The permanent income hypothesis revisited: reconciling evidence from aggregate data with the representative consumer behaviour

Malley, Jim; Molana, Hassan

Publication date:
1999

University of Dundee

Dundee Discussion Papers in Economics

The Permanent Income Hypothesis Revisited: Reconciling Evidence from Aggregate Data with the Representative Consumer Behaviour.

Jim Malley

and

Hassan Molana

Department of
Economic Studies,
University of Dundee,
Dundee.
DD1 4HN

Working Paper
No. 105
October 1999
ISSN: 1473- 236X

The Permanent Income Hypothesis Revisited

Reconciling Evidence from Aggregate Data
with the Representative Consumer Behaviour

Jim Malley and Hassan Molana
University of Glasgow University of Dundee

May 1999

ABSTRACT: The evidence on the excessive smoothness and sensitivity of consumption with respect to income is sufficiently overwhelming to refute the rational expectations version of the permanent income hypothesis known as the random walk model. This paper proposes an alternative model which ***(i)*** is compatible with the "*excess smoothness*" and the "*excess sensitivity*" phenomena, ***(ii)*** can be interpreted as a rule-of-thumb revision, or smoothing, scheme similar to that proposed by Friedman, and ***(iii)*** can also be derived as the solution to a forward-looking intertemporal optimising problem where the rational consumer maximises a time-nonseparable utility function subject to the life-time budget constraint. Data from Canada, the U.K. and the U.S. are used to examine the proposed model. The findings strongly support the theoretical generalisation of the PIH proposed in the paper.

KEYWORDS: permanent income; excess sensitivity; excess smoothness; intertemporal separability

JEL CLASSIFICATION: E21

Corresponding Author: Hassan Molana
Department of Economics, University of Dundee, Dundee DD1 4HN, U.K.
Tel: (00)44-(0)138-234-4375, Fax: (00)44-(0)138-234-4691
Email: h.h.molana@dundee.ac.uk

Acknowledgement: The authors would like to thank Julia Darby, Ralph Monaco and Anton Muscatelli for helpful comments and suggestions. The usual disclaimer applies.

1. INTRODUCTION

The relevance of agents' heterogeneous behaviour and their asymmetric access to information - and hence the importance of aggregation - are now being increasingly recognised in macroeconomic analysis (see, for instance, Lewbel, 1994; Goodfriend, 1992; Clarida, 1991; and Galí, 1990). Nevertheless, micro-based macro-models which rely on a representative agent's optimal behaviour continue to play a crucial role in providing intuitive explanations for various macroeconomic phenomena (for typical examples see Blanchard and Fischer, 1989). One of the most popular behavioural frameworks used in such models is the *Permanent Income Hypothesis* (PIH) - or its *Life Cycle* version - which explains how a typical household may choose its optimal consumption path under different circumstances. The popularity of the PIH, as proposed by Friedman, stems mainly from two factors. First, it approximates a household's consumption path by a rule-of-thumb smoothing or revision process which can also be derived by solving a constrained utility maximisation problem that explicitly incorporates the structure of intertemporal preferences. Second, it yields a relationship between consumption and income which has theoretically interpretable parameters and is empirically superior to those implied by the earlier somewhat *ad hoc* hypotheses, e.g. the *Absolute* and the *Relative Income Hypotheses*.

However, a glance through the recent literature on the consumption function indicates that the PIH can no longer be fully credited with the advantage of yielding a robust empirical relationship between consumption and income (see, for instance, Deaton, 1992; and Molana, 1992). Clearly, this accumulated negative evidence cannot be disregarded when the original version of the PIH is used to approximate the intertemporal consumption decisions of a representative household in micro-based macro-models. Nevertheless given its intuitively appealing foundations, it would be desirable to generalise the PIH so that its implications cohere with the empirical regularities of the relationship between consumption and income reported in the literature.

This paper re-examines the existing evidence which has persuasively thrown doubt on the data consistency of the PIH. To summarise, while the existence of a unit root in the level of consumption cannot be rejected, and the first difference of consumption can be safely regarded as a stationary stochastic process, changes in consumption tend to exhibit a rather strong first order autoregressive pattern. This can be shown to cause both the "*excess sensitivity*" and the "*excess smoothness*" of consumption with respect to income. These phenomena were first discussed by Flavin (1981) and Deaton (1987), respectively, and are the main empirical objections to the so called *Random Walk* model which was implied by Hall's (1978) interpretation of the PIH. However, it is also pointed out that the serially correlated nature of the changes in consumption can be an indication of the fact that consumption habits tend to persist (see, for instance, Muellbauer, 1988). We use this idea to provide a generalisation of the PIH which reconciles the theory with the evidence. More particularly, we derive the optimal intertemporal path of consumption implied by the PIH when the set-up is modified to take account of the time-nonseparablity of preferences. We show that the resulting path is consistent with existing evidence as well as being interpretable as a rule-of-thumb smoothing or revision scheme of the kind originally proposed by Friedman. Data from Canada, the U.K. and the U.S. are used to examine the proposed model and the evidence revealed strongly supports the theoretical modifications suggested in the paper.

The rest of the paper is organised as follows. Section 2 outlines the standard theory and briefly explains how consumption may exhibit excess sensitivity and excess smoothness with respect to income. In Section 3 the PIH is generalised by relaxing the assumption of intertemporal separability of preferences. Section 4 examines data from Canada, the U.K. and the U.S. to throw light on the empirical relevance of the framework developed in Section 3 and Section 5 concludes the paper.

2. THEORY AND EXISTING EVIDENCE

It is convenient to start by restating the standard definitions which are commonly used in the literature and which will also be used throughout this paper. Permanent income is defined as the annuity associated with the present value of the human and non-human wealth

$$Y_t^P = r\left(A_t + \sum_{j=0}^{\infty} \rho^{j+1} E_t X_{t+j}\right), \tag{1}$$

where Y^P denotes permanent income, X is real (after tax) labour income, A is the real value of stock of financial wealth, r is the real (after tax) interest rate, $\rho=1/(1+r)$, and E_t denotes the expectations operator conditional on the information at t. The period-by-period and life-time budget constraints are

$$A_{t+j+1} = (1/\rho)A_{t+j} + X_{t+j} - C_{t+j}; \qquad j \geq 0, \tag{2}$$

and

$$\sum_{j=0}^{\infty} \rho^{j+1} C_{t+j} = A_t + \sum_{j=0}^{\infty} \rho^{j+1} X_{t+j} \ , \tag{3}$$

where C denotes the real value of consumption. Note that A is measured at the beginning of period t and C and X are payments which are assumed to take place at the end of period t. Given the above definitions, it can also be shown that Y^P satisfies the following

$$r \sum_{j=0}^{\infty} \rho^{j+1} E_t C_{t+j} = Y_t^P \ , \tag{4}$$

and

$$Y_t^P = (1/\rho) Y_{t-1}^P - ((1-\rho)/\rho) C_{t-1} + V_t \, , \tag{5}$$

where V is the annuity associated with the present value of the revisions in future income due to news (see Flavin, 1981, for details)

$$V_t = r \sum_{j=0}^{\infty} \rho^{j+1} \left(E_t X_{t+j} - E_{t-1} X_{t+j} \right) . \tag{6}$$

Note that V will behave as an unpredictable disturbance term if expectations are formed rationally. Thus, because $E_{t-1}V_t = 0$, it follows that a household which consumes its permanent income will also expect it to remain constant. In other words, if we let $C_{t-1} = Y_{t-1}^P$ then $E_{t-1} Y_t^P = Y_{t-1}^P$ follows. This simple rule-of-thumb consumption revision scheme, which is consistent with the solution to an intertemporal utility maximisation, lies at the heart of Friedman's contribution. However, Friedman's actual account deviated from this simple framework and resulted in some confusion which was later noted by other writers[1]. The latest version of the PIH is now known as the Random Walk (RW) model, which is derived from Friedman's model when the "rational expectations hypothesis" is used to revise permanent income. To illustrate this here we follow Campbell and Deaton (1989) and assume that labour income X can be approximated by an $ARIMA$(1,1,0) process

$$\Delta X_t = \lambda \Delta X_{t-1} + \varepsilon_t \, , \tag{7}$$

where Δ is the first difference operator, λ is a constant parameter, $0<\lambda<1$, and ε is an independently distributed random disturbance[2]. Given that equations (6) and (7) also imply the following, respectively

$$V_t = \sum_{j=0}^{\infty} \rho^j \left(E_t \Delta X_{t+j} - E_{t-1} \Delta X_{t+j} \right), \tag{6'}$$

and

$$E_t \Delta X_{t+j} - E_{t-1} \Delta X_{t+j} = \lambda^j V_t ; \qquad j \geq 0 , \tag{7'}$$

we can substitute from (7') into (6') to obtain

$$V_t = \pi \varepsilon_t , \tag{8}$$

where $\pi = (1 - \lambda\rho)^{-1} > 1$. The optimal intertemporal path of consumption can now be obtained as the reduced form of equations (5) and (8) and the assumption that households consume their permanent income, that is $C_{t+j} = Y_{t+j}^P$. These yield the so called RW model

$$\Delta C_t = \pi \varepsilon_t . \tag{9}$$

A version of this model was originally derived and tested by Hall (1978). Afterwards, two studies, Flavin (1981) and Deaton (1987) raised severe doubt about the empirical validity of this model. Flavin showed that the cross equation restrictions between the generalisations of (9) and (7) are violated empirically since (current and past) changes in actual income turn out to be significant when they are included as additional regressors in (9). Deaton compared the sample variances of ε_t and ΔC_t and illustrated that the data implies $Var(\Delta C_t) < Var(\varepsilon_t)$ hence violating the theoretical requirement that $\pi>1$ should hold in (9). Many other studies have examined these issues empirically for data sets from various countries (see Pesaran, 1992; and Deaton, 1992, for further details on both theoretical and empirical aspects). Overall, the evidence supports the joint proposal by Flavin and Deaton that consumption exhibits an excessive degree of sensitivity and smoothness with respect to income beyond that implied by the PIH[3].

3. THE PIH REVISITED

The purpose of this section is to generalise the path of consumption implied by the RW model in (9) to resolve the inconsistency between the theory and the evidence noted above. Following the PIH approach, this is carried out in two steps. First, we posit a more general consumption path which can be described as a rule-of-thumb revision, or a smoothing, scheme. Next, we demonstrate that such a scheme does in fact coincide with an optimal consumption path derived by solving the appropriate intertemporal utility maximisation problem[4]. The empirical consistency of this path is then examined in Section 4.

3.1. A Rule-of-Thumb Smoothing Scheme

To define a simple smoothing rule, we first substitute from equations (4) and (8) into (5) to obtain

$$(1-\rho)\sum_{j=0}^{\infty}\rho^{j}E_{t}C_{t+j}=\left((1-\rho)/\rho\right)\sum_{j=0}^{\infty}\rho^{j}E_{t-1}C_{t+j-1}+\left((1-\rho)/\rho\right)C_{t-1}+\pi\varepsilon_{t},$$

which can be rearranged to get

$$\sum_{j=0}^{\infty}\rho^{j}\left(E_{t}\Delta C_{t+j}-E_{t-1}\Delta C_{t+j}\right)=\pi\varepsilon_{t}. \tag{10}$$

Equation (10) states that the present value of the revision in the consumption plan should be proportional to the present shock to income. The simplest revision rule consistent with (10) is one based on exponentially declining weights, namely

$$E_{t}\Delta C_{t+j}-E_{t-1}\Delta C_{t+j}=\mu^{j}k\pi\varepsilon_{t};\qquad j\geq 0, \tag{11}$$

where μ is a constant parameter reflecting the weight used to smooth the path of ΔC_t and $k=1-\mu\rho$ ensures that the path in (11) remains consistent with the budget constraint in (10)[5]. Clearly, the restrictions $0<\mu<1$ and $0<k<1$, and equation (11) are also consistent with the following *ARIMA*(1,1,0) path for consumption

$$\Delta C_{t}=\varphi_{t}+\mu\Delta C_{t-1}+k\pi\varepsilon_{t}, \tag{12}$$

where φ is a deterministic drift parameter.

3.2. Utility Maximisation

Within the life cycle framework, the structure of preferences over the life time consumption profile is usually approximated by an additively separable utility function

$$U_t = \sum_{j=0}^{\infty} \delta^j u_{t+j} \,, \tag{13}$$

where $\delta=1/(1+d)$, d is the subjective rate of time preference which discounts future utilities, and $0<\delta<1$. The intertemporal separability assumption implies

$$u_{t+j} = u(C_{t+j}), \tag{14}$$

where $u(.)$ is continuous and smoothly concave. By substituting (14) into (13) and choosing the path of C_{t+j} to maximise $E_t(U_t)$ subject to the budget constraint in equation (3) above, the following first order conditions are obtained

$$E_t u'\left(C_{t+j}\right) = \psi\left(\rho/\delta\right)^j; \qquad j \geq 1, \tag{15}$$

where $u'(C)$ denotes the marginal utility of consumption and ψ is the Lagrange multiplier. If we now let $\delta=\rho$ and $u(x)=-exp(-\gamma x)$ for $\gamma>0$, the above conditions can be shown to imply[6]

$$E_t\left(\Delta C_{t+j}\right) = (\gamma/2)\left(Var\left(C_{t+j-1}\right) - Var\left(C_{t+j}\right)\right); \qquad j \geq 1. \tag{16}$$

This is consistent with the following version of the RW model[7]

$$\Delta C_{t+j} = \eta_{t+j} + \xi_{t+j}; \qquad j \geq 1, \tag{17}$$

where

$$\eta_{t+j} = (\gamma/2)\left(Var\left(C_{t+j-1}\right) - Var\left(C_{t+j}\right)\right) \tag{18}$$

is the deterministic drift factor and ξ is an unpredictable random disturbance term if the rational expectations hypothesis is assumed. The consistency condition which ensures that (17) obeys the budget constraint[8] is $\xi_{t+j}=\pi\varepsilon_{t+j}$. However, this model has the shortcomings noted above.

The intertemporal separability assumption is nevertheless a rather arbitrary simplification which is usually assumed to facilitate analytical tractability. In fact, the strong correlation between current and past changes in consumption which are repeatedly reported can be interpreted as evidence against separability. The response to this issue in the literature is now growing and there are already a number of studies which address the implications, as well as the empirical validity, of the intertemporal separability assumption. These studies explore the possibility and consequences of allowing for intertemporally nonseparable preferences due to various behavioural phenomena, e.g. rational addiction, habit persistence, seasonality, subjective discounting and aversion to intertemporal trade-offs. Winder and Palm (1991) provide a detailed explanation of the technical and behavioural aspects of the problem[9]. Here, we present a simple generalisation by extending the instantaneous utility function to depend on both current and past consumption. The argument relies on the intuition that the choice of consuming $C_{t+j} \neq C_{t+j-1}$ involves two sources of satisfaction due to the level C_{t+j} and the change $\Delta C_{t+j} \neq 0$. This implies replacing $u(C_{t+j})$ in (14) with $u(C_{t+j}, \Delta C_{t+j})$. It can be further assumed that the separate effects of the arguments C_{t+j} and ΔC_{t+j} on the level of satisfaction are due to their relative weights, hence replacing (14) with

$$u_{t+j} = u\left(\phi C_{t+j} + \mu \Delta C_{t+j}\right); \qquad \phi>0; \qquad \mu>0, \tag{19}$$

where ϕ and μ are constant parameters reflecting the relative weights. Given that the normalisations $\phi+\mu=1$, $0<\phi<1$ and $0<\mu<1$ can be applied without loss of generality, (19) can be replaced by

$$u_{t+j} = u\left(C_{t+j} - \mu C_{t+j-1}\right), \qquad 0<\mu<1. \tag{20}$$

Using (20) instead of (14) and repeating the maximisation, we now obtain the following first order conditions

$$E_t\left(u'_{t+j} - \mu\delta u'_{t+j-1}\right) = \psi\left(\rho/\delta\right)^j; \qquad j \geq 0, \tag{21}$$

where u'_{t+j} is the marginal utility with respect to $\left(C_{t+j} - \mu C_{t+j-1}\right)$ and ψ is the Lagrange multiplier. Now, if we let $\delta=\rho$ as before, a sufficient condition for (21) to hold for all $j\geq 0$ is[10] $E_t\left(\Delta u'_{t+j}\right)=0$ which, on the assumption that $u(x)=-exp(-\gamma x)$ for $\gamma>0$, implies

$$E_t\left(\Delta C_{t+j}-\mu\Delta C_{t+j-1}\right)=\left(\gamma/2\right)\left(Var\left(C_{t+j-1}-\mu C_{t+j-2}\right)-Var\left(C_{t+j}-\mu C_{t+j-1}\right)\right). \tag{22}$$

The stochastic version of equation (22) is

$$\Delta C_{t+j}=\varphi_{t+j}+\mu\Delta C_{t+j-1}+\omega_{t+j}; \qquad j\geq 1, \tag{23}$$

where

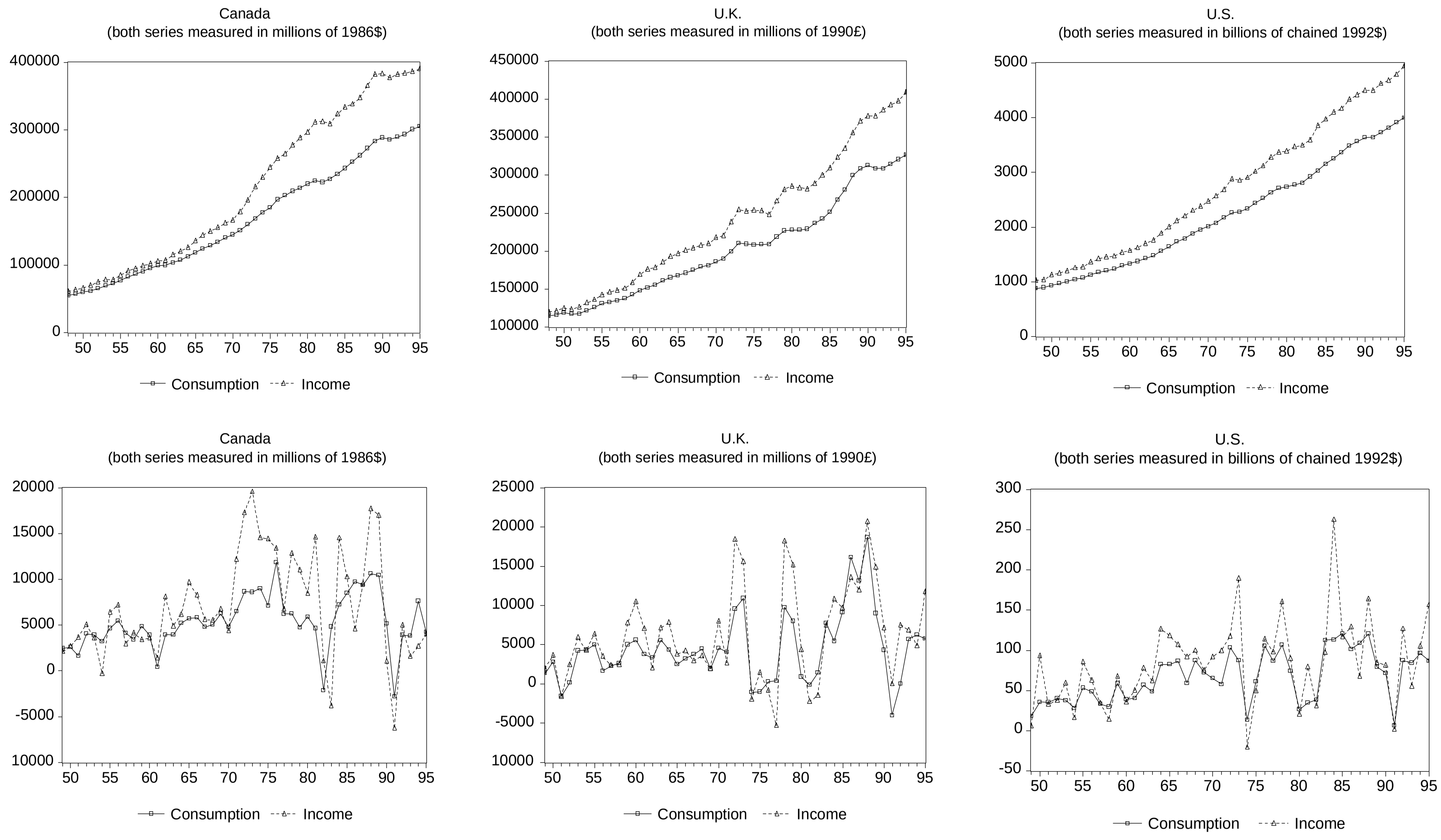
$$\begin{aligned}\varphi_{t+j}&=\left(\gamma/2\right)\left(\mu^2 Var(C_{t+j-2})+\left(1-\mu^2\right)Var(C_{t+j-1})-Var(C_{t+j})\right)\\ &\quad \gamma\left(Cov(C_{t+j},C_{t+j-1})-Cov(C_{t+j-1},C_{t+j-2})\right).\end{aligned} \tag{24}$$

Thus, φ is the deterministic drift parameter and ω is an unpredictable random disturbance term if the rational expectations hypothesis is assumed. The consistency condition which requires (23) to obey the budget constraint can then be shown to be $\omega_t=k\pi\varepsilon_t$, where $k=(1-\mu\rho)$. This provides the theoretical justification for our otherwise rule-of-thumb smoothing scheme in equation (12).

4. EVIDENCE FROM CANADA, THE U.K. AND THE U.S.A

In this section we use data from Canada, the U.K. and the U.S. to assess whether the theoretical generalisation suggested above is supported empirically. The data series used are annual observations for the period 1948-95 on consumers' expenditure on nondurable goods and services and personal disposable income measured at constant prices. These are denoted by *C* and *Y*, respectively. Before proceeding to explain our results, two points should be noted at the outset. First, the 1948-95 time span was chosen because it is the longest common period over which data are available for the three countries, while the annual frequency was used to avoid the problems associated with modelling the seasonal components of the series[11]. Second, while the underlying theory refers to total consumption - i.e. expenditure on nondurable goods and services plus the value of services from durable goods - and disposable labour income, empirical analysis are conducted using nondurable consumption and disposable total income[12].

Figure 1. Pattern of Levels and Changes in Consumer's Expenditure on Nondurable Goods and Services and Personal Disposable Income



Sources: Canadian Socio-Economic Information System, April ′97; U.K. Blue Book, July ′96 Release; U.S. NIPA, Survey of Current Business, May ′97.

Table 1. Volatility of Consumption and Income

	1949-95		1949-69		1970-95	
	MEAN	*S.D.*	*MEAN*	*S.D.*	*MEAN*	*S.D.*
$\Delta C1_t$	5327.1	2936.7	4070.5	1437.6	6342.0	3437.7
$\Delta Y1_t$	6999.6	5719.9	4795.9	2462.6	8779.5	6926.3
$\Delta C2_t$	4504.4	4456.3	3169.2	1808.5	5582.8	5587.1
$\Delta Y2_t$	6157.1	5879.7	4290.7	2740.5	7664.7	7232.7
$\Delta C3_t$	66.4	31.1	51.1	20.8	78.8	32.8
$\Delta Y3_t$	83.4	52.7	64.4	34.8	98.7	60.0

In Tables 1-6, Canada=1, U.K.=2 and U.S.=3.

The annual levels and changes in consumption and income shown in Figure 1 and summary statistics in Table 1 reveal that (*i*) the gap between *Y* and *C* has increased over time, (*ii*) income has, in general, been more volatile than consumption, and (*iii*) the volatility in both series rose drastically after the early 1970s[13]. Tables 2 and 3 provide further details on the time series behaviour of consumption and income. Given that the power of univariate unit root tests can vary considerably (see, for instance Pantula *et al*., 1994), several alternative tests are presented in Table 2. The results of the various tests uniformly suggest that both C_t and Y_t have a unit root and their first differences, ΔC_t and ΔY_t, are stationary. Further, the tests in Table 2 show that while ΔC_t and ΔY_t do not contain any stochastic trend, they exhibit a strong *AR*(1) pattern since their autocorrelation coefficients, reported in Table 3, are significant only at the first lag.

Table 2. Unit Root Tests (excluding a linear deterministic trend), 1950-1995

		Lag-0		Lag-1		Lag-2	
		STAT.	P-VALUE	STAT.	P-VALUE	STAT.	P-VALUE
	WS	2.85	1.00	0.15	0.99	0.01	0.99
C1	ADF	1.53	1.00	0.41	0.98	0.43	0.98
	PP	0.39	0.97	0.37	0.97	0.37	0.97
	WS	-3.92	0.001	-3.30	0.005	-3.16	0.007
ΔC1	ADF	-3.86	0.002	-3.40	0.011	-3.16	0.022
	PP	-22.0	0.007	-21.7	0.008	-22.0	0.007
	WS	1.81	1.00	0.16	0.99	0.10	0.99
Y1	ADF	0.49	0.98	-0.08	0.95	0.05	1.00
	PP	0.18	0.97	0.15	0.96	0.14	0.99
	WS	-4.09	0.001	-3.97	0.001	-2.34	0.077
ΔY1	ADF	-3.98	0.002	-3.87	0.002	-2.23	0.196
	PP	-23.2	0.005	-24.4	0.004	-21.6	0.008

Table 2 *Continued*

		Lag-0		Lag-1		Lag-2	
		STAT.	P-VALUE	STAT.	P-VALUE	STAT.	P-VALUE
	WS	2.42	1.00	0.21	0.99	0.37	1.00
C2	ADF	1.70	1.00	0.31	0.98	0.67	0.99
	PP	0.82	0.99	0.77	0.98	0.74	0.98
	WS	-3.67	0.002	-4.07	0.001	-4.02	0.001
ΔC2	ADF	-3.48	0.008	-3.97	0.001	-3.97	0.002
	PP	-19.1	0.015	-20.9	0.009	-20.9	0.009
	WS	2.67	1.00	0.80	1.00	0.83	1.00
Y2	ADF	1.75	1.00	0.88	0.99	1.41	1.00
	PP	0.84	0.99	0.81	0.99	0.81	0.99
	WS	-4.68	0.0001	-4.91	0.00004	-3.98	0.001
ΔY2	ADF	-4.50	0.0002	-4.88	0.00004	-3.93	0.002
	PP	-28.3	0.0015	-29.8	0.001	-27.5	0.002
	WS	3.38	1.00	0.45	1.00	0.31	1.00
C3	ADF	3.49	1.00	1.83	1.00	1.91	1.00
	PP	0.70	0.98	0.69	0.98	0.69	0.98
	WS	-3.83	0.001	-3.31	0.005	-2.99	0.012
ΔC3	ADF	-3.80	0.003	-3.36	0.012	-3.08	0.028
	PP	-21.5	0.008	-21.5	0.008	-21.3	0.008
	WS	2.92	1.00	0.97	1.00	0.54	1.00
Y3	ADF	2.20	1.00	1.79	1.00	1.59	1.00
	PP	0.64	0.98	0.64	0.98	0.64	0.98
	WS	-5.82	2.6E-06	-4.08	0.0047	-3.47	0.0029
ΔY3	ADF	-5.61	1.2E-06	-3.91	0.0020	-3.91	0.0112
	PP	-38.0	0.0001	-37.9	0.0001	-37.9	0.0001

WS, ADF and PP are the Weighted Symmetric (see Pantula *et al.*, 1994), the Augmented Dickey-Fuller (see Dickey & Fuller, 1979, 1981) and the Phillips-Perron (see Phillips & Perron, 1988) tests for unit roots. The P-values for the above tests are calculated using the tables reported in MacKinnon (1994). Note that the results of the above tests remain unaltered when a linear deterministic trend is added to the testing equations. To preserve space these results are not reported here but will be made available upon request.

Table 3. Testing the Autocovariance Structure of Stationary Variables (1952-1995)

	Order	AC	S.E.	B-P	L-B
	1	0.503	0.146	11.90 [0.001]	12.67 [0.000]
$\Delta C1_t$	2	0.261	0.179	15.10 [0.001]	16.16 [0.000]
	3	0.058	0.187	15.26 [0.002]	16.33 [0.001]
	1	0.470	0.146	10.38 [0.001]	11.06 [0.001]
$\Delta Y1_t$	2	0.104	0.175	13.51 [0.004]	11.61 [0.003]
	3	0.236	0.176	13.51 [0.004]	14.53 [0.002]
	1	0.560	0.146	14.71 [0.000]	15.67 [0.000]
$\Delta C2_t$	2	0.158	0.186	15.89 [0.000]	16.96 [0.000]
	3	-0.133	0.189	16.72 [0.001]	17.88 [0.000]

Table 3 ***Continued***	Order	AC	S.E.	B-P	L-B
	1	0.347	0.146	5.674 [0.017]	6.044 [0.014]
$\Delta Y2_t$	2	-0.083	0.163	5.998 [0.050]	6.397 [0.041]
	3	-0.153	0.163	7.099 [0.069]	7.622 [0.054]
	1	0.508	0.146	12.13 [0.000]	12.92 [0.000]
$\Delta C3_t$	2	0.245	0.180	14.94 [0.001]	15.98 [0.000]
	3	0.087	0.187	15.30 [0.002]	16.38 [0.001]
	1	0.116	0.146	0.630 [0.427]	0.671 [0.413]
$\Delta Y3_t$	2	0.092	0.148	1.029 [0.598]	1.105 [0.575]
	3	0.024	0.149	1.055 [0.788]	1.134 [0.769]

AC, S.E., B-P and L-B are the Autocorrelation Coefficient, Standard Error of the Autocorrelation Coefficient, Box-Pierce and Ljung-Box statistics for the corresponding lag. B-P and L-B are distributed $\chi^2_{(n)}$ where n is the number of lags. The numbers in square brackets are P-values.

Clearly, the evidence that ΔC_t is correlated with its own past is sufficiently strong to reject the hypothesis that the level of consumption follows a random walk process. However, further investigation of the above issues requires a measure of income innovation. The common practice to construct such a measure would be based upon empirical univariate time series approximations of the income process. But given that the purpose of this paper is to remain as close as possible to the original framework of the PIH, a preferable way in which to approximate income innovation is to use an exponential smoothing scheme similar to that suggested by Friedman. More explicitly, let ξ denote the unexpected component of current income, that is $\xi_t=Y_t-E_{t-1}Y_t$. The representative agent may then be assumed to use the following updating rule[14]

$$E_t \Delta Y_{t+j} - E_{t-1} \Delta Y_{t+j} = \eta^j \xi_t; \qquad j \geq 0, \tag{25}$$

where η is a constant parameter, $0<\eta<1$. This rule would be optimal if the actual income series were generated by the following *ARIMA*(1,1,0) process

$$\Delta Y_t = \varphi + \eta \Delta Y_{t-1} + \xi_t, \tag{26}$$

where ξ is an independently distributed random variable (see Muth, 1960). Estimates of (26) are reported in Table 4 and the relevant tests suggest that the hypothesis that the corresponding $\hat{\xi}_t$ is the realisation of an independently distributed random variable cannot be rejected. Thus, $\hat{\xi}_t$ provides a reasonably acceptable approximation of the income innovation for our purpose and may be used to re-examine the following points.

Table 4. OLS Estimation of the *ARIMA*(1,1,0) Income Generating Process, 1950-1995

$$\Delta Yi_t = \varphi i + \eta i \Delta Yi_{t-1} + \xi i_t ; \qquad i = 1,2,3$$

	$\hat{\varphi}1$	$\hat{\varphi}2$	$\hat{\varphi}3$	$\hat{\eta}1$	$\hat{\eta}2$	$\hat{\eta}3$
coef.	3767.0	4106.5	66.6	0.47	0.35	0.21
t-stat.	3.14	3.46	5.64	3.58	2.51	1.68

Diagnostic Tests

Statistics	*Canada*	*U.K.*	*U.S.*
S_1	0.74 [0.39]	2.29 [0.13]	0.26 [0.61]
S_2	1.98 [0.16]	2.85 [0.09]	0.003 [0.96]
S_3	0.59 [0.97]	5.26 [0.07]	0.56 [0.76]
S_4	0.27 [0.60]	0.09 [0.76]	0.46 [0.50]

Volatility of Consumption, Income and Unexpected Income

	1950-95 *S.D.*	1950-69 *S.D.*	1970-95 *S.D.*
$\Delta C1_t$	2937.4	1425.1	3437.7
$\Delta Y1_t$	5735.6	2445.3	6926.3
$\hat{\xi}1_t$	5046.9	2415.4	6311.3
$\Delta C2_t$	4480.5	1806.0	5587.1
$\Delta Y2_t$	5910.6	2757.1	7232.7
$\hat{\xi}2_t$	5529.3	2613.5	6897.9
$\Delta Y3_t$	52.0	32.9	60.0
$\Delta C3_t$	30.6	19.9	32.8
$\hat{\xi}3_t$	49.5	31.3	44.4

Autocovariance Structure of $\hat{\xi}i_t$

	Order	AC	S.E.	B-P	L-B
	1	0.061	0.147	0.176 [0.675]	0.187 [0.665]
$\hat{\xi}1_t$	2	-0.275	0.148	3.655 [0.161]	3.983 [0.136]
	3	0.207	0.159	5.624 [0.131]	6.181 [0.103]
	1	0.082	0.147	0.306 [0.580]	0.327 [0.568]
$\hat{\xi}2_t$	2	-0.211	0.148	2.373 [0.305]	2.581 [0.275]
	3	-0.077	0.155	2.649 [0.449]	2.889 [0.409]
	1	-0.047	0.147	0.100 [0.751]	0.107 [0.744]
$\hat{\xi}3_t$	2	0.216	0.148	2.245 [0.325]	2.446 [0.294]
	3	0.059	0.154	2.403 [0.493]	2.623 [0.453]

Numbers in square bracket are P-values; S_1 is the Lagrange multiplier $\chi^2_{(1)}$ statistic for residual first-order serial correlation; S_2 is the Ramsey RESET $\chi^2_{(1)}$ test for functional form misspecification (based on the square of fitted values); S_3 is the $\chi^2_{(2)}$ test for the normality (based on a test of skewness and kurtosis of the residuals); and S_4 is the $\chi^2_{(1)}$ statistic for heteroscedasticity (based on a regression of squared residuals on squared fitted values). A comparison of the distribution of standardised residuals with the normal distribution indicated two outliers (i.e. outside 3-standard deviations) for the U.S. (i.e. in 1974 (+ outlier) and 1984 (- outlier)) . Accordingly, the above regression for the U.S. incorporates dummy variables for these years which are likely to be capturing the 1973 oil shock effect and the high growth rate (highest since the Korean conflict) that followed the recovery from deep recession (deepest since the Great Depression) of the early 1980s.

First consider Deaton's "*smoothness paradox*". Given that $\hat{\xi}_t$ is derived by minimising the sample variance of ξ_t, we may check the smoothness problem by comparing the standard deviation of $\hat{\xi}_t$ with that of ΔC_t. The sample standard deviations are also reported in Table 4 and confirm that consumption is even less volatile than the income innovation. As noted by Deaton, this contradicts the implication of the RW model in equation (9) above since $\pi > 1$ ought to hold.

Next, when ΔC_t is partially predictable - since it exhibits a strong first order autoregressive pattern - but income innovations are independently distributed and hence are unpredictable, the RW model will show severe symptoms of misspecification because the two sides of equation (9) are not balanced. In particular, given that ΔY_t and ΔC_t exhibit similar autoregressive patterns, it follows that ΔY_t or ΔY_{t-1} will almost certainly have a significant coefficient if included as a regressor in equation (9). This evidence can then be interpreted as an indication of Flavin's "*excess sensitivity phenomenon*". To examine this we ought to regress ΔC_t on ΔY_{t-1} and $\hat{\xi}_t$. However, because on one hand $\hat{\xi}_t$ is a linear combination of ΔY_t and ΔY_{t-1} and, on the other hand, C_t and Y_t are both first difference stationary, regressing ΔC_t on ΔY_{t-1} and $\hat{\xi}_t$ will not be expected to generate well behaved (unpredictable) residuals unless C_t and Y_t do not cointegrate. This is because if C_t and Y_t were cointegrated then the above mentioned residuals would exhibit misspecification symptoms unless the underlying regression was also augmented by C_{t-1} and Y_{t-1} (see Hansen and Sargent, 1981, and Campbell and Shiller, 1987, for details). Recall, nevertheless, from Figure 1 that the gap between Y and C widens over our sample period, clearly indicating that the two series are unlikely to cointegrate. This is further confirmed by the formal cointegration tests reported in Table 5. As in Table 2, we again present several alternative tests to indicate the robustness of our findings.

The OLS estimates of regressions used to examine the sensitivity problem, as proposed above, are reported in Table 6. Clearly, the significant explanatory role of ΔY_{t-1} in each of countries reported in column **(I)** provides evidence of excess sensitivity and confirms the empirical inadequacy of the RW model in (9). Motivated by the theoretical extensions discussed in Section 3, column **(II)** reports the results of adding ΔC_{t-1} to the augmented RW model in column (**I**). The important result in this case is that the inclusion of ΔC_{t-1} renders the coefficient of ΔY_{t-1} statistically insignificant for all countries. Finally when ΔY_{t-1} is eliminated, column **(III)** shows that the specification implied by the theoretical generalisation of the PIH described in Section 3 is empirically superior for the U.K. and the U.S. and performs equally well as the augmented RW for Canada (e.g. compare the encompassing test statistics, S_5 and S_6 reported in

Table 5. Cointegration Tests (excluding a linear deterministic trend), 1950-1995

Engle-Granger cointegration tests

	$C1_t, Y1_t$			$C2_t, Y2_t$			$C3_t, Y3_t$		
Lags	0	1	2*	0	1	2*	0	1	2*
TestStat	-0.84	-1.36	-1.17	-2.20	-3.01	-2.47	-3.22	-2.27	-2.05
P-value	0.930	0.081	0.865	0.424	0.109	0.292	0.067	0.390	0.502

Johansen (trace) cointegration tests

	$C1_t, Y1_t$			$C2_t, Y2_t$			$C3_t, Y3_t$		
Lags	0	1*	2	0	1	2*	0	1*	2
Eigval1	0.055	0.058	0.047	0.225	0.225	0.157	0.273	0.212	0.198
Eigval2	0.006	0.003	0.006	0.010	0.007	0.049	0.132	0.019	0.022
H_0:r=0	2.56	2.46	1.82	10.88	10.19	8.18	18.85	10.04	9.00
P-value	0.922	0.925	0.940	0.368	0.428	0.608	0.037	0.441	0.535
H_0:r≤1	0.258	0.111	0.023	0.432	0.261	1.85	5.80	0.730	0.816
P-value	0.779	0.793	0.802	0.761	0.779	0.583	0.135	0.728	0.718

The P-values for the Engle-Granger (1987) and Johansen (1988) cointegration tests are based on the tables reported in MacKinnon (1994) and Osterwald-Lenum (1992) respectively. A **'*'** indicates the optimal lag-length chosen by the Akaike Information Criterion. Note that the results of the above tests remain unaltered when a linear deterministic trend is added to the testing specifications. To preserve space they are not reported here but will be made available upon request.

Table 6. OLS Estimates of Regression Equations Explaining ΔC_{it}, 1950-1995

	Canada			U.K.			U.S.		
regressors	*coefficient estimates*			*coefficient estimates*			*coefficient estimates*		
	(I)	**(II)**	**(III)**	**(I)**	**(II)**	**(III)**	**(I)**	**(II)**	**(III)**
Intercept	3846.8	3212.0	3114.1	2929.7	2714.7	2703.0	53.13	42.00	42.26
	(9.29)	(4.61)	(4.37)	(6.43)	(6.32)	(6.17)	(8.13)	(6.14)	(6.11)
ΔY_{it-1}	0.218	0.129	----	0.271	-0.015	----	0.176	-0.023	----
	(4.14)	(1.30)		(4.36)	(-0.13)		(1.99)	(-0.12)	
ΔC_{it-1}	----	0.237	0.425	----	0.435	0.418	----	0.422	0.428
		(1.19)	(3.69)		(2.38)	(4.62)		(4.49)	(5.30)
$\hat{\xi}_{it}$	0.349	0.328	0.312	0.630	0.556	0.559	0.498	0.414	0.382
	(4.53)	(4.35)	(3.97)	(8.04)	(7.60)	(8.00)	(7.13)	(1.77)	(3.57)
statistics									
Adj R^2	0.522	0.538	0.519	0.718	0.756	0.762	0.503	0.55	0.56
σ	2029.8	1996.8	2037.5	2380.8	2212.4	2187.0	21.53	20.49	20.26
S_1	2.22[0.14]	0.12[0.73]	2.25[0.13]	7.75[0.01]	2.99[0.08]	0.85[0.36]	6.15[0.01]	2.16[0.14]	2.03[0.15]
S_2	0.48[0.49]	0.03[0.86]	0.41[0.52]	1.68[0.20]	0.81[0.37]	0.82[0.37]	3.37[0.07]	3.59[0.06]	3.41[0.07]
S_3	3.61[0.17]	3.66[0.16]	3.43[0.18]	2.88[0.24]	0.93[0.63]	0.93[0.63]	10.8[0.01]	2.60[0.27]	3.52[0.17]
S_4	0.71[0.40]	1.55[0.21]	2.01[0.16]	5.57[0.18]	5.85[0.16]	5.88[0.02]	1.58[0.21]	0.67[0.41]	0.83[0.36]
S_5	1.56[0.12]	----	1.67[0.10]	2.79[0.01]	----	-0.13[0.90]	2.35[0.02]	----	-0.22[0.03]
S_6	2.44[0.13]	----	2.77[0.10]	7.79[0.01]	----	0.02[0.90]	5.51[0.02]	----	0.05[0.24]

Numbers in parentheses below coefficient estimates are t-ratios adjusted for heteroscedasticity; σ is the standard error of the regression; S_1 is the Lagrange multiplier $\chi^2_{(1)}$ statistic for residual first-order serial correlation; S_2 is the Ramsey RESET $\chi^2_{(1)}$ test for functional form misspecification; S_3 is the $\chi^2_{(2)}$ statistic for normality of the residuals; S_4 is the $\chi^2_{(1)}$ statistic for heteroscedasticity; and S_5 and S_6 are non-nested tests for the model in the corresponding column against it rival model. S_5 is JA test statistic proposed by Fisher and McAleer (1981) and has a t-distribution, whereas S_6 is the Encompassing test statistic suggested by Mizon and Richard (1986) and has a F(1,42) distribution.

the lower part of columns **(I)** and **(III)**)[15]. However, model (**III**) is nonetheless preferable to (**I**) for Canada, since (**I**) has a clear disadvantage in that it does not possess theoretically interpretable parameters. Thus, the results in Table 6 show that first, a regression based on (23) will not exhibit excess sensitivity and second, the excess smoothness phenomenon will no longer cause an inconsistency even if π>1 provided that μ and k satisfy the restriction $(1-\mu^2)>(k\pi)^2$. As a first approximation, our estimates in column (**III**) of Table 6 [$\mu1=0.425$, $k1\pi1=0.312$; $\mu2=0.418$, $k2\pi2=0.559$; and $\mu1=0.428$, $k1\pi1=0.382$] can be used. Clearly these estimates satisfy the above restriction thus providing further support for our interpretation[16].

Finally, it might be argued that because we have used annual data the time aggregation bias would shed doubt on the robustness of our results. In general, if consumers make planning decisions on a sub-annual basis then the errors in our model could have a moving average structure and might not orthogonal to the lagged changes in consumption and the income innovation (see Wickens and Molana, 1984). However since we were unable to obtain significant MA(1) or MA(2) errors in any country, we may safely conclude that the extent of this bias is minimal.

5. SUMMARY AND CONCLUSIONS

The evidence on the excess smoothness and sensitivity of consumption with respect to income is sufficiently overwhelming to motivate justifying these empirical phenomena theoretically. In particular, since this evidence is mainly directed towards refuting the rational expectations version of the permanent income model, it is interesting to ask whether an alternative optimal, forward-looking, behavioural model can be formulated within the PIH framework which is compatible with these phenomenon. The attractiveness of such a model lies mainly in its ability to reconcile the evidence from aggregate data with a simple but plausible theory based on the representative agent behaviour.

In an attempt to provide an answer to the above question we have shown first that a simple rule-of-thumb smoothing scheme similar to that suggested by Friedman's PIH implies that the change in consumption should depend on its own past and a drift factor containing surprise income. We have subsequently provided a theoretical justification for this rule by showing it to be consistent with the solution to a forward-looking intertemporal optimising problem where the rational agent is assumed to maximise a life-time utility function which allows for intertemporally nonseparable preferences. Finally, we have used data from Canada, the U.K. and the U.S. to test the empirical plausibility of this generalisation. Our evidence is strongly supportive in that data from all three countries are entirely consistent with a simple behaviour by consumers who choose to allocate a windfall gain or loss so as to maintain a smooth consumption path. It is therefore concluded that the empirical consistency of the PIH can be restored if it is generalised to yield a consumption path which relates the change in consumption to its own past and to income innovations.

References

Akaike, H. (1973). Information theory and the extension of the maximum likelihood principle. In Proceedings of the Second International Symposium on Information Theory, eds. Petrov, B.N. and Caski, F., Budapest: Akedemiai Kiado.

Attfield, C. Demery, D., and Duck, N. (1992). Partial adjustment and the permanent income hypothesis. *European Economic Review*, 36: 1205-1222.

Becker, G.S. and Murphy, K.M. (1988). A theory of rational addiction. *Journal of Political Economy*, 96: 675-700.

Blachard, O.J and Fischer, S. (1989). Lectures on Macroeconomics. Cambridge: MIT Press.

Caballero, R.J. (1990). Consumption puzzles and precautionary savings. *Journal of Monetary Economics*, 25: 113-136.

Campbell, J.Y. and Deaton, A.S. (1989). Why is consumption so smooth? *Review of Economic Studies*, 56: 357-374.

Campbell, Y.J. and Mankiw, N.G. (1991). The response of consumption to income: a cross-country investigation. *European Economic Review*, 35: 723-767.

Campbell, J.Y. and Shiller, R.J. (1987). Cointegration and tests of present value models. *Journal of Political Economy*, 95: 1062-1088.

Carroll, C. (1994). How does future income affect current consumption? *Quarterly Journal of Economics*, 108: 111-147.

Clarida, R.H. (1991). Aggregate stochastic implications of the life cycle hypothesis. *Quarterly Journal of Economics*, 106: 851-867.

Constantinides, G. M. (1990). Habit formation: a resolution of the equity premium puzzle. *Journal of Political Economy*, 98: 519-545.

Darby, M.R. (1974). The permanent income theory of consumption: a restatement. *Quarterly Journal of Economics*, 88: 228-50.

Deaton, A.S. (1987). Life cycle models of consumption: is the evidence consistent with the theory? In Advances in Econometrics, ed. Amsterdam: North Holland.

Deaton, A.S. (1992). Understanding Consumption. Clarendon Press, Oxford.

Dickey, D. and Fuller, W.A. (1979). Distribution of the estimates for autoregressive time series with a unit root, *Journal of the American Statistical Association*, 74: 427-431.

Dickey, D. and Fuller, W.A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root, *Econometrica*, 49: 1057-1072.

Dockner, E.J. and Feichtinger, G. (1993). Cyclical consumption patterns and rational addiction. *American Economic Review*, 83: 256-263.

Engle, R.F. and Granger, C.W.J. (1987). Cointegration and error correction: representation, estimation and testing, *Econometrica*, 55: 251-276.

Fisher, G.R. and McAleer, M. (1981). Alternative procedures and associated tests of significance for non-nested hypotheses. *Journal of Econometrics*, 16: 103-119.

Flavin, M.A. (1981). The adjustment of consumption to changing expectations about future income. *Journal of Political Economy*, 89: 974-1009.

Flavin, M.A. (1993). The excess smoothness of consumption: identification and interpretation. *Review of Economic Studies*, 60: 651-666.

Galí, J. (1990). Finite horizons, life-cycle savings, and time-series evidence on consumption. *Journal of Monetary Economics*, 26: 433-452.

Galí, J. (1991). Budget constraints and time-series evidence on consumption. *American Economic Review*, 81: 1238-1235.

Goodfriend, M. (1992). Information-aggregation bias. *American Economic Review*, 82: 509-519.

Hall, R.E. (1978). Stochastic implications of the life cycle-permanent income hypothesis: theory and evidence. *Journal of Political Economy*, 86: 971-987.

Hansen, L.P. and Sargent, T.J. (1981). Linear rational expectations models for dynamically interrelated variables. In Rational Expectations and Econometric Practice, eds. Lucas, R.E. and Sargent, T.J., Minneapolis: University of Minnesota Press.

Heaton, J. (1993). The interaction between time-nonseparable preferences and time aggregation. *Econometrica*, 61: 353-385.

Iannaccone, L.R. (1986). Addiction and satiation, *Economics Letters*, 21: 95-99.

Johansen, S. (1988). Statistical analysis of cointegration vectors, *Journal of Economic Dynamics and Control*, 12: 231-254.

Johnson, M.B. (1971). Household Behaviour. Harmondsworth: Penguin.

Lewbel, A. (1994). Aggregation and Simple Dynamics. *American Economic Review*, 84: 905-918.

Loewenstein, G. and Prelec, D. (1992). Anomalies in intertemporal choice: evidence and an interpretation. *Quarterly Journal of Economics*, 100: 573-597.

MacKinnon, J.G. (1994). Approximate asymptotic distribution functions for unit-root and cointegration tests, *Journal of Business and Economic Statistics*, 12: 167-176.

Mizon, G.E. and Richard, J-F. (1986). The encompassing principle and its applications to testing non-nested hypotheses. *Econometrica*, 54: 657-78.

Molana, H. (1992). Consumption function. In The New Palgrave Dictionary of Money and Finance, eds. Newman, P., Milgate, M. and Eatwell,J. London: Macmillan.

Muellbauer, J. (1983). Surprise in the consumption function. *Economic Journal*, 93: 34-49.

Muellbauer, J. (1988). Habits, rationality and myopia in the life-cycle consumption function. Annales d'Economie et de Statistique, 9: 47-70.

Muth, J.F. (1960). Optimal properties of exponentially weighted forecasts. *Journal of the American Statistical Association*, 55: 299-306.

Nelson, C.R. (1987). A reappraisal of recent tests of the permanent income hypothesis. *Journal of Political Economy*, 95: 641-46.

Oterwald-Lenum, M. (1992). Practitioner's corner: a note with quantiles for the asymptotic distribution of the maximum likelihood cointegration rank test statistic, *Oxford Bulletin of Economics and Statistics*, 54: 461-471.

Pantula, S.G., Gonzalez-Farias, G. and Fuller, W.A. (1994). A comparison of unit-root test criteria, *Journal of Business and Economic Statistics*, 12: 449-459.

Patterson, K.D. (1992). The service flow from consumption goods with an application to Friedman's permanent income hypothesis. *Oxford Economic Papers*, 38: 1-30.

Pesaran, H.M. (1992). Saving and consumption behaviour. In The New Palgrave Dictionary of Money and Finance, eds. Newman, P., Milgate, M. and Eatwell,J. London: Macmillan.

Phillips, P.C.B., and Perron, P. (1988). Testing for a unit root in time series regression, *Biometrika*, 75: 335-346.

Quah, D. (1990). Permanent and transitory movements in labour income: an explanation for "excess smoothness" in consumption. *Journal of Political Economy*, 98: 449-475.

Sargent, T.J. (1979). Macroeconomic Theory. New York: Academic Press.

Schwarz, G. (1978). Estimating the dimension of a model. Annals of Statistics, 6: 461-464.

West, K.D. (1988). The insensitivity of consumption to news about income. *Journal of Monetary Economics*, 21: 17-33.

Wickens, R. and Molana, H. (1984) Stochastic life-cycle model with varying interest rates and prices, *Economic Journal*, 94, supplement, pp.133-47

Winder, C. and Palm, F. (1991). Stochastic implications of the life cycle consumption model under rational habit formation. Memo, Limburg University.

Zellner, A. and Geisel, M.S. (1970). Analysis of distributed lag model with application to consumption function estimation. *Econometrica*, 38: 865-888.

Endnotes

[1] Johnson (1971) and Darby (1974) explain the theoretical issues, Muth (1960) and Sargent (1979) discuss the specification of the process for updating permanent income, and Zellner and Geisel (1970) examine the econometric specification of the model with a transitory consumption. See Molana (1992) for further details.

[2] Note that adding a deterministic drift term to the right-hand-side of (7) does not alter the results.

[3] For further aspects see West (1988), Caballero (1990), Campbell and Mankiw (1991), Flavin (1993) and Carroll (1994).

[4] For other justifications in the literature see Attfield *et al*. (1992), Clarida (1991), Caballero (1990), and Quah (1990).

[5] Galí (1991) uses a generalisation of this process and derives restrictions to test the relative smoothness of consumption.

[6] We utilise $E(exp(z))=exp(E(z)+(1/2)Var(z))$.

[7] See Pesaran (1992) and Nelson (1987) for further details.

[8] Note that are $\eta_{t+j}=0$ is not needed for consistency.

[9] For more details and various modelling aspects see Iannaccone (1986), Becker and Murphy (1988), Muellbauer (1988), Constantinides (1990), Loewenstein and Prelec (1992), Heaton (1993), and Dockner and Feichtinger (1993).

[10] Note that the first order conditions in (21) imply a second order difference equation for the marginal utility whose characteristic equation is given by $\mu\rho Z^2 - (1+\mu\rho)Z + 1= 0$. This has two distinct roots $z_1=1$ and $z_2=(1/\mu\rho)>1$. Of these, the only relevant (nonexplosive) root is z_1 which implies the constancy of marginal utility.

[11] Although the understanding of seasonality is pertinent to the subject in general, it does not concern the objective of this paper.

[12] This is a well known problem in the literature and the choice is rather restricted by data availability. Apart from very few exceptions the majority of studies use the same measurements. Therefore the results discussed below will be directly comparable with those reported in the literature. For exceptions see Patterson (1992) and Muellbauer (1983) which generate and use series which approximate, respectively, flow of services from durable goods and disposable labour income.

[13] Using per-capita series also produced the same results as those in Tables 1-6 and therefore, to preserve space, are not reported.

[14] As mentioned above, theory requires that we use disposable labour income X, rather than disposable total income Y. But the use of Y in empirical analysis, while imposed by the lack of appropriate data on X, has also been justified by noting that $Y_t = X_t + r_t A_{t-1}$ where the second term on the right-hand-side is the interest income from financial assets wealth, A, at rate r. When all permanent income is consumed and the real rate of interest r is assumed to remain constant, this term will also be expected to remain constant. Thus, $E_t Y_{t+j} - E_{t-1} Y_{t+j} = E_t X_{t+j} - E_{t-1} X_{t+j}$ since $E_t(rA_{t+j}) = E_{t-1}(rA_{t+j})$ for all $j \geq 0$. Flavin (1981) provides similar explanations of replacing labour disposable income with total disposable income in emprical work.

[15] Note that although the intercept term implied by the theory in (24) is time varying, restricting it to be fixed in the estimation only contributed to heteroscedastic errors in the U.K. case. Accordingly we re-estimated the U.K. specification under various GARCH-M specifications, also allowing the change in inflation to affect the residual variance. However, these estimates did not yield any substantial improvement and are thus not reported.

[16] This restriction is derived from $(1-\mu^2)\mathrm{Var}(\Delta C) = (k\pi)^2 \mathrm{Var}(\varepsilon)$ which is implied by equation (23).