Journal of Monetary Economics 21 (1988) 17–33. North-Holland

THE INSENSITIVITY OF CONSUMPTION TO NEWS ABOUT INCOME

Kenneth D. WEST*

Princeton University, Princeton, NJ 08544, USA

Received December 1986, final version received May 1987

This paper uses a variance bounds test to see whether consumption is too sensitive to news about income to be consistent with a standard permanent income model, under the maintained hypothesis that income has a unit root. It is found that, if anything, consumption is less sensitive than the model would predict. This implication is robust to the representative consumer having private information about his future income that the econometrician does not have, to wealth shocks, and to transitory consumption. This suggests the importance in future research on the model of allowing for factors that tend to make consumption smooth.

1. Introduction

A standard rational expectations version of the permanent income model of consumption implies that the unanticipated component of consumption equals the unanticipated change in the expected present discounted value of labor income [Flavin (1981)]. Flavin's (1981) and Kotlikoff and Pakes's (1984) tests, however, indicated that post World War II aggregate U.S. consumption responds too strongly to news about income for this model to be correct. Flavin (1981), for example, found that the consumption response to an income innovation was over three times the value predicted by the model.

Flavin (1981) and Kotlikoff and Pakes (1984) accounted for the observed upward movement in per capita income by detrending their income series. Mankiw and Shapiro (1985) have pointed out that if income has a unit root with drift rather than a time trend, then the use of time trends in empirical tests will tend to spuriously suggest excess sensitivity of consumption to income.[1] Mankiw and Shapiro left open the question of whether or not

*I thank Joe Altonji, Ben Bernanke, Flint Brayton, Angus Deaton, an anonymous referee and participants in seminars at Northwestern and Princeton Universities for helpful comments and discussions, and the National Science Foundation for financial support.

[1] Mankiw and Shapiro (1985) conclude this in the sense that one will tend to spuriously find that lagged income helps predict changes in consumption. It follows from the sign of the biases reported in table 2 in Mankow and Shapiro (1985) and from the algebra in Flavin (1981, p. 993), however, that one will also tend to spuriously find excess sensitivity of changes in consumption to the income innovation.

0304-3932/88/$3.50

consumption is excessively sensitive, if in fact income has unit root. Deaton (1986) has argued that if such is the case, there is some evidence that consumption is in fact less sensitive to news about income than the model predicts – precisely the opposite conclusion that is reached when detrending is used.

This paper uses a variance bounds test to consider in detail the issue of the sensitivity of consumption to news about income, largely under the maintained hypothesis that the income process has a unit root. In section 2, I develop the implications of the model for the relationship between the relevant consumption and income variances. All of the papers cited in the previous paragraphs assumed that the representative consumer uses only lagged income to forecast future income, and exploited the resulting prediction that the unanticipated consumption component is equal to a certain function of the innovation in the univariate income process. This implication will not hold, however, if the representative consumer uses additional data such as, say, tax or labor market variables to forecast his income. In this case, the variance of the relevant consumption component will be less than the variance of this function of the univariate innovation. One can, however, use just consumption and income data to calculate precisely how much less variable consumption should be, and thus determine whether consumption is in fact too smooth.

In section 3, the paper uses some post World War II quarterly data to test both the inequality and equality derived in section 2, under the assumption that income has a unit root. As in the estimates reported in Deaton (1986), it is found that the relevant consumption variance is indeed less than the relevant income variance. The evidence does not strongly suggest, however, that this implied insensitivity of consumption results from additional information used by the consumer to forecast income. In various ARIMA specifications for the univariate income process, the point estimate of how much less variable consumption should be, given the consumer's superior information, is never more than a third, and is usually less than a tenth, of the point estimate of the difference in the variances. Neither wealth shocks nor white noise transitory consumption help explain the residual difference. The difference is significantly different from zero at the 5 percent level in almost all specifications.

This means that allowing for a unit root in the income process implies that the aggregate data are not quite as inconsistent with the permanent income model as is suggested when one allows instead for a time trend [Flavin (1981)]. On the other hand, the model by no means comfortably characterizes the data. It would seem that if one accepts the unit root specification, consumption is even smoother than the model predicts.

A final introductory word is appropriate, on why variance tests are useful in studying the permanent income model. An alternative would be to test the cross-equation restrictions of the model. Hansen and Sargent (1981) have pointed out that the cross-equation restrictions of a linear rational expecta-

tions model summarize all the restrictions of the model. So if these are obeyed, so, too, are any variance inequalities implied by the model. Indeed, one can show that unpredictability of changes in consumption and a transversality condition on wealth imply the variance inequality studied here, a fact noted independently in Hansen, Sargent and Roberds (1987).

The additional power of the tests of cross-equation restrictions does not, however, seem to be of critical importance in studying the permanent income model. Tests of the model have tended to suggest that whether or not one detrends, the model can be rejected by formal statistical tests [e.g., Blinder and Deaton (1986), Campbell (1985), Christiano, Eichenbaum and Marshall (1987), Flavin (1981), Hall (1978), Nelson (1985), Watson (1986)]. It is natural, then, to ask what stylized facts about consumption appear to be inconsistent with the model. In this connection, a variance test can be very revealing. It suggests that if income has a unit root, there is not much appeal to the argument that consumption is excessively sensitive to news about income. Rather, in future research that maintains the assumption of a unit root, it is important to allow for factors that tend to make consumption even smoother than the permanent income model predicts.

2. The model and test

The model is as in Flavin (1981). It is assumed that consumption equals permanent income, with permanent income the infinite horizon annuity value of the sum of human and non-human wealth:

$$c_t = rw_t + y_{tI}, \tag{1}$$

$$y_{tI} = r(1+r)^{-1}\sum_{0}^{\infty}(1+r)^{-j}\mathrm{E}y_{t+j}|I_t, \tag{2}$$

$$w_t = (1+r)w_{t-1} + y_{t-1} - c_{t-1}. \tag{3}$$

Here, c_t is consumption, r is the constant real interest rate, w_t is non-human wealth at the beginning of period t, y_{tI} is the annuity value of human wealth, y_t is labor income, $\mathrm{E}(\cdot|I_t)$ denotes expectations conditional on the consumer's information set I_t, assumed equivalent to linear projections. Summations in (2) and throughout the paper run over j. When 'income' is used without qualification, it should be understood to refer to labor income y_t.

Flavin (1981) showed that the model implies that the change in consumption equals the unpredictable change in the annuity value of labor income:

$$c_t - \mathrm{E}c_{t-1}|I_{t-1} = \Delta c_t = y_{tI} - \mathrm{E}y_{tI}|I_{t-1}. \tag{4}$$

So

$$\mathrm{var}(\Delta c_t) = \mathrm{E}(y_{tI} - \mathrm{E}y_{tI}|I_{t-1})^2 \equiv \sigma_i^2.$$

Let $H_t = \{1, y_t, y_{t-1}, \ldots\}$ be the information set determined by current and lagged labor income. Define

$$y_{tH} \equiv r(1+r)^{-1} \sum_{0}^{\infty} (1+r)^{-j} \mathrm{E} y_{t+j} | H_t.$$

Let σ_H^2 denote the variance of the innovation in y_{tH},

$$\sigma_H^2 = \mathrm{E}(y_{tH} - \mathrm{E} y_{tH} | H_{t-1})^2.$$

If $H_t = I_t$ – the representative consumer uses nothing but lagged income to forecast future income – then the model implies that $\mathrm{var}(\Delta c_t) = \sigma_H^2$. This is examined in Deaton (1986). (The mechanics of calculating σ_H^2 are explained below.) Suppose instead that H_t is a subset of I_t, because consumers use additional data to form better forecasts of future income. These data might be private signals about future income seen by the consumer, or observable macroeconomic variables such as, say, taxes or unemployment rates. It follows from Proposition 1 in West (1987) that in this case $\sigma_I^2 \leq \sigma_H^2$.[2] The forecasts made from H_t, which use less information, tend to be noisier. The model implies, then, that $\mathrm{var}(\Delta c_t) \leq \sigma_H^2$. Intuitively, the reason for this is that the permanent income model says that consumers try to smooth consumption in the face of income fluctuations. Additional information above and beyond that in the income series therefore will tend to make consumption smoother. Consumption being insensitive to income, in the sense that $\mathrm{var}(\Delta c_t) \leq \sigma_H^2$, is perfectly consistent with the model.

The model does, however, say how much smaller $\mathrm{var}(\Delta c_t)$ should be than σ_H^2. The difference between σ_I^2 and σ_H^2 is proportional to the variance of $y_{tI} - y_{tH}$. To understand why, observe first that by the law of iterated expectations, $\mathrm{E} y_{tI} | H_t = y_{tH}$. The permanent income model says that $y_{tI} = c_t - rw_t$, so

$$y_{tI} - y_{tH} = c_t - rw_t - \mathrm{E}(c_t - rw_t | H_t).$$

Now, $\mathrm{var}[c_t - rw_t - \mathrm{E}(c_t - rw_t | H_t)]$ is a measure of how much of the movement of $c_t - rw_t$ is not predictable by (is orthogonal to) past income. Naturally, the model says that this variance is larger the greater is the extent to which the consumer uses information above and beyond that in H_t in choosing consumption and wealth, i.e., the greater is the difference between σ_H^2 and σ_I^2.

[2] One key technical condition used in West (1987) is worth noting. This is that arithmetic differences suffice to induce stationarity in all the variables in I_t. This is consistent with most of the permanent income literature. Exceptions are Nelson (1985) and Watson (1986), who assume that log differences are required. Incidentally, if I_t contains variables in addition to lagged y_t, the variance–covariance matrix of the consumption and income innovations will not be singular, a problem noted in Hall (1986).

Specifically, the relation between σ_I^2 and σ_H^2 is[3]

$$\sigma_H^2 = \sigma_I^2 + \left[(1+r)^2 - 1\right]\mathrm{var}(y_{tI} - y_{tH}). \tag{5}$$

The permanent income model therefore implies

$$\sigma_H^2 = \sigma_c^2 + \sigma_v^2, \tag{6}$$

where

$$\sigma_c^2 \equiv \mathrm{var}(\Delta c_t) \quad \text{and} \quad \sigma_v^2 \equiv \left[(1+r)^2 - 1\right]\mathrm{var}(c_t - rw_t - y_{tH}).$$

Of course, if the model is incorrect (e.g., there are liquidity constraints), then, in general,

$$\sigma_c^2 \neq \sigma_I^2 \equiv \mathrm{E}(y_{tI} - \mathrm{E}y_{tI}|I_{t-1})^2,$$

and

$$\sigma_v^2 \neq \left[(1+r)^2 - 1\right]\mathrm{var}(y_{tI} - y_{tH}).$$

Eq. (6) may become clearer if the procedure used in part of the empirical work is detailed. Suppose that the univariate y_t process follows an ARIMA$(p, 1, q)$ process,

$$\Delta y_t = m + \phi_1 \Delta y_{t-1} + \cdots + \phi_p \Delta y_{t-p} + e_t + \cdots + \theta_q e_{t-q}. \tag{7}$$

Hansen and Sargent (1980) show that

$$y_{tH} = \text{constant} + \delta_1 y_t + \cdots + \delta_{p+1} y_{t-p} + \pi_1 e_t + \cdots + \pi_{q-1} e_{t-q+1};$$

the δ_i and π_i are functions of r, the ϕ_i and the θ_i (e.g., $\delta_1 = [1 - \phi_1(1+r)^{-1} - \cdots - \phi_p(1+r)^{-p}]^{-1}$, $\pi_1 = \delta_1[\theta_1(1+r)^{-1} + \cdots + \theta_q(1+r)^{-q}]$). Then

$$y_{tH} - \mathrm{E}y_{tH}|H_{t-1} = (\delta_1 + \pi_1)e_t \equiv \psi e_t,$$

[3]See eqs. (9) to (11) below for the intuition behind the $(1+r)^2 - 1$ term in eq. (5). Eqs. (5) and (9) are established in West (1987) (although that paper only studies in detail the implications of the inequality $\sigma_H^2 \geq \sigma_I^2$, for stock prices and dividends). Incidentally, eq. (5) does not say that $\sigma_H^2 - \sigma_I^2$ depends on r in any simple way, since $y_{tI} - y_{tH}$ potentially varies with r in a complicated manner.

where

$$\psi = \left[1 + \theta_1(1+r)^{-1} + \cdots + \theta_q(1+r)^{-q}\right]$$

$$/\left[1 - \phi_1(1+r)^{-1} - \cdots - \phi_p(1+r)^{-p}\right]. \tag{8}$$

So $\sigma_H^2 = \psi^2\sigma_e^2$, and σ_H^2 may be calculated from r and the usual estimates of the Δy_t process. One can then test $\mathrm{var}(\Delta c_t) \leq \sigma_H^2$. To calculate σ_v^2, one first computes the variance of $c_t - rw_t - y_{tH}$, using the y_t, the estimates of the δ_i and π_i, and, if $q > 0$, the residuals from the estimates of the Δy_t process, to compute y_{tH} for each t. This is then multiplied by the proportionality factor $(1+r)^2 - 1$.

3. Empirical results

The Blinder and Deaton (1986) data were used, and were kindly supplied by Angus Deaton. The data were real (1972 dollars), seasonally adjusted, and per capita, 1953:2 to 1984:4. The consumption data were for non-durables and services, excluding shoes and clothing. These data were divided by 0.7855, the mean fraction of such consumption in total consumption over this period, before any statistics were calculated. Additional details on the data, as well as on the empirical results, are in an appendix available on request from the author.

Blinder and Deaton constructed separate series for labor income and disposable income. I measured rw_t, income from non-human wealth, in two ways. The first followed Campbell (1985) and set rw_t to the difference between the two income series. The second set w_t to the MPS series for household net worth, converted to real (1972) per capita dollars, and then calculated the implied rw_t. The estimates of $\mathrm{var}(c_t - rw_t - y_{tH})$ that resulted from the first measure are called σ_{v1}^2, those from the second measure are called σ_{v2}^2. A quarterly real interest rate of 0.5 percent was assumed throughout the results reported below. Point estimates (though not standard errors) were also calculated for quarterly interest rates of 0.25, 0.75, 1.0 and 1.25 percent. These are not reported, since the results were very similar, but are available on request.

As just noted, the test of the inequality and equality variance relations requires estimates of the parameters of the univariate Δy_t process. This was done assuming that Δy_t follows an ARMA(p, q) process, with $0 \leq p, q \leq 2$. This wide variety of processes was used to make sure that the results were not sensitive to the exact specification chosen. The ARMA parameters were estimated by non-linear least squares, with the presample disturbances set to zero. The Monte Carlo evidence in Ansley and Newbold (1980) suggests that this technique has attractive small sample properties when roots are not near

the unit circle, as appears to be the case in these data. All variances were calculated with the appropriate degrees of freedom adjustment.

The estimated parameter vector included not only the autoregressive coefficients, but all the variances that needed to be computed. The covariance matrix of the estimated vector was calculated using the methods of Newey and West (1987). The technique properly accounts for the uncertainty about all the elements of the parameter vector, and allows, for example, arbitrary serial correlation of the difference between $c_t - rw_t$ and y_{tH}, and for arbitrary heteroskedasticity of the disturbances conditional on past values of Δy_t.[4] A tenth-order Newey and West (1987) correction was used because the asymptotic theory requires that the order of the correction be the square root of the sample size, which was about 120. A small amount of experimentation with fifth- and fifteenth-order corrections indicated that the calculated standard errors are not sensitive to the order of the correction.

Table 1 contains the estimates of the univariate Δy_t process. Application of Box–Jenkins techniques would probably suggest an AR(1), or perhaps an MA(1): neither ϕ_2 nor θ_2 are significantly different from zero at the 5 percent level in any specification. Except for $(p, q) = (0, 0)$ (which is the only specification that has a Q statistic significantly different from zero), the implied values of ψ are very similar. They range from about 1.4 to about 1.9.

Table 2 contains the results on the tests of the innovation variances. As may be seen in column 4, the null hypothesis that $\sigma_H^2 - \sigma_c^2$ is zero can be comfortably rejected at the 5 percent level for all specifications. The permanent income model does not fare as well when one tests instead the null that $\sigma_H^2 = \sigma_c^2 + \sigma_v^2$. See columns 5 and 6 when rw_t is measured as the difference between disposable and labor income, columns 7 and 8 when it is measured from the MPS wealth series. The estimates of σ_{vi}^2, $i = 1, 2$, are fairly insensitive to choice of p and q. Except when $(p, q) = (0, 0)$, the point estimates of σ_{v1}^2 and σ_{v2}^2 are never more than one sixth the estimate of $\sigma_H^2 - \sigma_c^2$. The differences reported in columns 6 and 8 are significantly different from zero at the 5 percent level in all specifications except $(p, q) = (1, 2)$ and $(p, q) = (2, 2)$, where the differences are significant at the 10 percent level.

In sum, then, column 4 suggests that σ_c^2 is less than σ_H^2, which is what the permanent income model predicts. Unfortunately, it appears from columns 6 and 8 that the implied insensitivity of consumption to news about income is unlikely to result purely from the use by the consumer of additional variables to forecast income.

The remainder of this section briefly considers two minor modifications to the model (1)–(3), and four technical modifications to the procedure used.

[4] The $(p + q + 1)$ past values of Δy_t that are used as instruments in non-linear least squares are $\partial e_t/\partial m$, $\partial e_t/\partial \phi_1, \ldots, \partial e_t/\partial \phi_p$, $\partial e_t/\partial \theta_1, \ldots, \partial e_t/\partial \theta_q$. Heteroskedasticity of e_t conditional on these past values of Δy_t was accounted for.

Table 1

Estimates of the Δy_t process.[a]

(1) (p,q)	(2) m	(3) ϕ_1	(4) ϕ_2	(5) θ_1	(6) θ_2	(7) ψ	(8) σ_e^2	(9) Q
(0,0)	14.08					1.00	790.9	57.46
	(3.67)						(150.2)	(0.01)
(1,0)	8.17	0.44				1.79	636.1	37.74
	(2.35)	(0.06)				(0.20)	(127.4)	(0.30)
(0,1)	13.98			0.40		1.40	659.5	35.57
	(3.78)			(0.07)		(0.07)	(129.1)	(0.30)
(1,1)	7.07	0.52		−0.10		1.86	640.6	37.65
	(3.12)	(0.12)		(0.14)		(0.28)	(126.8)	(0.19)
(2,0)	8.33	0.43	0.01			1.78	643.6	37.73
	(2.71)	(0.07)	(0.11)			(0.23)	(128.1)	(0.19)
(0,2)	13.91			0.45	0.11	1.55	633.3	36.42
	(3.83)			(0.07)	(0.11)	(0.15)	(127.5)	(0.23)
(2,1)	4.64	0.86	−0.17	−0.44		1.81	646.2	37.20
	(3.84)	(0.53)	(0.24)	(0.54)		(1.35)	(128.3)	(0.17)
(1,2)	5.24	0.65		−0.22	−0.12	1.90	640.6	36.10
	(5.63)	(0.39)		(0.40)	(0.27)	(1.34)	(127.6)	(0.20)
(2,2)	6.42	0.50	0.07	−0.07	−0.13	1.88	649.2	35.60
	(14.07)	(1.61)	(0.72)	(1.64)	(0.23)	(1.29)	(127.7)	(0.19)

[a]Sample period is 1953:3–1984:4 for $p=0$, 1953:4–1984:4 for $p=1$, 1954:1–1984:4 for $p=2$. Heteroskedasticity consistent asymptotic standard errors are in parentheses. The Box–Pierce Q statistic is asymptotically distributed as a χ^2 $(33-p-q)$ random variable, with marginal significance level given in parentheses.

Table 2

Empirical results.[a]

(1) (p,q)	(2) σ_H^2	(3) σ_c^2	(4) $\sigma_H^2-\sigma_c^2$	(5) σ_{v1}^2	(6) $\sigma_H^2-\sigma_c^2-\sigma_{v1}^2$	(7) σ_{v2}^2	(8) $\sigma_H^2-\sigma_c^2-\sigma_{v2}^2$
(0,0)	790.9	250.7	540.2	151.4	388.7	82.4	457.8
	(150.2)	(30.6)	(128.7)	(29.6)	(132.6)	(21.8)	(144.6)
(1,0)	2028.2	246.1	1782.1	161.6	1620.5	95.5	1686.6
	(670.2)	(30.2)	(666.5)	(31.4)	(657.4)	(22.9)	(677.9)
(0,1)	1285.3	250.7	1034.5	154.0	880.5	85.7	948.9
	(316.7)	(30.6)	(301.7)	(30.4)	(300.6)	(21.7)	(313.1)
(1,1)	2226.3	246.1	1980.2	164.9	1815.3	99.4	1880.8
	(817.9)	(30.2)	(740.1)	(33.3)	(734.4)	(24.4)	(753.9)
(2,0)	2042.4	236.9	1805.5	161.9	1643.6	96.3	1709.2
	(689.6)	(31.1)	(687.5)	(31.9)	(679.2)	(23.2)	(669.2)
(0,2)	1577.9	250.7	1327.2	155.9	1171.3	87.9	1239.3
	(482.0)	(30.6)	(478.1)	(29.8)	(468.2)	(21.9)	(488.2)
(2,1)	2126.4	236.9	1889.6	164.4	1725.1	99.1	1790.4
	(3348.9)	(32.1)	(694.8)	(36.0)	(692.4)	(26.3)	(708.8)
(1,2)	2322.5	246.1	2076.4	169.3	1907.1	104.3	1972.1
	(3413.8)	(30.2)	(1041.4)	(52.9)	(1068.1)	(46.1)	(1079.5)
(2,2)	2304.9	236.9	2068.0	167.8	1900.2	103.3	1964.7
	(3324.9)	(31.1)	(969.3)	(49.8)	(994.4)	(47.0)	(1009.1)

[a]Sample period is 1953:3–1984:4 for $p=0$, 1953:4–1984:4 for $p=1$, 1954:1–1984:4 for $p=2$. Asymptotic standard errors are in parentheses. Units are 1972 dollars, squared.

None of these appear likely to explain the insensitivity. The modifications to the model:

(1) *Wealth shocks.* Let us modify the budget constraint (3) to allow for shocks to wealth, say, unanticipated capital gains [Campbell (1985)]:

$$w_t = (1+r)w_{t-1} + (y_{t-1} - c_{t-1}) + a_t,$$

where a_t is a white noise random variable. This implies that eq. (4) becomes

$$\Delta c_t = y_{tI} - \mathrm{E}y_{tI}|I_{t-1} + ra_t.$$

If the wealth shock a_t is negatively correlated with the innovation in the present value of labor income, then $\mathrm{var}(\Delta c_t)$ will be less than σ_I^2. Such a shock therefore potentially explains the results in table 1.

To accommodate this possibility, subtract $ra_t = r[w_t - (1+r)w_{t-1} - (y_{t-1} - c_{t-1})]$ from Δc_t. This yields

$$x_t \equiv -(y_t + rw_t - c_t) + (1+r)(y_{t-1} + rw_{t-1} - c_{t-1}) + \Delta y_t$$

$$= \Delta c_t - ra_t = y_{tI} - \mathrm{E}y_{tI}|I_{t-1}.$$

One can then calculate the variance of x_t instead of Δc_t.

This was done for all the specifications in table 1. When the first measure of rw_t was used (difference between disposable and labor income), the estimates of σ_x^2 were slightly higher than those reported in the σ_c^2 column in table 2; when the second measure was used (r times MPS wealth), the estimates were slightly lower. For $(p,q) = (0,0)$, σ_v^2 was slightly over one third of $\sigma_H^2 - \sigma_x^2$; no other estimates were more than one sixth of $\sigma_H^2 - \sigma_x^2$.

(2) *Transitory consumption.* Suppose that $c_t = rw_t + y_{tI}$ + transitory consumption, where transitory consumption is a zero mean stationary variable. If transitory consumption is uncorrelated with any of the variables used to forecast income, then $\mathrm{var}(\Delta c_t) = \sigma_I^2 + \mathrm{var}$(linearly filtered transitory consumption) [see Flavin (1981) for the exact formula when transitory consumption is white noise] and so $\mathrm{var}(\Delta c_t)$ is bigger than σ_I^2. Also, $\mathrm{var}(c_t - rw_t - y_{tH}) = \mathrm{var}(y_{tI} - y_{tH}) + \mathrm{var}$(transitory consumption) is larger than $\mathrm{var}(y_{tI} - y_{tH})$. As noted in Deaton (1986), then, such transitory consumption cannot explain excess smoothness of consumption. The same applies to transitory consump-

tion positively correlated with news about income (say, because of within quarter multiplier effects).[5]

The four technical modifications to the procedure used:

(1) *Monte Carlo estimates of significance levels.* It is possible that there is a strong finite sample bias towards rejection, even when the model is true. To investigate this possibility, a small Monte Carlo experiment was performed. For ARMA(1, 0) and ARMA(0, 1) processes, the permanent income model was used to generate one hundred artificial samples of consumption and income data of size 125. The ARMA parameters matched those reported in table 1.[6] For each sample, the relevant variances were estimated as described at the beginning of this section, and the estimated $\sigma_{vi}^2/(\sigma_H^2 - \sigma_c^2)$ was calculated. There was a tabulation of the number of times this fraction was positive and less than that implied by the table 2 estimates. This experiment, then, is intended to get an idea of how likely it is that the point estimates will suggest that only a fraction of the difference between σ_H^2 and σ_c^2 is explained by the consumer having additional variables to forecast income, when in fact the entire difference is so explained.

The results are in table 3. To read the table, consider the entries in line 1. The column 2 entry is 0.091 = 161.6/1782.1 = (table 2, line 2, column 5)/(table 2, line 2, column 4). Now, 100 samples were generated with the true (population) value of $\sigma_c^2 = 246.1$, the true $\sigma_{v1}^2 = 161.6$. The column 3 entry of 0.02 indicates that in only 2 of these 100 was the estimated σ_{v1}^2 less than 0.091 of the estimated $\sigma_H^2 - \sigma_c^2$. The column 4 entry in line 1 is 0.054 = 95.5/1782.1 = (table 2, line 2, column 7)/(table 2, line 2, column 4). The column 5 entry of 0.00 indicates that in none of the samples generated with $\sigma_c^2 = 246.1$, $\sigma_{v2}^2 = 95.5$, was the estimated σ_{v2}^2 less than 0.054 of the estimated $\sigma_H^2 - \sigma_c^2$.

[5]Another extension to the model deserved mention, namely, allowing for variations in expected returns. While this is a possible avenue for future research on consumption variability [see Christiano (1987)], this is not pursued here. The basic reason is that consumption models that allow for such variations still find evidence against the model [e.g., Grossman and Shiller (1981)]. This suggests that simply generalizing the model to allow for this variation will not persuasively reconcile the consumption and income data, especially since Michener (1984) has argued that in general equilibrium, this variation will make consumption *more* sensitive to income than the permanent income model predicts.

[6]Specifically, for the AR(1) process [the MA(1) simulation was analogous]: write the Δy_t process as $\Delta y_t = \mu + \phi_1 \Delta y_t + e_t$. It was assumed that $I_t = \{1, z_{t-j}, x_{t-j}\}$, where $e_t = z_t + x_{t-60}$, with x_t and z_t mutually and serially uncorrelated zero mean normal random variables. It is routine to use the formulas in Hansen and Sargent (1980) to calculate y_{tI}. The values of μ and ϕ_1 were chosen to match those estimated in the data, those of σ_z^2 and σ_x^2 so that σ_c^2 and σ_{vi}^2, $i = 1$ or 2, would match those estimated in the data and reported in table 2. A different random number seen was used to initiate the generation of the z_t and x_t for each of the four different specifications in table 3.

Table 3
Monte Carlo marginal significance levels.

(1) (p, q)	(2) v_1 fraction	(3) Monte Carlo M.S.L.	(4) v_2 fraction	(5) Monte Carlo M.S.L.
(1,0)	0.091	0.02	0.054	0.00
(0,1)	0.149	0.00	0.083	0.00

The significance levels in columns 3 and 5 of table 3 are consistent with those implied by columns 6 and 8 in table 2. In particular, the Monte Carlo experiment suggests that the odds are less than 0.05 that the results for the ARMA(1,0) and ARMA(0,1) specifications are purely due to sampling error, rather than to a shortcoming of the model. Since only 100 samples were used to establish the Monte Carlo significance levels, this experiment does not establish the small sample distribution of the table 2 estimates with any great degree of precision. But the experiment also does not suggest that there is a systematic bias towards rejection of the model.

(2) *Estimates for subsamples.* Point estimates (though not standard errors) of all the entries in table 2 were calculated for samples ending in 1973:3 and beginning in 1974:1. This was done to guard against the possibility that the first OPEC shock caused an unexpected shift in the stochastic process for c_t and y_t, thereby biasing the estimates in an unpredictable way. In each of the two subsamples, however, the point estimates of the σ_{vi}^2 were only a fraction of the point estimates of the corresponding $\sigma_H^2 - \sigma_c^2$. In particular, the ratio of σ_{vi}^2, $i = 1$ or 2, to $\sigma_H^2 - \sigma_c^2$ never exceeded one fifth. This suggests that biases induced by any such a shift in the stochastic processes for c_t and y_t are unlikely to explain the table 2 results.

(3) *Non-parametric estimates of* σ_H^2. In a different context, Cochrane (1986) has argued that the use of low-order ARMA models can cause large errors in estimation of quantities like σ_H^2. Angus Deaton has pointed out to me that σ_H^2 can be approximated by the frequency domain quantity that Cochrane (1986) suggested for a different purpose.

Write the moving average representation of Δy_t as $\Delta y_t - \mathrm{E}\Delta y_t = d(L)e_t$, $d(L) = 1 + d_1 L + d_2 L^2 + \cdots$, with L the lag operator. Hansen and Sargent (1980) show that $\sigma_H^2 = \{d[(1+r)^{-1}]\}^2 \sigma_e^2$, where $d[(1+r)^{-1}] = 1 + d_1(1+r)^{-1} + d_2(1+r)^{-2} + \cdots$. Consider approximating $\{d[(1+r)^{-1}]\}^2 \sigma_e^2$ by $[d(1)]^2 \sigma_e^2$, which may be a reasonable approximation since r is very small. If Δy_t is AR(1) with first-order serial correlation coefficient of 0.44 (the estimate

in these data), for example, $[d(1)]^2 = 3.19$, $\{d[(1+r)^{-1}]\}^2 = 3.16$ when $r = 0.005$.[7]

Now, $[d(1)]^2\sigma_e^2$ is just the spectral density of Δy_t at frequency zero. Thus we can use the spectral density to approximate $\{d[(1+r)^{-1}]\}^2$, without parametrically specifying the Δy_t process. It should be noted that this approximation is applicable even if y_t is stationary around a time trend. It is also applicable if y_t is a mixture of stationary and unit root processes, as in Watson (1986).

I estimated this using what Anderson (1971, p. 512) calls a modified Bartlett estimator. This estimator is simply a weighted sum of the sample autocovariances of Δy_t. (Recall that Δy_t's spectral density evaluated at frequency zero is simply the sum of its autocovariances.) I tried summing Δy_t's first 5, 10, 15 and 20 sample autocovariances. The smallest estimate happened to occur when 20 were used. (I report the smallest because this gives the model any possible benefit of the doubt.) It was 1630, in the middle of the table 2 estimates of σ_H^2. Using the asymptotic normal approximation to the finite sample distribution [Anderson (1971, p. 540)], a 95 percent confidence interval for this estimate is about (857, 16582). Unsurprisingly, the non-parametric estimate is somewhat noisier than are the parametric ones. The values of most of the point estimates of $\sigma_c^2 + \sigma_v^2$ in table 2 are nonetheless outside this confidence interval.[8]

(4) *Using data from every fourth quarter, rather than every quarter.* This obviously will reduce any biases induced by seasonal adjustment. It also may reduce any biases from moving average components due to time aggregation: if instantaneous consumption is a continuous time random walk, it is well known that measured $c_t - c_{t-1}$ is MA(1) with a coefficient of $\frac{1}{4}$ [Christiano and Eichenbaum (1986)]; it is straightforward to verify that in such a case, measured $c_t - c_{t-4}$ is MA(1) with a coefficient of $\frac{1}{22}$.[9]

The relationship that was used to derive eq. (6) above is [West (1987)]

$$\sum_1^\infty (1+r)^{-j}\left[d(1+r)^{-1}\right]e_{t+j}$$

$$= \sum_1^\infty (1+r)^{-j}\Delta c_{t+j} + (c_t - rw_t - y_{tH}). \tag{9}$$

[7]Even though $[d(1)]^2 > \{d[(1+r)^{-1}]\}^2$ in this example, there is no presumption of an upward bias in general.

[8]The 95 percent confidence interval is not valid if the true value of the spectral density is zero, as would be the case if y_t is stationary around a time trend. The interval is valid, however, if y_t is a mixture of trend stationary and unit root components, as in Watson (1986).

[9]See Christiano, Eichenbaum and Marshall (1987) for a careful, rigorous test of an explicit continuous time consumption model.

This can be rewritten as

$$y_{tH} - y_t + \sum_1^\infty (1+r)^{-j}\left[d(1+r)^{-1}\right]e_{t+j}$$

$$= \sum_1^\infty (1+r)^{-j}\Delta c_{t+j} + (c_t - rw_t - y_t). \tag{10}$$

Under the null, $\mathrm{E}\Delta c_{t+j}(c_t - rw_t - y_t) = 0$ for all $j > 0$; since e_t is the univariate innovation in y_t, $\mathrm{E}e_{t+j}(y_{tH} - y_t) = 0$ for all $j > 0$. Upon calculating the variance of each side of (10), then, and multiplying by $(1+r)^2 - 1$, we obtain

$$\left[(1+r)^2 - 1\right]\mathrm{var}(y_{tH} - y_t) + \left[d(1+r)^{-1}\right]^2\sigma_e^2$$

$$= \sigma_c^2 + \left[(1+r)^2 - 1\right]\mathrm{var}(c_t - rw_t - y_t)$$

$$\equiv \sigma_c^2 + \sigma_{v*}^2. \tag{11}$$

The variance σ_{v*}^2 may be consistently (though inefficiently) estimated as $[(1+r)^2 - 1]$ times the sample variance of every fourth observation on $c_t - rw_t - y_t$; σ_c^2 can be consistently (though inefficiently) estimated as $\frac{1}{4}$ times the sample variance of every fourth observation on $c_t - c_{t-4}$. There does not appear to be any simple way of estimating the left-hand side of (11) using every fourth observation. We have, however, $\lim_{n\to\infty}(1/n)\mathrm{var}(y_t - y_{t-n}) = [d(1)]^2\sigma_e^2$ [Cochrane (1986)]. Consider approximating the left-hand side of (11) with

$$\tfrac{1}{4}\mathrm{var}(y_t - y_{t-4}) \equiv \sigma_{H*}^2.$$

This obviously can be estimated using data from every fourth observation. Note that the approximation ignores the $[(1+r)^2 - 1]\mathrm{var}(y_{tH} - y_t)$ term and therefore may underestimate the left-hand side of (11). In particular, if Δy_t is an AR(1) or MA(1) with a positive ϕ or θ (either of which seems plausible, in light of the table 1 estimates), it may be shown that $[(1+r)^2 - 1]\mathrm{var}(y_{tH} - y_t) + [d(1+r)^{-1}]^2\sigma_e^2 \geq \sigma_{H*}^2$. For either of these two ARMA specifications, then, and perhaps more generally, if $\hat{\sigma}_{H*}^2 \geq \hat{\sigma}_{v*}^2 + \hat{\sigma}_c^2$, the implication is once again that consumption is too insensitive to news about income.

The variances were calculated for each quarter in four separate tests. (A more powerful test would of course result from pooling the four sets of

Table 4

Empirical results, using every fourth observation.[a]

(1) Quarter	(2) σ^2_{H*}	(3) σ^2_c	(4) $\sigma^2_{H*} - \sigma^2_c$	(5) σ^{2*}_{v1}	(6) $\sigma^2_{H*} - \sigma^2_c - \sigma^{2*}_{v1}$	(7) σ^{2*}_{v2}	(8) $\sigma^{2*}_H - \sigma^2_c - \sigma^{2*}_{v2}$
First	1817.1	422.4	1394.7	154.7	1240.0	82.6	1312.1
	(475.1)	(156.8)	(440.5)	(33.5)	(451.8)	(23.0)	(456.2)
Second	1731.1	328.1	1403.0	148.0	1255.0	92.4	1310.6
	(317.5)	(104.8)	(326.8)	(30.4)	(338.9)	(24.9)	(343.8)
Third	1313.2	356.9	956.4	156.2	800.1	85.5	870.9
	(276.2)	(123.0)	(279.2)	(35.0)	(291.1)	(20.6)	(291.5)
Fourth	1481.5	464.0	1017.6	165.2	852.3	84.7	932.9
	(360.4)	(176.3)	(327.2)	(40.0)	(336.4)	(23.4)	(342.5)

[a]See notes to table 2.

estimates, and performing a joint test. This, however, seemed pointless, in light of the result.) Thirty observations were available for the first quarter of the year, thirty-one observations for the other quarters. A fifth-order Newey and West (1987) correction was used in calculating the standard errors.

The results are in table 4. The point estimates of σ_c^2 are slightly higher than in table 2, indicating positive serial correlation in Δc_t. The point estimates $\sigma_{v^*}^2$ are of course quite similar to the table 2 estimates of σ_v^2 for $(p, q) = (0,0)$. The estimates of $\sigma_{H^*}^2$ are, however, so high that consumption once again appears to be insensitive to news about income. The statistical significance of the rejections is quite strong, though with only thirty or thirty-one observations the asymptotic normal approximation perhaps should not be taken very seriously.

4. Conclusions

The variance bounds test applied here suggests that consumption is even less sensitive to news about income than the permanent income model predicts. The test maintained the assumption that income has a unit root (although there was one non-parametric estimate that is valid even if income is stationary around a time trend). If, then, income does have a unit root, as is argued in Mankiw and Shapiro (1985) and Deaton (1986), a stylized fact is that consumption is insensitive to news about income. This does not suggest (to me) liquidity constraints, as is considered in, for example, Flavin (1985). Extensions of the model that seem more likely to be consistent with consumption insensitivity include non-separability of preferences, so that consumption expenditures in a given period yield utility in future periods [e.g., Eichenbaum, Hansen and Singleton (1986)], costs of adjusting consumption [e.g., Bernanke (1985)] and habit persistence [e.g., Deaton (1986)].

References

Anderson, T.W., 1971, The statistical analysis of time series (Wiley, New York).
Ansley, Craig F. and Paul Newbold, 1980, Finite sample properties of estimators for autoregressive moving average models, Journal of Econometrics 13, 159–183.
Bernanke, Ben S., 1985, Adjustment costs, durables and aggregate consumption, Journal of Monetary Economics 15, 41–68.
Blinder, Alan S. and Angus Deaton, 1985, The time series consumption function revisited, Brookings Papers on Economic Activity 2, 465–511.
Campbell, John Y., 1985, Does saving anticipate declining labor income? An alternative test of the permanent income hypothesis, Manuscript.
Christiano, Lawrence J., 1987, Is consumption insufficiently sensitive to innovations in income?, American Economic Review 77, 337–341.
Christiano, Lawrence J. and Martin Eichenbaum, 1986, Temporal aggregation and structural inference in macroeconomics, NBER technical working paper no. 60.
Christiano, Lawrence J., Martin Eichenbaum and David Marshall, 1987, The permanent income hypothesis revisited, Federal Reserve Bank of Minneapolis working paper no. 335.
Cochrane, John H., 1986, How big is the random walk component in GNP?, Manuscript.

Deaton, Angus, 1986, Life cycle models of consumption: Is the evidence consistent with the theory? NBER working paper no. 1910.

Eichenbaum, Martin S., Lars Peter Hansen and Kenneth J. Singleton, 1986, A time series analysis of representative agent models of consumption and leisure under uncertainty, NBER working paper no. 1981.

Flavin, Marjorie A., 1985, Excess sensitivity of consumption to current income: Liquidity constraints or myopia?, NBER working paper no. 1341.

Grossman, Sanford J. and Robert J. Shiller, 1981, The determinants of the variability of stock prices, American Economic Review 71, 222–227.

Hall, Robert E., 1978, Stochastic implications of the life cycle–permanent income hypothesis: Theory and evidence, Journal of Political Economy 86, 971–987.

Hall, Robert E., 1986, Consumption, Manuscript.

Hansen, Lars Peter and Thomas J. Sargent, 1980, Formulating and estimating dynamic linear rational expectations models, Journal of Economic Dynamics and Control 2, 7–46.

Hansen, Lars Peter and Thomas J. Sargent, 1981, Exact linear rational expectations models: Specification and estimation, Federal Reserve Bank of Minneapolis Research Department staff report no. 71.

Hansen, Lars Peter, Thomas J. Sargent and William Roberds, 1987, Time series implications of present value budget balance and of martingale models of consumption or taxes, Manuscript.

Kotlikoff, Laurence J. and Ariel Pakes, 1984, Looking for news in the noise: Additional stochastic implications of optimal consumption choice, NBER working paper no. 1492.

Mankiw, N. Gregory and Matthew D. Shapiro, 1985, Trends, random walks and tests of the permanent income hypothesis, Journal of Monetary Economics 16, 165–174.

Michener, Ron, 1984, Permanent income in general equilibrium, Journal of Monetary Economics 13, 297–306.

Nelson, Charles R., 1985, A reappraisal of recent tests of the permanent income hypothesis, NBER working paper no. 1687.

Newey, Whitney K. and Kenneth D. West, 1987, A simple, positive semidefinite, heteroskedasticity and autocorrelation consistent covariance matrix, Econometrica, forthcoming.

Watson, Mark W., 1986, Univariate detrending methods with stochastic trends, Journal of Monetary Economics 18, 49–76.

West, Kenneth D., 1987, Dividend innovations and stock price volatility, Econometrica, forthcoming.