.

[81] Kimball, M. S. (1990), 'Precautionary Savings in the Small and the Large', Econometrica, 58, pp. 53-73.

[82] Kuznets, S. (1946), *National Product since 1869.* New York: National Bureau of Economic Research.

[83] Leland, H. E. (1968), 'Saving and Uncertainty: The Precautionary Demand for Saving', *Quarterly Journal of Economics*, 82, pp. 465-73.

[84] Lucas, R. E. (1976), 'Econometric Policy Evaluation: a Critique'. In The Phillips Curve and Labour Markets, K. Brunner and A. H. Meltzer (ed.). Amsterdam: North Holland.

[85] Mankiw, N. G., and Shapiro, M. (1985), 'Trends Random Walks, and Tests of the Permanent Income Hypothesis', *Journal Of Monetary Economics,* 16, pp. 165-74.

[86] Mariger, R. P. (1987), 'A Life-Cycle Consumption Model with Liquidity Constraints: Theory and Empirical Results' *Econometrica*, 55, pp. 533-57.

[87] Mas-Colell, A., Whinston, M. D. and Green, J. R. (1995), *Microeconomic Theory.* Oxford University Press.

[88] Modigliani, F. (1949), 'Fluctuations in the Saving-Income Ratio: A Problem in Economic Forecasting', *Studies in Income and Wealth*, 11, New York: National Bureau of Economic Research, pp. 371-443.

[89] - (1986), 'Life cycle, Individual Thrift, and the Wealth of Nations', *American Economic Review*, 76, pp. 297-313.

[90] - and Brumberg, R. (1980), 'Utility and Aggregate Consumption Functions: An Attempt at Integration' in Abel, A. ed., *The Collected Papers of Franco Modigliani.* Cambridge: MIT Press.

[91] Molana, H. (1992), *Consumption Function.* In The *New Palgrave Dictionary of Money and Finance*, J. Eatwell, M. Milgate and P. Newman (eds.) London: Macmillan.

[92] - (1993), 'The Role of Income in the Consumption Function', *Scottish Journal of Political Economy*, 40, pp. 335-52.

[93] Muellbauer, J. N. J. (1988), 'Habits, Rationality and Myopia in the Life-Cycle Consumption Function', *Annales d'Economie et de Statistique*, 9, 47-70.

[94] - (1994), 'The Assessment: Consumer Expenditure', *Oxford Review of Economic Policy,* 10, 1-41.

[95] - and Lattimore, R. (1994), 'The Consumption Function: A Theoretical and Empirical Overview' in Pesaran, H. and Wickens, M.R. (eds) *Handbook of Applied Econometrics.*

[96] - and Murphy, A. (1993), 'Income Expectations, Wealth and Demography in the Aggregate UK Consumption Function', Unpublished Paper Presented to the HM Treasury Academic Panel, Nuffield College, Oxford University.

[97] Pagan, A. R. (1974), 'A Generalized Approach to the Treatment of Autocorrelation', Australian Economic Papers, 13, pp. 267-80.

[98] Pesaran, M.H. (1974), 'On the General Problem of Model Selection', Review of Economic Studies, 41, pp. 153-71.

[99] - (1992), *Savings and Consumption Behaviour.* In *The New Palgrave Dictionary of Money and Finance*, J. Eatwell, M. Milgate and P. Newman (eds.) London: Macmillan.

[100] Pischke, J. S. (1995), 'Individual Income, Incomplete Information and Aggregate Consumption', *Econometrica*, 63, pp. 805-40.

[101] Pratt, J. W. (1964), 'Risk Aversion in the Small and the Large', *Econometrica*, 32, pp. 122-36.

[102] Quah, D. (1990), 'Permanent and Transitory Movements in Labor Income: An Explanation for 'Excess Smoothness' in Consumption', *Journal of Political Economy*, 98, pp. 449-75.

[103] Ramsey, F. P. (1928), 'A Mathematical Theory of Saving', *Economic Journal*, 38, pp. 543-59.

[104] Romer, D. (1996), *Advanced Macroeconomics.* McGraw-Hill Advanced Series in Economics.

[105] Rothschild, M. and Stiglitz, J. E. (1970), 'Increasing Risk I: A Definition' *Journal of Economic Theory*, 2, pp. 225-43.

[106] -, - (1971), 'Increasing Risk II: Its Economic Consequences', *Journal of Economic Theory*, 3, pp. 66-84.

[107] Sargent, T. J. (1987), *Macroeconomic Theory, 2nd. Edition.* Academic Press, Inc.

[108] Seater, J. J. (1997), 'An Optimal Control Solution to the Liquidity Constraint Problem', *Economic Letters*, 54, pp. 127-34.

[109] Shazam Econometrics Computer Program v. 8.0 (1997), *User's Reference Manual*, McGraw-Hill.

[110] Shiller, R. (1978), 'Rational Expectations and the Dynamic Structure of Macroeconomic Models: A Critical Review' *Journal of Monetary Economics*, 4, pp. 1-44.

[111] Sims, C. A., Stock, J. H., and Watson, M. W. (1990), 'Inference in Linear Time Series Models with Some Unit Roots', *Econometrica*, 58, pp. 113-44.

[112] Skinner, J. (1988), 'Risky Income, Life-Cycle Consumption and Precautionay Saving', *Journal of Monetary Economics*, 22, pp. 237-55.

[113] Speight, A. E. H. (1990), *Consumption, Rational Expectations and Liquidity: Theory and Evidence.* Harvester Wheatsheaf.

[114] Stock, J. H. and West, K. D. (1988), 'Integrated Regressors and Tests of the Permanent Income Hypothesis', *Journal of Monetary Economics*, 21, pp. 85-95.

[115] Walker, C. M. (1996), 'Financial Management, Coping with Debt in Households under Financial Strain', *Journal of Economic Psychology*, 17, pp. 789-807.

[116] Wall, K. D., Preston, A. J., Bray, J. W. and Peston, M. H. (1975), 'Estimates of a Simple Control Model of the UK Economy', Chapter 14 in Renton, G. A. (ed.) *Modelling the Economy.* London: Heinemann Educational Books..

[117] Wallis, N. (1980), 'The Overlapping Generations Model of Fiat Money' in Kareden, J. and Wallace, N. (eds) *Models of Monetary Economies.* Minneapolis: Federal Reserve Bank of Minneapolis.

[118] West , K. D. (1988), 'The Insensitivity of Consumption to News About Income', Journal of Monetary Economics, 21, pp. 17-33.

[119] Whiteman, C. H. (1983), *Linear Rational Expectations Models: A User's Guide.* Minneapolis: University of Minnesota Press.

[120] - (1987), *Problems in Macroeconomic Theory: Solutions to Exercises from Thomas J. Sargent's Macroeconomic Theory, Second Edition.* Academic Press Inc.

[121] Zeldes, S. P. (1989a), 'Consumption and Liquidity Constraints: An Empirical Investigation', *Journal of Political Economy*, 97, pp. 305-46.

[122] - (1989b), 'Optimal Consumption with Stochastic Income: Deviations from Certainty Equivalence', *Quarterly Journal of Economics*, 104, pp. 275-98.