Review of Economic Studies (1989) 56, 357-374 0034-6527/89/00240357$02.00

Why is Consumption So Smooth?

JOHN CAMPBELL
and
ANGUS DEATON
Princeton University and NBER

First version received February 1988; *final version accepted December* 1988 (*Eds.*)

For thirty years it has been accepted that consumption is smooth because permanent income is smoother than measured income. This paper considers the evidence for the contrary position, that permanent income is in fact less smooth than measured income, so that the smoothness of consumption cannot be straightforwardly explained by permanent income theory. The paper argues that in postwar U.S. quarterly data, consumption is smooth because it responds with a lag to changes in income.

INTRODUCTION

One of the most striking features of aggregate consumption behaviour is that aggregate consumption is *smooth* relative to aggregate income. Shifts in aggregate income are associated with relatively small shifts in aggregate consumption, and variations in consumption about trend are smaller than variations in income about trend. The textbook explanation for these facts is that consumption is determined by permanent income not by current income, and permanent income is smooth relative to current income. Innovations to income generate relatively small innovations to permanent income, and thus to consumption. Hence, if the smoothness of consumption relative to income is taken to refer to the relative variance of innovations (a definition we adopt throughout this paper), smoothness is explained by permanent income theory. Indeed, the smoothness of consumption is one of the principal *raisons d'etre* for the permanent income formulation.

However, there is no logical necessity that permanent income must be smoother than current income. Beveridge and Nelson (1981) showed that the innovation variance of permanent income will exceed that of current income if the growth rate of current income is predominantly positively auto-correlated, while Nelson and Plosser (1982) argued that the growth rates of many macroeconomic time series, including real and nominal GDP, follow first-order moving average processes with a positive moving average coefficient. Deaton (1987) drew out the implications of these results for the consumption function, arguing that permanent income is indeed *noisier* than current income so that permanent income theory fails to offer any straightforward and well-supported account of why consumption is smoother than income.

In the present paper, we attempt to do three things. First, in Section 2, we build a bivariate model of saving and labour income that recognizes the fact that consumers are likely to be able to predict their incomes better than would an outsider with access only to the past history of the series. The more information that consumers have, the smoother will be their consumption, so that a test of excess smoothness must be robust to possible econometric misspecification of the consumer's information set. We show that such a test can be constructed around a simple bivariate vector autoregression (VAR), and we find that the puzzle remains, consumption is smoother than it ought to be.[1] Section 3 considers the possibility that the result comes from an insufficiently general specification of the stochastic process that we use to characterize income and saving. We present a non-parametric test of excess smoothness which, although less clearcut, is consistent with the earlier results. In particular, although the standard errors are large enough to permit belief in the conventional view, the data provide remarkably little evidence in its support. If permanent income is really smoother than measured income, there is remarkably little sign of it in the aggregate data.

Our third objective, and a somewhat more positive one, is to offer some account of why it is that consumption is so smooth; if smooth permanent income does not explain the smoothness of consumption, what does? Our VAR results, consistent with the work of Flavin (1981) and Nelson (1987), show a positive correlation between the change in consumption and the lagged change in income, a correlation that would be zero if the permanent income model were true. We show in Section 4 that it is precisely this correlation, the "excess sensitivity" of changes in consumption to anticipated changes in income, that accounts for the fact that consumption is excessively *insensitive* to unanticipated changes in income. The sort of story that this suggests is one in which consumption is slow to adjust to innovations in income, so that changes in consumption are related to averages of previous innovations, thus explaining both the smoothness and the correlation. We believe that the reason consumption is so smooth is because the permanent income theory is false, false in a way that has been widely recognized, although not previously linked to the smoothness of consumption. A belief in slowly-adjusting consumption reconciles all of the evidence and permanent income can be allowed to be less smooth than measured income without contradicting the known smoothness of consumption.

Section 1 is a preliminary one that establishes the context in which the smoothness puzzle is examined. Our formulation of the permanent income model is a standard one, but we introduce a log-linearization of the model that in many respects is easier to handle. 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 logarithmic forms tend to fit the data better. The linearization given here is designed to reap the benefits of logarithms without sacrificing the inherently additive structure of the permanent income model. The first section also gives simple forms of the smoothness result, for both linear and logarithmic formulations, and these provide a convenient starting point for the more sophisticated calculations in the later sections.

1. CONSUMPTION, PERMANENT INCOME, AND INNOVATIONS

Following Flavin (1981), we shall take the following equation as representing the permanent income theory:

$$c_t = \frac{r}{1+r}[A_t + \textstyle\sum_{i=0}^{\infty}(1+r)^{-i}E_t y_{t+i}]. \tag{1}$$

Here c_t is real *per capita* consumption at time t, r is the real rate of interest, assumed to be a constant, A_t is real *per capita* non-human wealth at the end of period t, so that $rA_t/(1+r)$ is capital income, E_t is the expectation operator for expectations formed at t, and y_t is real *per capita* labour income received at time t. Point expectations, a constant

1. West (1988b) uses an alternative methodology to account for the possibility that consumers have superior information. He also reaches the same conclusion.

real interest rate, and the infinite horizon are all adopted to enable us to illustrate as simply as possible the issues with which we are concerned. The evolution of assets over time is governed by

$$A_{t+1}=(1+r)(A_t+y_t-c_t). \tag{2}$$

The first difference of equation (1) can be written, using (2), in the form

$$\Delta c_{t+1}=r\textstyle\sum_{i=1}^{\infty}(1+r)^{-i}(E_{t+1}-E_t)y_{t+i} \tag{3}$$

so that changes in consumption are driven by innovations in labour income. More precisely, in this infinite horizon model, the change in consumption is simply the annuity value of the present discounted value of change in the expected value of future labour incomes. As in more general models of consumption under uncertainty, the change in consumption depends on neither the past history of nor previously anticipated changes in labour income. This of course is the well-known result of Hall (1978).

We shall also make use of an alternative but equivalent expression first derived in Campbell (1987). Saving, s_t, is defined by

$$s_t=rA_t/(1+r)+y_t-c_t=z_t-c_t \tag{4}$$

where $z_t=rA_t/(1+r)+y_t$ is total real income, and is explained by the "saving for a rainy day" equation

$$s_t=-\textstyle\sum_{i=1}^{\infty}(1+r)^{-i}E_t\Delta y_{t+i}. \tag{5}$$

According to the permanent income model, saving is the discounted present value of expected future declines in labour income. (The equivalence of (5) and (1) can readily be seen by using (4) to substitute for s_t, and then "unscrambling" the changes in labour income.)

It is standard practice in the consumption literature to combine the linear equations (1) through (5) with a linear time-series model for current labour income y_t. To take a simple example that illustrates the algebra as well as the phenomenon with which we shall be concerned, the following simple autoregression in first differences gives a good fit to the labour income series. Our data period runs from 1953(2) to 1984(4), so that with differencing and lags there are 125 observations: by OLS

$$\underset{(3\cdot2)}{\Delta y_t=8\cdot2}+\underset{(5\cdot5)}{0\cdot442}\,\Delta y_{t-1}+\varepsilon_t, \qquad \sigma_\varepsilon=25\cdot2. \tag{6}$$

Flavin (1981) and Hansen and Sargent (1981*a*) show that if y_t follows an *AR* of the form $A(L)y_t=a+\varepsilon_t$, equation (3) implies that $A\{1/(1+r)\}\Delta c_t=r\varepsilon_t/(1+r)$, so that in this case, with $A(L)=1-1\cdot442L+0\cdot442L^2$, the formula gives

$$\Delta c_t=\frac{(1+r)}{0\cdot558+r}\,\varepsilon_t. \tag{7}$$

The multiplier on the right-hand side of (7) is 1·79 when r is zero, and decreases only slowly with r, for example to 1·76 when r is 10% per annum. Thus (7) predicts that the standard deviation of changes in consumption should be at least 1·76 times larger than the standard deviation of the innovation of labour income, estimated by (6) to be 25·2 ($1972 *per capita* per annum). In fact, the standard deviation of the change in consumption is 27·3, and even this is an overestimate since *purchases* rather than consumption of durables are included in the total. The standard deviation of changes in consumption of non-durables and services is 12·4, and scaling this by the ratio of the mean of total consumption to the mean of consumption of non-durables and services gives a figure of only 15·8.[2]

In spite of the fact that autoregressions like (6) fit the data well, there must be some doubt as to whether the change in labour income is a stationary time series. Low growth rates towards the end of the period prevent there being a significant upward trend in the size of income changes, but there has certainly been an increase in variability over time. We also have some prejudice in favour of proportional over linear growth, at least in the long-run. In consequence, we find it more reasonable to suppose that the first difference of the logarithm of labour income is stationary. However, the permanent income model relates consumption and changes in consumption to the level and changes in the level of labour income, so that some reformulation is needed in order to work in logarithms. The basic idea is to work with the *ratios* of saving and consumption to labour income, and to relate them to expectations about the ratios of future to current income. Since we assume that the rate of growth of labour income is stationary with mean μ, say, expressions of the form $E_t(y_{t+j}/y_t)$ are readily decomposed into an expected growth component $\exp(j\mu)$, and a residual. And because this residual is likely to be small relative to the growth component, it is possible to adopt a convenient linearization that yields a logarithmic version of the model. The details are confined to a brief Appendix; here we report the loglinear forms for the key equations (3) and (5), and present the equations to be used in the empirical work.[3]

The "rainy day" equation, (5), in which saving anticipates future declines in labour income, has a very similar form in logarithms, in which the saving *ratio* anticipates future *logarithmic* declines in income, viz.,

$$(s_t/y_t) \approx -\textstyle\sum_{i=1}^{\infty} \rho^i E_t \Delta \log y_{t+i} - \kappa \tag{8}$$

where κ is a constant given in the Appendix. The discount factor ρ in (8) is not $1/(1+r)$, but $(1+\mu)/(1+r)$, or approximately $1+\mu-r$. One of the costs of moving from linear to logarithmic specifications is the need to assume that $r>\mu$, that the real interest rate is larger than the rate of growth of real labour income. The discounted present value of linear growth will exist provided only that the discount rate is positive, but this is clearly not the case for proportional growth. We believe that the stronger assumption is warranted by the increase in realism of the logarithmic model for incomes.

Changes in consumption can also be related to changes in expectations about rates of income growth. The logarithmic counterpart to equation (3) takes the form

$$\frac{\Delta c_{t+1}}{y_t} \approx \frac{r}{r-\mu} \textstyle\sum_{i=1}^{\infty} \rho^i (E_{t+1}-E_t) \Delta \log y_{t+i} \tag{9}$$

so that the ratio of the change in consumption to labour income is proportional to the change in the present value of future rates of growth of income, where once again, the discount rate is the excess of the real interest rate over the rate of growth.

2. These calculations make no allowance for the presence of transitory consumption. But adding transitory consumption to the model only serves to deepen the smoothness puzzle, not to resolve it. If we add a white noise error, "transitory consumption", to the original model (1), we obtain moving average errors in equation (3). But if, as seems plausible, this error is uncorrelated with the innovation to the income process, it makes changes in consumption *noisier* than justified by changes in permanent income alone. Much the same can be said for unanticipated capital gains; if equation (2) is modified to include an asset shock, the effect is to add the shock, multiplied by the rate of interest, to the change in consumption in equation (3). The result, once again, is to increase the predicted variability of changes in consumption, except in the unlikely event that unanticipated capital gains are negatively correlated with innovations in income.

3. Note that, as illustrated by the calculations in this section, very similar results are obtained for these data whether one works in levels or in logarithms. Nelson (1987) also compares logarithmic and level specifications.

Equations (8) and (9) make somewhat different approximations, and since we shall use both in the analysis, it is important that they be reconciled. The derivation of equation (8) requires that the saving ratio be small, and in the Appendix we show that this will only be true if both r and μ/r are small. If so, $r/(r-\mu)$ in (9) is approximately equal to unity, while if the lagged value of (8) is divided by ρ and subtracted from (8), we have

$$\frac{s_t}{y_t}-\Delta\log y_t-\frac{s_{t-1}}{\rho y_{t-1}}\approx-\sum_{i=0}^{\infty}\rho^i(E_t-E_{t-1})\Delta\log y_{t+i}\approx-\frac{\Delta c_t}{y_{t-1}}. \tag{10}$$

The approximate equivalence of the first and last expressions may be checked directly, and is satisfied in our data. Our concern will be to test whether either of the outermost expressions is equal to that on the inside.

The simple autoregressive scheme that works for first differences in (6) also works for rates of growth of labour income, viz., and again by OLS, with measurements in percentage points at an annual rate,

$$\underset{(3\cdot3)}{\Delta\log y_t=1\cdot052}+\underset{(5\cdot5)}{0\cdot443}\,\Delta\log y_{t-1}+\varepsilon_t,\qquad \sigma_\varepsilon=3\cdot138 \tag{11}$$

so that we have

$$\sum_{i=0}^{\infty}\rho^i(E_t-E_{t-1})\Delta\log y_{t+i}=\sum_{i=0}^{\infty}(0\cdot443\rho)^i\varepsilon_t=\frac{\varepsilon_t}{1-0\cdot443(1+\mu-r)}. \tag{12}$$

The sample average quarterly rate of growth of labour income, μ, is 1·805% per annum. It follows that the coefficient multiplying ε_t on the right-hand side of (12) is 1·80 when $r=0$ and, for example, 1·77 when r is 10% per annum. Hence we have "excess smoothness" in consumption if the standard deviation of either of the two outermost expressions in (10) is less than 1·77 times 3·138 (the estimate of σ_ε in (11)), which is 5·55% per annum. In fact, the standard deviation of the ratio of the change in consumption of non-durables and services to lagged income is 3·27% per annum, while that of the expression involving saving ratios is 3·57% per annum.

Table I gives these and other results for a variety of different data series. The first panel shows means and standard deviations for quarterly growth rates of both total (i.e. inclusive of capital) disposable income z_t, and disposable labour income y_t. The mean and standard deviation of the present value of innovations in labour income are calculated as above, in percentage points at an annual rate. The second panel gives rates of growth of total consumption, and the two measures of consumption change. Apart from sign, the two are close to one another, and both have standard deviations much smaller than predicted from the AR (1) behaviour of the changes in log income.

The final panel calculates similar statistics for consumption of nondurables and services, which may be a better measure of true consumption than is total consumption expenditure; the narrower measure grows somewhat less rapidly than the total, and is very much smoother. Two scale factors are considered. The first, 1·274, is the ratio of the mean of total consumption to the mean of consumption excluding durables. Once again, the two change measures are close in both mean and variance, and the variances are much less than those predicted by the theory. The second scale factor, 1·495, comes from Campbell (1987), and is the reciprocal of the marginal propensity to consume estimated from a simple bivariate regression of consumption of non-durables and services on total income. This "cointegrating" factor is the number λ that will make stationary the saving series $z_t-\lambda c_t$. Under this definition of consumption, the two approximations diverge significantly, particularly in means. One of the difficulties here is that purchases

TABLE I

Means, actual and predicted standard deviations

Variable	Mean (% per annum)	Standard deviation (% p.a.)	Scale factor λ
$\Delta \log z_t$	2·155	3·293	
$\Delta \log y_t$	1·805	3·516	
$\sum \rho^i \Delta E_t \Delta \log y_t$	0·000	5·587	
	Total consumption		
$\Delta \log c_t$	2·204	3·141	
$\Delta c_t / y_{t-1}$	2·307	3·272	
$\Delta \eta_t$	−2·155	3·574	
	Scaled consumption of non-durables and services		
$\Delta \log c_{nt}$	1·991	1·950	—
$\lambda \Delta c_{nt} / y_{t-1}$	2·073	2·019	1·274
$\Delta \eta_t$	−1·923	2·469	1·274
$\lambda \Delta c_{nt} / y_{t-1}$	2·433	2·370	1·495
$\Delta \eta_t$	−1·194	2·488	1·495

Notes. z_t is total real disposable income *per capita* and y_t is real disposable labour income *per capita*, as constructed by Blinder and Deaton (1985). In constructing the parameter ρ, we assume a constant discount rate of 6% per annum. The calculations in the third line assume that $\Delta \log y_t$ follows an AR (1) process with coefficient $\phi = 0{\cdot}443$. $\Delta \eta_t$ is shorthand for $(s_t/y_t) - \Delta \log y_t - \rho^{-1}(s_{t-1}/y_{t-1})$. The data run from 1953(1) to 1985(4).

of durable goods are becoming a steadily larger share of total consumption, so simple scaling is not an adequate substitute for calculating the consumption of durables.

Even so, these results clearly show that the excess smoothness result can be described in logarithms just as well as in levels. If the simple AR (1) model of the growth of labour income adequately captures the way in which consumers forecast their future income, then permanent income is much noisier than current income and any measure of consumption.[4]

2. SUPERIOR INFORMATION AND BIVARIATE MODELS

The illustrative calculations of the previous section are subject to an important objection. They all assume that income is appropriately modelled, by consumers and in our statistical procedures, as a *univariate* stochastic process. In terms of equations (8) and (10), which define our log-linear version of the permanent income model, we replaced expectations that should have been conditional on *consumers'* information with expectations that were conditional on current and lagged values of income. It seems likely that consumers in fact form their expectations using a richer information set than just the history of income itself. Consumers have information about monetary policy, asset prices, growth of specific sectors of the economy, and so on. It is most implausible that none of this information is relevant for forecasting income.

If consumers do have extra information, the effect is to smooth permanent income relative to the "permanent income" measure calculated on the basis of univariate forecasts. In the extreme case of perfect foresight, this result is obvious; permanent income and

4. We note that these results are quite robust to splitting the full sample period 1953/1 to 1985/4. We split the sample at the end of 1963, and every quarter thereafter to the end of 1973. The subsample first-order autoregression coefficients never fell below 0·38 and are frequently larger than those estimated over the full sample. Chow tests never reject the hypothesis that the AR (1) model is stable across subsamples, although in the later years the mean growth rate does seem to be somewhat lower and the innovation variance larger.

consumption will be constant even though an analysis of measured income alone might seem to warrant considerable variability in permanent income. More generally, it is useful to think of superior information as meaning that consumers receive news about (univariate) innovations sometime before they occur. Permanent income will still be revised, but because the changes are some way in the future, they are discounted and so affect permanent income by less than would be the case if they were to affect income at once. A formal proof of the smoothing effect of superior information can be found in West (1988*a*).

Given these results, it could be argued that consumption is smooth, not because there is anything wrong with the permanent income theory, but because consumers have sufficient information to make permanent income smooth despite the fact that the univariate analysis of income suggests that it is not. At first sight, it would seem very hard to rule out such a possibility. However many forecasting variables we include in our econometric model, we can never be sure that we are capturing all the relevant information that is used by consumers. Instead of trying to do so, the solution is to use consumers' own behaviour to reveal their expectations. In particular, if the permanent income theory is true, consumers reveal their estimates of permanent income through their consumption and savings decisions. Campbell (1987) shows how this fact can be exploited in testing the permanent income theory; here we apply the same method to examine the smoothness question.

To see how the approach works, start from the basic log-linear model (8), with saving defined as in equation (4) as the difference between total income and consumption. Rewrite (8) as

$$(s_t/y_t) \approx -\textstyle\sum_{i=1}^{\infty} \rho^i E\,[\Delta \log y_{t+i} \,|\, I_t] - \kappa \tag{13}$$

where I_t is the consumers' information set. I_t is, of course, unknown, but define the information set H_t consisting of all current and lagged values of s_t/y_t and $\Delta \log y_t$. This information set, while generally smaller than I_t, is larger than that of the previous section because, not only does it include income and its past history, but also the history of s_t/y_t, a variable which summarizes consumers' information about future incomes. Take expectations of equation (13) conditional on H_t; this gives

$$(s_t/y_t) \approx -\textstyle\sum_{i=1}^{\infty} \rho^i E[\Delta \log y_{t+i} \,|\, H_t] - \kappa. \tag{14}$$

Because s_t/y_t is in H_t, the left-hand sides of (13) and (14) are identical, while, since H_t is contained in I_t, the right-hand side is the same apart from the change of conditioning set. Subtracting $s_{t-1}/\rho y_{t-1}$ and $\Delta \log y_t$ from (14) (note that both variables belong to H_t), we obtain a version of the original equation (10),

$$\frac{s_t}{y_t} - \Delta \log y_t - \frac{s_{t-1}}{\rho y_{t-1}} \approx -\textstyle\sum_{i=0}^{\infty} \rho^i [E(\Delta \log y_{t+i} \,|\, H_t) - E(\Delta \log y_{t+i} \,|\, H_{t-1})]. \tag{15}$$

Once again, the left-hand sides of (10) and (15) are identical while the right-hand side contains the innovations with respect to H_t rather than the consumers' information set I_t.

Provided that expectations can be approximated by linear projections, equations (14) and (15) suggest a simple way to test the permanent income theory even when we do not know what information set is used by consumers. We can estimate a VAR system containing income growth ($\Delta \log y_t$) and the savings ratio (s_t/y_t). The VAR forecasts of future income growth can be weighted together to form the right-hand side of (14), which can then be compared with the left-hand side; this was the strategy in Campbell (1987). Alternatively, we can weight together the *revisions* in VAR forecasts on the right-hand side of (15), and compare the result with the savings innovation on the left-hand side. If the standard deviation of the left-hand side of (15) is less than the standard deviation of the right-hand side, calculated from the VAR, then we can say that consumption is smoother than permanent income. Such a result will be valid whether or not consumers have superior information.

The basic VAR model that we consider is

$$\begin{pmatrix} \Delta \log y_t - \mu \\ s_t/y_t - \sigma \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} \Delta \log y_{t-1} - \mu \\ s_{t-1}/y_{t-1} - \sigma \end{pmatrix} + \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix} \tag{16}$$

where σ is the mean saving ratio. In matrix notation, (16) is

$$x_t = Ax_{t-1} + u_t. \tag{17}$$

Most of our analysis can be done with the simple first-order vector autoregression (16), but we shall occasionally incorporate additional lags. (More general specifications altogether are considered in Section 3.) If there are M lags, x_t is a $2M$ vector containing current $\Delta \log y_t$, its first $(M-1)$ lags, followed by the equivalent M terms for the saving ratio. The matrix A is then a $2M$ square matrix of coefficients. Equation (16) asserts that the saving ratio and the change in income follow a jointly stationary process with feedbacks both from lagged income growth to saving, and from lagged saving to income growth. The existence of the "cross" effects not only allows for the information structure, but is also consistent with rather general univariate time-series representations of the two series.

The first-order vector autoregression (16) generates parameter estimates that are shown in the first panel of Table II. The first order auto-regression in $\Delta \log y_t$ from equation (11) appears again here in the growth rate, although there is also a small but significant negative feedback from the lagged saving ratio to changes in income. Such a term is predicted by the "rainy-day" equation and by the existence of consumers' superior information; consumers with early notice of income changes will signal them in advance through their saving behaviour. The saving rate is also well described by an AR(1), especially when consumption is total consumption; when non-durable and service consumption is used, the "own" autoregressive parameter is close to unity, and there is a much larger feedback from lagged income changes.

TABLE II

Bivariate representations of income growth and savings

	(1)		(2)		(3)	
			VAR parameters			
a_{11}, a_{12}	0·443	−0·177	0·454	−0·067	0·458	−0·066
	(0·08)	(0·06)	(0·08)	(0·03)	(0·08)	(0·03)
a_{21}, a_{22}	0·072	0·803	0·212	0·972	0·266	0·956
	(0·07)	(0·05)	(0·07)	(0·03)	(0·08)	(0·03)
			ARMA parameters and AR roots			
ϕ_1, ϕ_2 (AR)	1·247	−0·369	1·425	−0·455	1·414	−0·455
	(0·09)	(0·07)	(0·08)	(0·08)	(0·08)	(0·03)
θ_1 (MA)	−0·799	(0·05)	−0·954	(0·02)	−0·952	(0·02)
r_1, r_2 (roots)	0·763	0·483	0·943	0·482	0·918	0·496

Notes. (1) refers to calculations in which consumption is taken to be total consumption, including purchases of durables. Panel (2) uses only nondurables and services consumption inflated by a factor of 1·274. Panel (3) is the same as panel (2) but with an inflation factor of 1·495.

Univariate representations for both series can readily be derived using standard techniques, see e.g. Granger and Newbold (1977, p. 217). Both series have ARMA (2, 1) representations, and the parameters are shown for the income growth rate process; because of the feedbacks between income and the saving ratio, the representation varies with the definition of consumption. However, all three representations have an autoregressive root that is between 0·48 and 0·50, while the second autoregressive root is always very close to cancelling with the moving average root. In consequence all three ARMA (2, 1) processes are very close to being AR (1) processes with a positive autoregressive parameter of 0·48, and this is in accord with the simple model (11) from which we started.

We consider next the relationship between the estimated vector auto-regression and the theoretical properties of income, saving and consumption. The version of the permanent income model we use is that given by equation (15). In the notation of the VAR (17), this can be rewritten in the convenient form

$$\frac{s_t}{y_t} - \Delta \log y_t - \frac{s_{t-1}}{\rho y_{t-1}} = (e_2' - e_1')x_t - \rho^{-1}e_2'x_{t-1} \tag{18}$$

where e_1' and e_2' are the vectors (1, 0) and (0, 1) respectively, so that $e_1'x_t = \Delta \log y_t$ and $e_2'x_t = s_t/y_t$. (Note that (18) also holds when the VAR contains more lags, provided e_1 and e_2 are appropriately redefined.) Substitution of the VAR (17) into the right-hand side gives

$$\frac{s_t}{y_t} - \Delta \log y_t - \frac{s_{t-1}}{\rho y_{t-1}} = [(e_2' - e_1')A - \rho^{-1}e_2']x_{t-1} + (e_2' - e_1')u_t. \tag{19}$$

From (17), and writing $E_t(\cdot)$ for $E(\cdot \mid H_t)$, it is clear that $(E_t - E_{t-1})x_{t+i} = A^i u_t$, so that, since $e_1'x_t = \Delta \log y_t$,

$$\textstyle\sum_{i=0}^{\infty} \rho^i (E_t - E_{t-1}) \Delta \log y_{t+i} = \sum_{i=0}^{\infty} e_1' \rho^i A^i u_t. \tag{20}$$

Now, the permanent income theory (15) implies that the left-hand side of (19) should be equal to minus the left-hand side of (20). This requires that two sets of constraints be satisfied, viz.,

$$(e_2' - e_1')A - \rho^{-1}e_2' = 0, \tag{21a}$$

$$-\textstyle\sum_{i=0}^{\infty} e_1' \rho^i A^i = e_2' - e_1'. \tag{21b}$$

The interpretation of these two sets of restrictions is an important part of our story. Equation (21a) guarantees that the left-hand side of (19) is independent of either lagged income growth or the lagged saving ratio. It is therefore a test of the unpredictability of consumption, and includes a test of the absence of "excess sensitivity", of the lack of relationship between changes in consumption and lagged values of income. By contrast, the restrictions in (21b) ensure that the change in consumption is that which is warranted by the change in the present value of income. If (21b) is satisfied, there can be no "excess smoothness" of consumption.

The orthogonality condition (21a) implies that the A matrix can be written in the form

$$A = \begin{pmatrix} \alpha & \beta \\ \alpha & \beta + \rho^{-1} \end{pmatrix} \tag{22}$$

where α and β are not restricted. Since ρ is non-zero, the determinant of $(I - \rho A)$ will be non-zero if, and only if $\beta \neq 0$, i.e. if saving Granger-causes income. Such Granger-causality is guaranteed if consumers have superior information, and is supported by the empirical evidence.[5] Given $\beta \neq 0$, the inverse $(I-\rho A)^{-1}$ exists, so that (21a) can be rearranged to yield

$$-e_1'(I-\rho A)^{-1} = e_2' - e_1' \tag{23}$$

which is identical to (21b). Hence, provided that the lagged saving ratio has predictive power for the change in labour income, the orthogonality condition and the condition for smoothness are identical. If consumption changes cannot be predicted by past changes in either consumption or income, then the consumption change must be equal to the revision in permanent income. Intuitively, this equality follows from the intertemporal budget constraint; the only random walk for consumption that satisfies the intertemporal budget constraint is the one whose innovations are equal to the innovations in permanent income.

Table III presents the relevant data, again for the three alternative definitions of consumption. The first column presents standard tests of orthogonality using VAR's with one lag of the two variables, and VAR's with five lags of each. All of these test statistics show rejections, with the possible exception of the five-lag VAR for the larger scaling of consumption of non-durables and services. Even allowing for the asymptotic nature of these results, they are not favourable for the hypothesis and are in line with other findings using similar (or the same) data.

TABLE III

Tests for excess smoothness and excess sensitivity

	Wald test (p-value)	Predicted innovation standard deviation	Actual innovation standard deviation	Ratio (s.e.)
		Total consumption		
VAR-1	19·6 ($5{\cdot}4\times10^{-5}$)	5·044	3·309	0·656 (0·10)
VAR-5	45·9 ($1{\cdot}5\times10^{-6}$)	4·768	3·017	0·633 (0·15)
		Non-durables and services consumption ×1·274		
VAR-1	18·4 ($1{\cdot}0\times10^{-4}$)	3·864	2·290	0·593 (0·21)
VAR-5	31·4 ($5{\cdot}0\times10^{-4}$)	4·686	2·135	0·456 (0·20)
		Non-durables and services consumption ×1·495		
VAR-1	10.0 (0·0066)	3·188	2·382	0·747 (0·16)
VAR-5	18·2 (0·0509)	4·022	2·257	0·561 (0·13)

Notes. The Wald test is the test of the estimated parameters in the VAR for conformity with the restrictions (21a). The predicted innovation is the standard deviation of the last term in (20), i.e. the square root of the quadratic form $e_1'(I-\rho A)^{-1}\Omega(I-\rho A')^{-1}e_1$ where Ω and A are estimated from the unrestricted VAR. The actual innovation is the standard deviation of $u_{2t}-u_{1t}$, or the square root of $(e_2'-e_1')\Omega(e_2-e_1)$. The standard error of the ratio of these quantities is computed as $\sqrt{(D'\theta D)}$ 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 numerically. All calculations use a ρ of 0·9895, which corresponds to a real interest rate of 6% per annum.

The remainder of the table compares theoretical and actual innovation variances, or standard deviations. For each case, we calculated the theoretical innovation variance of $(s_t/y_t - \Delta \log y_t)$, which is (approximately) equal to the theoretical innovation in $\Delta c_t/y_{t-1}$, and is calculated from the last term in (20). For example, in the one-lag case, the VAR's are given in Table II, the variance-covariance matrices of the residuals are calculated in the usual way, and a value for the real interest rate of 6% per annum is assumed. The

5. See Campbell (1987) for technical discussion of the case where $\beta=0$.

actual innovation in $(s_t/y_t - \Delta \log y_t)$ is, by (19), simply $u_{2t} - u_{1t}$, the standard deviation of which requires only the variance-covariance matrix. In every case, the theoretical innovation variance is larger than the actual innovation variance, and in all but one case, is more than twice as large. Consumption is markedly smoother than it ought to be if the permanent income theory were correct.

3. A NON-PARAMETRIC TEST FOR EXCESS SMOOTHNESS

It is possible that the results in the previous section are specific to the simple VAR specification, and that a more general model would give different results. While such criticisms can always be made, there is a specific cause for concern here. It is possible that the growth of labour income, while behaving very much like an AR(1) in the short-run, may have quite different long-run properties. In particular, the concern is that there may be a deterministic trend in labour income to which the series slowly reverts. A first- or fifth-order VAR might miss such slow trend reversion, but if consumers are aware of it, consumption will not respond much to innovations in income, and a test based on the VAR's will give a misleading impression of excess smoothness. In this section, we show how to construct a non-parametric test of the model. Unfortunately, nothing comes for free, and we shall have to make additional assumptions in order to relax those on functional forms, and our data will be stretched to the limit to provide results.

Instead of the VAR (16) and (17), we suppose that the x_t process has a possibly infinite-order moving average representation given by

$$x_t = \varepsilon_t + B_1\varepsilon_{t-1} + B_2\varepsilon_{t-2} + \cdots + B_r\varepsilon_{t-r} + \cdots. \tag{24}$$

From this, we wish to be able to calculate the discounted present value of innovations in income at time t, so that we need the counterpart of (20), which becomes

$$\textstyle\sum_{i=0}^{\infty} \rho^i (E_t - E_{t-1})\Delta \log y_{t+i} = \sum_{i=0}^{\infty} e_1'\rho^i B_i \varepsilon_t. \tag{25}$$

In order to construct a non-parametric estimator of (25) we have to assume that the discount factor ρ is unity. Such an assumption is not desirable, since the log-linear permanent income model is not well-defined when $\rho = 1$. However, note that the right-hand side of (25) is the discounted present value of the innovations to income, and that this sum will typically not be very sensitive to reasonable variation in ρ. One case in which the result would be sensitive to the precise value of ρ is when the time-series process exhibits trend reversion only in the very long run, while the first difference is positively autocorrelated in the short run. In this case, however, our assumption that $\rho = 1$ will tend to underestimate the eventual effects of innovations in measured income, since the long-run trend reversion terms are not discounted.

From (25), with $\rho = 1$, we can write

$$\operatorname{var}\{\textstyle\sum_{i=0}^{\infty} (E_t - E_{t-1})\Delta \log y_{t+i}\} = e_1' \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} B_j \Omega B_k' e_1 \tag{26}$$

where Ω is the variance-covariance matrix of the vector ε. Define C_i as the covariance between x_t and x_{t-i}, so that we have

$$C_i = E(x_t x_{t-i}') = C_{-i}' = \textstyle\sum_{j=0}^{\infty} B_{i+j}\Omega B_j' \tag{27}$$

where the last expression comes from substitution from (24) and taking expectations. The final step is to calculate the doubly infinite sum of the C_i matrices; after some tedious but elementary algebra, we reach

$$\textstyle\sum_{i=-\infty}^{\infty} C_i = \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} B_j \Omega B_k'. \tag{28}$$

Hence, comparing (26) and (28), we have an expression for the variance of innovations in permanent income in terms of the variances and covariances of x_t,

$$\operatorname{var}\{\textstyle\sum_{i=0}^{\infty}(E_t - E_{t-1})\Delta \log y_{t+i}\} = e_1' \textstyle\sum_{i=-\infty}^{\infty} C_i e_1. \tag{29}$$

Note that the expression on the right involves only the autocovariances of the first element of x_t, i.e. the growth rate of income, so that (29) can be written

$$\operatorname{var}\{\textstyle\sum_{i=0}^{\infty}(E_t - E_{t-1})\Delta \log y_{t+i}\} = \sigma^2 \textstyle\sum_{i=-\infty}^{\infty} \gamma_i = \sigma^2 V, \quad \text{say}, \tag{30}$$

where σ^2 is the variance of $\Delta \log y_t$, and γ_i is its ith autocorrelation coefficient. This equation tells us that the quantity V is exactly what we want to know; if $V > 1$, changes in permanent income are noiser than changes in measured income, while if $V < 1$, the conventional wisdom holds good, and permanent income is relatively smooth.

The quantity V has been proposed by Cochrane (1988) as a measure of the *persistence* of the fluctuations in a single series. If measured income follows a univariate process, the persistence of its fluctuations is precisely what governs the relation between the variability of measured income and the variability of permanent income. The result in equation (30) shows that we can also use V even when income is part of a multivariate process which consumers use for forecasting. The assumption that $\rho = 1$ is important for this result. Additional information beyond the past history of income does not further smooth consumption when there is no discounting, so that, in this case, the univariate and multivariate measures of persistence are identical.

From the point of view of this paper, the main virtue of the measure V is that there exists a direct non-parametric estimator. Technically, V is the normalized spectral density at zero, so that there exists a literature on how it may be estimated, see for example Cochrane (1988), Campbell and Mankiw (1987*a*, 1989), and the excellent account in Priestley (1982, pp. 432–471). If we write the sample autocorrelations $\hat{\gamma}_j$ as

$$\hat{\gamma}_j = \frac{T}{T-j} \textstyle\sum_{t=j+1}^{T} (\Delta \log y_t - \hat{\mu})(\Delta \log y_{t-j} - \hat{\mu}) / \textstyle\sum_{t=1}^{T} (\Delta \log y_t - \hat{\mu})^2, \tag{31}$$

where $\hat{\mu}$ is the sample mean of $\Delta \log y_t$, then the estimate of V based on a triangular (Bartlett) window of size k is

$$\hat{V}^k = 1 + 2 \textstyle\sum_{j=1}^{k} [1 - j/(k+1)] \hat{\gamma}_j. \tag{32}$$

The window weights in (32) give linearly-declining weights to higher-order autocorrelations up to and including the kth. Provided this window size is increased with the sample size, $\hat{V}^k$ is a consistent estimator of V, and has an asymptotic t-value given by[6]

$$t(\hat{V}^k) = \sqrt{[0{\cdot}75T/(k+1)]}. \tag{33}$$

The choice of window size k is obviously important in practice. Including too few autocorrelations may obscure trend reversion manifested in higher autocorrelations. Including too many will tend to find excessive trend reversion, since as k approaches the sample size T, the estimate approaches zero in *any* data set. This occurs because the sample mean is removed from (31), which biases downward the sample autocorrelations $\hat{\gamma}_j$. Hence, while large k is desirable, k must be small relative to T.[7]

6. Strictly speaking, this standard error is correct only if we know the variance of the process, σ^2.

7. Cochrane (1988) proposes a bias correction for $\hat{V}^k$; see also Campbell and Mankiw (1989). We use uncorrected $\hat{V}^k$ estimates here, noting that bias correction would increase the estimates in Table IV by a factor of approximately $T/(T-k)$.

Table IV presents the estimates of the spectral density and the implied persistence measures. The autocorrelations are shown for lag lengths up to ten quarters, and it is interesting to note that after the first four, which, as predicted by the simple AR (1), are positive and geometrically declining, there is a sequence of *negative* coefficients, which are not predicted by the AR (1), and which may indicate some tendency for the effects of the innovations to be negated some considerable time after they have occurred. Nevertheless, the estimate of V is consistently above 2 for all window sizes up to nearly 50; the subsequent decline is difficult to interpret because of the mechanical and meaningless decline in the estimate as the window size approaches the sample size. The standard error also increases with the window size, reaching uncomfortably large values when the window size is only 40. The outcome therefore depends a good deal on one's prior views. If we believe that all the relevant trend reversion should show up in the first 10 or 20 autocorrelations, then the results provide quite convincing evidence that permanent income is relatively noisy. By contrast, if we suspect that trend reversion is very slow indeed, and may not show up for fifteen or twenty years after the original shock, then there are not sufficient data to make a judgment; the estimates do not suggest trend reversion but the standard errors are large. Even so, we wish to emphasize as strongly as we can that the conventional explanation for the smoothness of consumption requires that V be less than unity. For a proposition that was taken as self-evident for three decades, Table IV provides remarkably little supporting evidence.

TABLE IV

Non-parametric estimates of persistence

Labour income, $\Delta \log y_t$

Autocorrelations at lag:

1	2	3	4	5	6	7	8	9	10
0·446	0·193	0·156	0·056	−0·152	−0·017	−0·010	−0·070	−0·024	0·123

Persistence estimates:

Window size	$\hat{V}^k$	s.e. ($\hat{V}^k$)	$\tilde{V}^k$
10	2·212	0·755	2·401
20	2·334	1·100	2·132
30	2·523	1·445	2·883
40	2·443	1·609	0·827
50	1·986	1·459	−0·530
60	1·424	1·144	−2·302

Total income, $\Delta \log z_t$

Autocorrelations at lag:

1	2	3	4	5	6	7	8	9	10
0·373	0·124	0·103	−0·039	−0·206	−0·074	−0·068	−0·120	−0·031	0·141

Persistence estimates:

Window size	$\hat{V}^k$	s.e. ($\hat{V}^k$)	$\tilde{V}^k$
10	1·589	0·542	1·406
20	1·518	0·716	0·906
30	1·529	0·871	1·291
40	1·438	0·947	−0·068
50	1·088	0·800	−0·642
60	0·793	0·637	−1·039

A number of possible objections to these results should be dealt with. First, we computed the first ten autocorrelations of disposable income growth in sub-periods 1953(3)-1969(1), 1969(2)-1984(4), 1953(3)-1972(4), and 1973(1)-1984(4). While there are some differences in individual autocorrelations, the overall pattern as well as the persistence estimates for window size 10 are very similar to those for the full sample period. Second, since the series for labour income is a "manufactured" series, see Blinder and Deaton (1985) for details, we also present similar estimates for the "official" series on total income, even though this is not the concept that we need. The pattern is somewhat similar, and although the persistence estimates are smaller there is certainly no suggestion that the results are extremely sensitive to the definition of income.

Third, it has been argued that the Bartlett window in (31), by down-weighting higher-order autocorrelations, which is where trend reversion would show up, gives estimates that are biased upwards, i.e. towards finding persistence. Note however that this effect is offset by the downward bias in sample autocorrelations discussed above. Even so, we calculated estimates $\tilde{V}^k$, using the formula (31) but without the declining weights, so that each autocorrelation has a unit weight up to the kth, but zero thereafter. This estimator, the "truncated window" estimator, is also shown in Table IV; standard errors can be calculated by multiplying the standard errors of $\hat{V}^k$ by $\sqrt{3}$ (which is one of the reasons why the Bartlett window is usually preferred). At least for the labour income series, these new estimates simply reconfirm the old. There is no evidence of trend reversion in the first 30 autocorrelations; beyond that it is impossible to tell because the estimates have very large standard errors.

We provide a fourth and final check on our results by some Monte Carlo experiments reported in Table V. The first panel reports Bartlett and truncated window estimates of V^k for various window sizes corresponding to a model in which the growth rate of income is an AR (1) with a parameter of 0·45. The right panel repeats the exercise for a trend-reverting series, the *levels* of which are generated by a (just) stationary AR (2) around a deterministic trend. This process, with parameters 1·44 and −0·45, was chosen because it accurately reflects the low-order autocorrelations of the measured income series. Note first that the estimates for the AR (1) model are biased downward and that the bias worsens as the window size is enlarged; the theoretical values of V^k increase from 2·37 when $k=10$ to 2·49 at $k=20$ and eventually to 2·59 at 60. Similarly, the

TABLE V

Experimental results on persistence estimates

	Truth is AR (1) in the first differences		Truth is AR (2) with deterministic trend	
Window	$\hat{V}^k$ (s.d.)	$\tilde{V}^k$ (s.d.)	$\hat{V}^k$ (s.d.)	$\tilde{V}^k$ (s.d.)
10	2·19 (0·68)	2·30 (1·28)	2·05 (0·59)	1·97 (1·08)
20	2·16 (1·04)	2·10 (1·88)	1·84 (0·79)	1·28 (1·33)
30	2·04 (1·26)	1·56 (2·16)	1·55 (0·84)	0·72 (1·43)
40	1·85 (1·39)	1·01 (2·30)	1·30 (0·84)	0·37 (1·53)
50	1·65 (1·45)	0·68 (2·25)	1·08 (0·81)	0·11 (1·51)
60	1·45 (1·44)	0·22 (2·08)	0·90 (0·76)	−0·17 (1·56)

Notes. The AR (1) model has a parameter of 0·45, while the AR (2) parameters are 1·44 and −0·45. The results are based on 500 replications using generated series of length 130, taken as the last 130 observations of a series of size 650. True V's are 2·64 for the AR (1) model and zero for the AR (2) model. Standard deviations are standard deviations of the *empirical* distribution of the estimates.

estimates for the AR (2) with deterministic trend are too large; the long-run return to trend is only captured at window sizes that are larger than the sample size can bear. Even so, the difference between the left and right panels is apparent, and is in the right direction; at all window sizes the AR (1) is estimated to be more persistent than the AR (2) in the levels, and the difference becomes quite marked when the window size is 20 or more.

Comparing Table V with our "real" results in Table IV, labour income generates even larger estimates of persistence than does the theoretical AR (1), although the standard deviations in Table V suggest that our results are well within the sampling distribution. The Monte Carlo results for the AR (2) in levels are even further from our actual estimates, no matter which of the two windows is used.

In summary, if we have to pick on the basis of estimates without attention to standard errors, then we must choose persistence, and consumption is much too smooth to be consistent with the permanent income model. If instead our perspective is one of testing, there is insufficient evidence to reject the hypothesis that labour income is ultimately non-persistent, although the reversion to trend must be very slow indeed. Once again, however, the main point is there is simply no support in these results for the proposition that permanent income is smoother than measured income.

4. EXCESS SMOOTHNESS AND EXCESS SENSITIVITY

Both parametric and non-parametric results are consistent with the supposition that consumption is smoother than it ought to be given rational expectations about permanent income. However, since Flavin's important 1981 paper, it has been generally supposed that the empirical evidence shows consumption to be excessively *sensitive* to innovations in measured income. In fact, and in spite of the language, the findings of this paper are entirely consistent with Flavin's and with other similar findings in the literature. In this final section we provide a reconciliation in terms of the VAR models of Section 2, and we show how exactly the same feature of the data is simultaneously responsible for both excess smoothness and excess sensitivity.

Although Table III is a useful summary of our findings from the VAR's, rather more insight can be obtained from a slightly different approach. If we examine the three different A matrices shown in Table II and compare their structures with the structure required for orthogonality in (22), it is apparent that the restriction on the second column is approximately satisfied, so that the rejection of the model reflects the fact that $a_{11} \neq a_{21}$. This informal impression is easily confirmed by calculating the test statistics for each of the two hypotheses separately. The data can therefore be characterized straightforwardly by writing A, not as in (22), but as

$$A = \begin{pmatrix} \alpha & \beta \\ \alpha - \chi & \beta + \rho^{-1} \end{pmatrix} \tag{34}$$

where $\chi > 0$ is the "excess sensitivity" parameter. Note that, if A satisfies (34), we have, from (19),

$$\frac{s_t}{y_t} - \Delta \log y_t - \frac{s_{t-1}}{\rho y_{t-1}} = -\chi \Delta \log y_{t-1} + (e_2' - e_1') u_t \approx -\frac{\Delta c_t}{y_{t-1}} \tag{35}$$

so that $\chi > 0$ reflects the fact that changes in consumption respond positively to last period's changes in income. It is this stylized fact of the data that has caused the model to fail in a number of studies using U.S. time-series data. Note also that the AR (1) structure for $\Delta \log y_t$ with a positive coefficient implies that $\Delta \log y_{t-1}$ predicts $\Delta \log y_t$, so that $\chi > 0$ can also be interpreted as a response of consumption to anticipated changes in current income, as would be generated by the existence of liquidity constraints.

Such a response to predictable movements in current income is what Flavin (1981) called "excess sensitivity". She estimated an excess sensitivity parameter using indirect least-squares on a just-identified two-equation system. But the covariances which identify excess sensitivity are just those between the current change in consumption and lagged levels of or changes in income, which in our system are captured by the parameter χ. In consequence, it is not the case that excess sensitivity to anticipated changes in income means that consumption is excessively sensitive to unanticipated changes in income. Indeed, as we now show, the opposite is true. The implications of the existence of excess sensitivity for forecasts of future income can most clearly be seen by evaluating (20) when A has the structure given by (34):

$$\textstyle\sum_{i=0}^{\infty} \rho^i (E_t - E_{t-1}) \Delta \log y_{t+i} = \sum_{i=0}^{\infty} e_1' \rho^i A^i u_t = e_1'(I - \rho A)^{-1} u_t = -(u_{2t} - u_{1t})/(1 - \rho\chi). \quad (36)$$

The theoretical innovation is therefore the actual innovation multiplied by $(1-\rho\chi)^{-1}$, which, since χ is a little less than 0·4, means that the predicted standard deviation will be about fifty percent too large, which is essentially what is shown in Table III. In the bivariate framework used here, the untoward sensitivity of changes in consumption to *anticipated* changes in income inevitably implies that consumption will respond by less than is warranted given the innovation in income. There is no contradiction between excess sensitivity and excess smoothness; they are the same phenomenon.

We believe that the answer to the question with which we began, "Why is consumption so smooth?" lies not in the truth of the permanent income hypothesis, but rather in its inadequacy. Whatever it is that causes changes in consumption to be correlated with lagged changes in income, whether it is that measured consumption and income are time averages of continuous processes, or that real interest rates vary, or that the marginal utility of consumption depends on other variables besides consumption, or that some consumers are liquidity constrained or that consumers adjust slowly through inertia or habit formation, that same failure of the model is responsible for the smoothness of consumption relative to permanent income.[8]

APPENDIX

A logarithmic version of the permanent income model

Start from equation (1) in the text, i.e.

$$c_t = \frac{r}{1+r}[A_t + \textstyle\sum_{i=0}^{\infty} (1+r)^{-i} E_t y_{t+i}]. \quad \text{(A1)}$$

Subtract capital income from both sides of (A1) to give

$$c_t - rA_t/(1+r) = y_t - s_t = \frac{r}{1+r}[y_t + \textstyle\sum_{i=1}^{\infty} (1+r)^{-i} E_t y_{t+i}]. \quad \text{(A2)}$$

Dividing through by current labour income y_t gives

$$1 - (s_t/y_t) = \frac{r}{1+r}[1 + \textstyle\sum_{i=1}^{\infty} (1+r)^{-i} E_t (y_{t+i}/y_t)]. \quad \text{(A3)}$$

Now, for all $j > 0$,

$$(y_{t+j}/y_t) = \exp[j\mu + \textstyle\sum_{k=1}^{j} (\Delta \log y_{t+k} - \mu)] \approx e^{j\mu}[1 + \sum_{k=1}^{j} (\Delta \log y_{t+k} - \mu)] \quad \text{(A4)}$$

8. Some of these alternative explanations are explored by Bean (1986), Campbell and Mankiw (1987*b*), Constantinides (1988), Deaton (1987), and Hall (1988), among others.

where $\mu = E(\Delta \log y_t)$ is the (unconditional) mean of the rate of growth of labour income. If the final term in (A4) is substituted in the right-hand side of (A3) and terms rearranged, we reach

$$1-(s_t/y_t) \approx \frac{r}{(1+r)(1-\rho)}[1+\sum_{i=1}^{\infty} \rho^{-i} E_t(\Delta \log y_{t+i} - \mu)] \tag{A5}$$

where $\rho = [(1+\mu)/(1+r)] \approx 1+\mu - r$. Provided the ratio of saving to labour income is sufficiently small, we can take logs of both sides and approximate once again to give:

$$(s_t/y_t) \approx -\sum_{i=1}^{\infty} \rho^i E_t \Delta \log y_{t+i} - \kappa, \tag{A6}$$

$$\kappa = \log(r/(1+r)) - \log(1-\rho) - \mu\rho/(1-\rho). \tag{A7}$$

Equation (A6) is equation (8) in the main text. From (A5), we have

$$E(s_t/y_t) \approx 1 - \frac{r}{(1+r)(1-\rho)} \approx r - \frac{\mu}{r} \tag{A8}$$

so that the saving ratio is small if r is small and if μ/r is small.

Acknowledgement. Revised version of "Is consumption too smooth?" dated January 1987. We are grateful to Larry Christiano, Marjorie Flavin, Fumio Hayashi, Charles Nelson, Ken West, and two referees who made helpful comments on previous versions. Campbell's research was supported by the National Science Foundation and by the John M. Olin Fellowship at the National Bureau of Economic Research.

REFERENCES

BEAN, C. R. (1986), "The estimation of "surprise" models and the "surprise" consumption function", *Review of Economic Studies*, **53**, 497-516.

BEVERIDGE, S. and NELSON, C. R. (1981), "A new approach to decomposition of economic time series data into permanent and transitory components with particular attention to measurement of the business cycle", *Journal of Monetary Economics*, **7**, 151-174.

BLINDER, A. S. and DEATON, A. S. (1985), "The time series consumption function revisited", *Brookings Papers on Economic Activity*, **2**, 465-511.

CAMPBELL, J. Y. (1987), "Does saving anticipate declining labor income? An alternative test of the permanent income hypothesis", *Econometrica*, **55**, 1249-1273.

CAMPBELL, J. Y. and MANKIW, N. G. (1987*a*), "Are output fluctuations transitory?", *Quarterly Journal of Economics*, **102**, 857-880.

CAMPBELL, J. Y. and MANKIW, N. G. (1987*b*), "Permanent income, current income, and consumption" (NBER Working Paper No. 2436).

CAMPBELL, J. Y. and MANKIW, N. G. (1989), "International evidence on the persistence of economic fluctuations", *Journal of Monetary Economics* (forthcoming).

COCHRANE, J. H. (1988), "How big is the random walk in GNP?", *Journal of Political Economy*, **96**, 893-920.

CONSTANTINIDES, G. (1988), "Habit formation: a resolution of the equity premium puzzle" (University of Chicago, unpublished paper).

DEATON, A. S. (1987), "Life-cycle models of consumption: is the evidence consistent with the theory?" in Bewley, T. F. (ed.) *Advances in Econometrics, Vol II* (Amsterdam: North-Holland) 121-148.

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

GRANGER, C. W. J. and NEWBOLD, P. (1977) *Forecasting economic time-series* (New York: Academic Press).

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

HALL, R. E. (1988), "Intertemporal substitution in consumption", *Journal of Political Economy*, **96**, 339-357.

HANSEN, L. P. and SARGENT, T. J. (1981*a*), "A note on Wiener-Kolmogorov prediction formulas for rational expectations models", *Economics Letters*, **8**, 255-260.

HANSEN, L. P. and SARGENT, T. J. (1981*b*), "Exact linear rational expectations models: specification and estimation" (Federal Reserve Bank of Minneapolis, Staff Report 71).

NELSON, C. R. (1987), "A reappraisal of recent tests of the permanent income hypothesis", *Journal of Political Economy*, **95**, 641-646.

NELSON, C. R. and PLOSSER, C. I. (1982), "Trends and random walks in macroeconomic time series: some evidence and implications", *Journal of Monetary Economics*, **10**, 139-162.

PRIESTLEY, M. R. (1982) *Spectral analysis and time series* (London: Academic Press).

WEST, K. D. (1988*a*), "Dividend innovations and stock price variability", *Econometrica*, **56**, 37-61.

WEST, K. D. (1988*b*), "The insensitivity of consumption to news about income", *Journal of Monetary Economics*, **21**, 17-23.