THE KOREAN ECONOMIC REVIEW
Volume 12, Number 1, Summer 1996.

CHARACTERIZING THE FAILURE OF THE PERMANENT INCOME HYPOTHESIS*

CHULSOO KIM**

This paper characterizes the failure of the permanent income hypothesis (PIH) in terms of excess sensitivity and excess smoothness. First, this paper examines a model explaining the excess sensitivity of consumption to current income by computing all the testable implications of the model and characterizing its failure in terms of the extent to which consumption deviates from the excess sensitivity of consumption. The model has the smallest deviation when the fraction of agents who consume their current income rather than permanent income is about 0.5. Second, this paper develops an excess smoothness test which is independent of the presence of unit roots in the labor income process. Consumption is shown to be too smooth to be justified by the PIH.

I. INTRODUCTION

This paper characterizes the failure of the permanent income hypothesis (PIH) in terms of excess sensitivity and excess smoothness. Numerous authors such as Flavin (1981) and Campbell (1987) have statistically rejected the PIH using predictability tests while Kim (1996) finds that consumption deviates marginally from the PIH. Instead of simply accepting or rejecting the PIH, this paper attempts to characterize its failure in two ways. First, we examine the hypothesis of excess sensitivity of consumption as the null hypothesis rather than the alternative hypothesis and compare the specification error or deviation from the PIH with that

* I would like to thank an anonymous referee, Orazio P. Attanasio, Steven N. Durlauf, Robert E. Hall, Pok-sang Lam, Edward Steinmueller, Joseph E. Stiglitz and James H. Stock for helpful discussion and comments and John Y. Campbell for supplying me with his data set. Of course, the remaining errors are mine. I gratefully acknowledge the financial support from the 1996 Sookmyung Women's University Research Fund, and the Lynde and Harry Bradley Foundation Graduate Fellows Program and Center for Economic Policy Research, Stanford University. This is a revision of CEPR Technical Paper # 188.

** Department of Economics, Rutgers University and Department of Economics, Sookmyung Women's University.

from the Keynesian consumption hypothesis. Second, we devise a statistical procedure to test the hypothesis of excess smoothness of consumption which is independent of the unit root assumption in labor income, and show that consumption is in fact too smooth to be consistent with the PIH.

This paper sets up the signal extraction approach to the hypothesis of excess sensitivity of consumption to current income and measures the extent to which consumption deviates from the hypothesis. The signal extraction approach enables us to examine whether the model is reasonable rather than whether the model is exactly correctly specified. The hypothesis of excess sensitivity of consumption to current income has been suggested to be a reasonable alternative to the PIH. Authors such as Campbell and Mankiw (1989, 1990, 1991), Hayashi (1982) and Flavin (1981) have tested the PIH against the excess sensitivity of consumption, rejected the PIH, and hence accepted *by default* the excess sensitivity of consumption. Only Flavin (1985) has tested the excess sensitivity of consumption as the null hypothesis; she shows that consumption is too sensitive to unemployment, not to current income. Since the excess sensitivity of consumption to income is not usually accepted as the null hypothesis, its role as the alternative hypothesis to the PIH must be re-examined. This paper tests the excess sensitivity of consumption itself as the null hypothesis rather than the alternative hypothesis and concludes that the deviation estimate becomes the smallest when λ, the fraction of agents who consume current income, is about 0.5.

This paper constructs an alternative excess smoothness test which is free from the unit root assumption in labor income. Deaton (1987), West (1988), Campbell and Deaton (1989), and Flavin (1993) conclude that consumption is too smooth to be justified by the PIH, assuming that labor income has a unit root. In other words, consumption does not respond enough to the new information on permanent income. First, this section resolves Quah's (1990) dispute of the claims of Campbell-Deaton and West. Campbell and Deaton show that even if econometricians do not observe all the variables agents observe, in theory econometricians can infer the omitted information from the observables. Quah presents an example where the omitted information in practice causes econometricians to incorrectly conclude that consumption is too smooth. Quah, however, incorrectly refutes Campbell and Deaton because he fails to recognize that the omitted information can be finessed in theory. The omitted information problem remains in practice even if it does not in theory. Second, this section presents a remedy to the main criticism of the literature on the excess smoothness of consumption: the unit root assumption in labor income. If labor income has a unit root, shocks to current labor income become permanent, and therefore, permanent income becomes more volatile. This is why consumption may appear too smooth compared to permanent income if we assume a unit root in labor income. Since the existence of unit roots is still a controversial issue, their conclusion is not yet convincing. This paper shows that consumption is, in fact, too smooth regardless of the unit root problem.

Ⅱ. EXCESS SENSITIVITY OF CONSUMPTION

We examine if consumption is too sensitive to current income to be consistent with the PIH. This section tests the hypothesis of excess sensitivity of consumption to current income as a null hypothesis as opposed to all the possible alternative hypotheses. Numerous authors such as Flavin (1981), Hayashi (1982), and Campbell and Mankiw (1989, 1990, 1991) have found that consumption is more sensitive to current income than is warranted by the PIH. They examine the PIH as the null hypothesis against the excess sensitivity of consumption as the alternative hypothesis. In general, they reject the PIH and therefore by default accept the excess sensitivity of consumption. This conclusion, however, needs more justification. In particular, this conclusion does not imply that consumption is too sensitive to current income; all it says is that the PIH does not hold. If one includes other variables, one might as well claim that consumption is more sensitive to those other variables as Flavin (1985) finds. In other words, consumption may deviate from the excess sensitivity of consumption as much as from the PIH itself. That is why we need to examine the excess sensitivity of consumption as a null hypothesis rather than an alternative hypothesis and measure the extent to which consumption deviates from the hypothesis of excess sensitivity of consumption. These ideas can be explored by the signal extraction approach originally developed by Durlauf and Hall (1988, 1989a, 1989b) and extended by Kim (1996).

We define a general consumption function which includes the excess sensitivity of consumption hypothesis, the PIH and the Keynesian consumption hypothesis. Thus, this paper can compare deviations from the PIH with deviations from the Keynesian consumption hypothesis and the excess sensitivity of consumption hypothesis. This exercise can therefore shed light on the old debate of which hypothesis is more consistent with the data.

Define the following information structures. $L_x(t)$ = Linear space generating information available to econometricians at t. $L_\phi(t)$ = Linear space generating information available to the representative agent at t. $L_\delta(t)$ = Linear space generating information orthogonal to the representative agent's forecast error at t. We denote the projection of Y_t onto $L_\xi(s)$ as $Y_{t|s}(\xi)$. We assume $L_x(t) \subseteq L_\phi(t)$. Namely, econometricians do not observe all the information agents observe.

Suppose there are two groups of agents who receive labor income y^l_{1t}, y^l_{2t} and capital income y^k_{1t}, y^k_{2t}. Total labor income y^l_t is the sum of the labor incomes of these two groups: $y^l_t = y^l_{1t} + y^l_{2t}$, and total capital income is the sum of the individual capital incomes: $y^k_t = y^k_{1t} + y^k_{2t}$. Let us assume that the first group receives a fixed share λ of the total labor income, so $y^l_{1t} = \lambda y^l_t$ and $y^l_{2t} = (1 - \lambda) y^l_t$ for $0 \leq \lambda \leq 1$, and that the first group consumes their current income whereas the second group consumes their permanent income.[1] Thus, $C_{1t} = y^k_{1t} + y^l_{1t}$, and $C_{2t} = y^p_{2t}$ where

$$y^p_{2t} \equiv y^k_{2t} + \left(\frac{r}{1+r}\right) \sum_{i=0}^{\infty} \left(\frac{1}{1+r}\right)^i E(y^l_{2t+i} \mid \Phi_t).$$

The budget constraint holds for each group: for $i = 1, 2$,[2]

$$y^k_{it+1} = (1 + r)y^k_{it} + r(y^l_{it} - C_{it}).$$

Note that $y^k_{1t} = \bar{y}^k_1$ for all t since the first group neither saves nor dissaves, and therefore their capital stock level does not change.[3][4] The individual budget constraints imply that the budget constraint for the aggregate agent holds:

$$y^k_{t+1} = (1 + r)y^k_t + r(y^l_t - C_t).$$

We define a general consumption function as, for $0 \leq \lambda \leq 1$,

$$c^e_t(\lambda) = C_{1t} + C_{2t} = y^k_{1t} + y^l_{1t} + y^p_{2t} = \lambda Y_t + (1 - \lambda) Y^p_t \tag{2.1}$$

where the aggregate income Y_t and the aggregate permanent income Y^p_t are defined as:

$$Y_t = y^k_t + y^l_t,$$

$$Y^p_t \equiv y^k_t + \left(\frac{r}{1+r}\right) \sum_{i=0}^{\infty} \left(\frac{1}{1+r}\right)^i E(y^l_{t+i} \mid \Phi_t).$$

---

[1] Kim (1995) finds that λ is not constant since the fraction of agents who consumes current income should vary negatively with income if agents are liquidity constrained. We simply assume λ is constant in this paper to conform to the past literature.

[2] A more familiar budget constraint would be
$W_{t+1} = (1 + r) W_t + y^l_t - C_t$,
where W_t is real non-human wealth. Since this paper wishes to use the data on capital income, the budget constraint is formulated in terms of capital income where $y^k_t = rW_t$.

[3] This model differs from Campbell and Mankiw (1989, 1990, 1991) only in that the ratio of labor incomes, instead of total incomes, is constant between the two groups. Since the capital income for the first group is constant, the constant ratio of total incomes puts an unnecessary restriction on the labor income processes. Further, this assumption enables us to write consumption as a convex combination of current total income and permanent income as in (2.1).

[4] This model is a rough proxy for the liquidity constrained behavior since the liquidity constrained agents would eventually consume all of the capital stock.

Note that $\lambda = 0$ corresponds to the PIH, $\lambda = 1$ corresponds to the Keynesian consumption hypothesis and $0 < \lambda < 1$ corresponds to the excess sensitivity of consumption hypothesis. Therefore, we can measure the extent to which consumption deviates from each hypothesis.

Define for $0 \leq \lambda \leq 1$,

$$Y_t^* \equiv y_t^t + \left(\frac{r}{1+r}\right) \sum_{i=0}^{\infty} \left(\frac{1}{1+r}\right)^i y_{t+i}^l,$$

$$c_t^*(\lambda) \equiv \lambda Y_t + (1 - \lambda) Y_t^*. \tag{2.2}$$

Y_t^* is the *ex post* realization of Y_t^p, and c_t^* is the *ex post* realization of c_t^e. By construction,

$$E(Y_t^* \mid \Phi_t) = Y_t^p,$$

and

$$E(c_t^* \mid \Phi_t) = c_t^e.$$

The null and the alternative hypotheses are, for $0 \leq \lambda \leq 1$:

$$H_0: C_t = c_t^e(\lambda)$$

$$H_1: C_t = c_t^e(\lambda) + S_t(\lambda),$$

where $S_t(\lambda)$ is a specification error or deviation from the general consumption function. $S_t(\lambda)$ contains all the testable implications of the general consumption function. As long as $S_t(\lambda) \neq 0$, the general consumption function does not hold for λ. This approach has no need to specify any specific alternative hypothesis; the alternative hypothesis is any behavior other than that predicted by the general consumption function. The model then becomes, for $0 \leq \lambda \leq 1$,

$$C_t = c_t^e(\lambda) + S_t(\lambda), \tag{2.3}$$

$$c_t^*(\lambda) = c_t^e(\lambda) + (1 - \lambda) v_t, \tag{2.4}$$

where $Y_t^* = Y_t^p + v_t$, and v_t is the aggregate representative agent's forecast error $: v_{t|t}(\phi) = 0$. This becomes a signal extraction problem because the number of observables, C_t, $c_t^*(\lambda)$, are less than the number of underlying components, $c_t^e(\lambda)$, $S_t(\lambda)$, v_t. Our approach finds optimal smoothing estimates of $S_t(\lambda)$. In other words, an econometrician extracts the deviation estimate of $S_t(\lambda)$ from the vantage point of $t = \infty$. The optimal smoothing estimates differ from the optimal prediction estimates in that the former utilizes all the information available up to time ∞ whereas the later utilizes the information only up to time t.

Let us rearrange (2.3) and (2.4) in terms of observables and unobservables, and divide by C_t:

$$1 - \frac{c_t^*(\lambda)}{C_t} = \frac{S_t(\lambda)}{C_t} - (1 - \lambda)\frac{v_t}{C_t}. \tag{2.5}$$

If there is no deviation, the deviation-consumption ratio, $\frac{S_t(\lambda)}{C_t}$, will be zero. Thus, this paper measures the extent to which $\frac{S_t(\lambda)}{C_t}$ deviates from 0 or the null hypothesis. The deviation-consumption ratio will be 0 if the general consumption function holds exactly and will be 1 if $C_t = S_t(\lambda)$. The ratio measures the extent to which the general consumption function fails to explain the consumption series. This paper will find the optimal smoothing estimate of $\frac{S_t(\lambda)}{C_t}$. If we regress $1 - \frac{c_t^*(\lambda)}{C_t}$ onto $L_\delta(t)$, which is orthogonal to the aggregate representative agent's forecast errors v_t, we can obtain $\frac{S_t(\lambda)}{C_t}{}_{\mid t}(\delta)$.

$L_\delta(t)$, of course, includes all the variables known to the aggregate representative agent at time t since they are orthogonal to the agent's forecast errors by construction. Further, there exists a subset of $L_\delta(t)$ which may not be known to the aggregate representative agent at *time* t, but is known to econometricians at *time* ∞.[5] In other words, we wish to obtain an optimal smoothing estimate rather than an optimal prediction estimate. The agents' forecast error, v_t, is AR(1) under the null hypothesis:

$$v_t - \left(\frac{1}{1+r}\right)v_{t+1} = \left(\frac{r}{1+r}\right)\sum_{i=0}^{\infty}\left(\frac{1}{1+r}\right)^{i+1}[E(y^l_{t+1+i} \mid \Phi_{t+1}) - E(y^l_{t+1+i} \mid \Phi_t)]$$
$$\equiv \frac{1}{1+r}\varepsilon_{t+1}$$

where ε_t is the information innovation for the aggregate representative agent at time t. It implies that $v_t = \sum_{i=1}^{\infty}\left(\frac{1}{1+r}\right)^i\varepsilon_{t+i}$, and therefore $\{\varepsilon_i\}_{i=0}^{t}$ is orthogonal to v_t. Namely, the information innovation is known to the representative agent at

[5] Note that $L_\pi(t) \subseteq L_\phi(t)$ does not imply $L_\pi(\infty) \subseteq L_\phi(t)$.

time t. Thus, $L_\delta(t)$ should include all the information available to econometricians at time t and the current and past information innovations for the representative agent:

$$L_\delta(t) \supseteq L_x(t) \oplus L_\varepsilon(t).$$

The information innovations under the null hypothesis can be obtained from (2.1) in changes:

$$\Delta C_t = \lambda \Delta y_t^l + (1-\lambda)\left(\frac{r}{1+r}\right)\sum_{i=0}^{\infty}\left(\frac{1}{1+r}\right)^i [E(y_{t+i}^l \mid \Phi_t) - E(y_{t+i}^l \mid \Phi_{t-1})] \quad (2.6)$$
$$= \lambda \Delta y_t^l + (1-\lambda)\varepsilon_t$$

where $\Delta C_t \equiv C_t - C_{t-1}$. Under the null hypothesis, $\Delta C_t - \lambda \Delta y_t^l$ is the information innovation for the aggregate representative agent at time t and should be included in $L_\delta(t)$. Thus, if $L_\delta(t)$ contains current and past C_t, y_t^l, then the optimal prediction estimate becomes identical to the optimal smoothing estimate. $L_\delta(t)$ thus includes all variables known to econometricians, including, at least, current and past consumption and labor income.

$\dfrac{S_t(\lambda)}{C_t}\bigg|_t(\delta)$ can be obtained by projecting equation (2.5) onto $L_\delta(t)$. This paper computes for $0 \leq \lambda \leq 1$,

$$\text{Measure } A: \sqrt{E\left[\frac{S_t(\lambda)}{C_t}\bigg|_t(\delta)\right]^2}$$

and

$$\text{Measure } B: E\left[\,\left|\frac{S_t(\lambda)}{C_t}\bigg|_t(\delta)\right|\,\right].$$

Since $\dfrac{S_t(\lambda)}{C_t}$ can be negative, this paper has to either square it or take an absolute value of it. They can give us some sense of how significant the model deviation is. This paper finds λ which minimizes $\dfrac{S_t(\lambda)}{C_t}\bigg|_t(\delta)$ since λ cannot be estimated in levels.

Kim (1996) shows that Measure A is a minimum bound for $E\left[\frac{S_t(\lambda)}{C_t}\right]^2$. Also, Measure A can be decomposed into variance and mean of $\frac{S_t(\lambda)}{C_t}{}_{|t}(\delta)$:

$$E\left[\frac{S_t(\lambda)}{C_t}{}_{|t}(\delta)\right]^2 = Var\left[\frac{S_t(\lambda)}{C_t}{}_{|t}(\delta)\right] + \left[E\frac{S_t(\lambda)}{C_t}{}_{|t}(\delta)\right]^2.$$

The failure of the model can be characterized not only in terms of its variations but also its levels. Even when the deviation has a big level with small variations, measure A can detect the deviation while the past literature would have missed it. The decomposition can also indicate how slow the deviation process is.

The signal extraction approach has three advantages over the past literature such as Campbell and Mankiw (1989, 1990, 1991), Flavin (1981, 1985) and Hayashi (1982). First, the signal extraction approach finds λ in levels whereas the past literature estimates λ in changes. Although (2.6) can be estimated in changes by instrumental variables, the estimates may be misleading since estimating in changes misses a slow moving deviation. Second, our approach tests all the testable implications of the excess sensitivity of consumption hypothesis whereas the past literature examines only one specific implication. For example, (2.6) does not imply (2.1). Third, the past literature does not test (2.6) as the null hypothesis as opposed to the alternative hypothesis. In particular, λ could become insignificant if other variables are included on the right hand side of (2.6). We test (2.1) as the null hypothesis against all the possible alternative hypotheses.

Campbell and Mankiw (1989, 1990, 1991) build a model similar to the above economy except that they assume the first group receives a fixed share of total income, not total labor income. They derive

$$\Delta C_t = \mu + \lambda\Delta Y_t + (1-\lambda)\varepsilon_t, \tag{2.7}$$

and reconcile the rejection of the PIH with the claim that 35 - 65 percent of agents consume their current income, not permanent income. They estimate λ with instrumental variables, find $\lambda \neq 0$, and reject the PIH with nondurable consumption. Flavin (1981) also finds $\lambda = 0.355$ in (2.7). Since she uses nondurable consumption and the ratio of nondurable consumption to total consumption is 0.45, she claims it is significant departure from the PIH. Hall and Mishkin (1982) find similar conclusions with micro data. Hayashi (1982) assumes that the discount rate for the future labor income is not equal to the rate of return from nonhuman wealth due to the risk premium. He accepts the PIH if consumption is based on the expenditures of nondurables and services and imputed service flows from consumer durables; he rejects the PIH if consumption includes expenditures on consumer durables and excludes service flows from consumer durables.

For purposes of comparison, this paper uses the seasonally adjusted quarterly data (in 1972 dollars) for the period 1953:2 -1984:4 from Blinder and Deaton (1985). Blinder and Deaton make several adjustments to the National Income and Product Accounts (NIPA) and break down real disposal income into the capital and labor components. All the series are on a real per capita basis. See Blinder and Deaton for a detailed description. This paper uses only consumption of nondurables and services, assuming that the agent's utility function is separable between durable goods and nondurable goods. Therefore, consumption of nondurables and services can be analyzed separately from durable consumption. Since λ is the fraction of agents who consume current income, we need to scale consumption of nondurables and services by the mean ratio of total consumption to consumption of nondurables and services, which is 1.2737 in our data set.[6] This paper sets the sample range from 1954:3 to 1984:4 and sets $\frac{1}{1+r} = 0.99$ as Campbell (1987). This corresponds to 4.04 percent on an annual basis. The test results are robust to the reasonable value of interest rates as Campbell finds.

The data is exponentially detrended throughout this section. For the representative agent model to make any sense, the data has to be detrended since the representative agent model does not address the issue of growth when the discount rate is equal to the interest rate. Even with growth in labor income, the representative agent model dictates that consumption should be a random walk without a drift under the PIH. The consumption data simply rejects a random walk without a drift if we do not detrend it. On the other hand, detrending may cause erroneous inferences, as Mankiw and Shapiro (1985) show that erroneous detrending will lead to excessive rejection of the PIH when the data contains a unit root. The erroneous inferences would, however, lead to higher deviation estimates than the true deviation in our framework and would make our estimates more conservative.

$\{c_t^*\}_{t=0}^{T}$ needs to be constructed from the finite observable data series. As Shiller (1981) calculates P_t^*, (2.2) implies

$$c_t^*(\lambda) = \frac{1}{1+r} c_{t+1}^* (\lambda) + \lambda(Y_t - \frac{1}{1+r} Y_{t+1}) + (1-\lambda)(y_t^k - \frac{1}{1+r} y_{t+1}^k + \frac{r}{1+r} y_t^l). \qquad (2.8)$$

Let c_T^* represents a terminal value which this paper sets equal to the endpoint of the series. Then $\{c_t^*\}_{t=0}^{T-1}$ can be constructed recursively by (2.8). Kim's (1996)

[6] Using the ratio of the mean of total consumption to the mean of consumption of nondurables and services, which is 1.2786, we obtain similar results and omit them here.

Monte Carlo simulations show that the finite sample problem for the use of the terminal value which Flavin (1983) and Kleidon (1986) point out does not cause a significant problem in this procedure.

For $0 \leq \lambda \leq 1$, (2.5) is regressed onto $\frac{y_t}{C_t}, \frac{y_t^l}{C_t}$ for lags 0 to 5. Table 1 reports the results for lag 5 which is chosen by Akaike's information criterion. The second and third columns present Measure A and Measure B. The minimum deviation is 1.72 % at $\lambda = 0.5$ with Measure A, and 1.40 % at $\lambda = 0.6$ with Measure B.[7] We believe that these deviations are small enough to be negligible, considering numerous assumptions and simplifications we made for the sake of tractability of the model. We argue that the excess sensitivity of consumption hypothesis tracks the US data very well. These are also comparable to the past estimates for λ. The Keynesian consumption hypothesis with $\lambda = 1.0$ has deviations of 2.84% with Measure A and 2.41% with Measure B. The PIH with $\lambda = 0$ has deviations of 3.22% with Measure A and 2.60% with Measure B. Although we do not claim that the Keynesian consumption hypothesis fits the US data better than the PIH, our results definitely suggest that consumption is more sensitive to current income than the PIH implies. The fourth and fifth columns present the mean-variance decomposition of the deviation. Deviation results mostly from its variations with lower λ while deviation results more from its levels with higher λ. The sixth column presents the χ^2 statistics of excluding all regressors, using the Newey and West's (1987) heteroskedasticity autocorrelation consistent covariance matrix estimator. They are all significant at 1% level. In other words, if we used standard statistics, we would have rejected all the model even if their deviations are very small. That is why we need to measure deviation rather than simply accept or reject the model.

Note that this section describes tests of whether consumption is too sensitive to current income, not of whether some fraction of agents are liquidity constrained or myopic. Although it may be plausible to argue that consumption is more sensitive to current income since agents are liquidity constrained, a formal model is needed to link the two. Flavin (1985) estimates

$$\Delta C_t = \mu + \lambda \Delta Y_t + (1-\lambda)\varepsilon_t + \gamma \Delta z_t,$$

where z_t is unemployment and a proxy for the presence of liquidity constraints. She finds that by including z_t, the estimate of λ becomes insignificant, and concludes that agents are not myopic, but liquidity constrained. Kim (1995) also

[7] Although this paper does not claim that $\lambda = 0.5$ (Measure A =1.72%) is statistically better than, for example, $\lambda = 0.7$ (Measure A =1.87%), these deviation estimates provide us with measures as to how good or bad the hypothesis fits the data.

[Table 1] Measuring Deviations in the Excess Sensitivity of Consumption

λ	$E\left[\left\| \frac{S_t}{C_{t\|t}}(\delta) \right\|\right]$	$\sqrt{E\left[\frac{S_t}{C_{t\|t}}(\delta)\right]^2}$	$Var\left[\frac{S_t}{C_{t\|t}}(\delta)\right]$	$\left(E\left[\frac{S_t}{C_{t\|t}}(\delta)\right]\right)^2$	χ^2
0.0	0.0260	0.0322	0.0010	0.0000	42.9349
0.1	0.0230	0.0281	0.0008	0.0000	36.9485
0.2	0.0204	0.0243	0.0006	0.0000	31.6517
0.3	0.0180	0.0210	0.0004	0.0000	27.4043
0.4	0.0158	0.0185	0.0003	0.0000	24.6876
0.5	0.0142	0.0172	0.0002	0.0001	24.1358
0.6	0.0140	0.0173	0.0002	0.0001	26.5713
0.7	0.0157	0.0187	0.0001	0.0002	33.0356
0.8	0.0182	0.0213	0.0002	0.0003	44.8053
0.9	0.0210	0.0246	0.0002	0.0004	63.3775
1.0	0.0241	0.0284	0.0003	0.0005	90.4068

shows that consumption is too sensitive to current income because agents are liquidity constrained not because agents are myopic. If agents are in fact liquidity constrained, we should observe that more agents are liquidity constrained during recessions whereas less agents are liquidity constrained during booms. Thus, we should observe negative correlations between λ, the fraction of agents who consume current income, and output fluctuations. If agents are simply myopic, on the other hand, the two variables should be independent. Kim (1995) finds that a higher fraction of agents consume their current income during recessions.

Ⅲ. EXCESS SMOOTHNESS OF CONSUMPTION

Numerous authors such as Deaton (1987), Campbell and Deaton (1989), West (1988) and Flavin (1993) have shown that consumption is too smooth to be justified by the PIH. In other words, consumption does not respond enough to new information on permanent income. Consumption is smooth not because the PIH holds but because the PIH does not hold. The past excess smoothness of consumption literature, however, has one major short-coming. The excess smoothness of consumption holds only if labor income has a unit root, which is still a controversial assumption. This section therefore presents an alternative excess smoothness test which is free from the unit root assumption and shows that consumption is in fact too smooth.

First, we resolve the dispute of whether the omitted information may have caused consumption to be too smooth. Deaton (1987), Campbell and Deaton (1989) and West (1988) show that consumption is too smooth to be justified by the PIH, assuming a unit root in the labor income series. Quah (1990), however, presents an example where the omitted information causes econometricians to conclude incorrectly that consumption is too smooth. He assumes that agents observe the permanent and transitory components in the labor income fluctuations. If the labor income fluctuations have only permanent components, we know that consumption is too smooth because all shocks become permanent. If they have only transitory components, on the other hand, we know that consumption is not too smooth because all shocks are temporary and will die off. Quah, thus, decomposes the labor income fluctuations into permanent and transitory components in such a ratio that would generate a degree of smoothness which is consistent with the PIH.

Quah (1990), however, is incorrect to claim that econometricians cannot recover the underlying innovations since they are not fundamental. Quah uses incorrect variables: Quah should have used either saving and labor income, or consumption, labor income and capital income, but not consumption and labor income since $L_{\{C_t, y^l_t\}}(t) \neq L_{\{l_t, y^l_t\}}(t)$ and $L_{\{C_t, y^l_t\}}(t) \subseteq L_{\{C_t, y^l_t, y^k_t\}}(t)$ where l_t is saving. It is true, as Quah (1990) claims, that agents observe innovations that are not fundamental for the joint consumption-labor income process. If agents study the joint process of $\{C_t, y^l_t, y^k_t\}$ or $\{l_t, y^l_t\}$, however, they can observe innovations which are fundamental. Since the VAR with consumption and labor income is not invertible, Quah incorrectly claims that Campbell and Deaton are wrong. The appendix shows that Quah uses incorrect variables and that the VAR with consumption and capital income is, in fact, invertible. In other words, econometricians can recover the underlying innovations.

The reason that Campbell and Deaton (1989) and Quah (1990) reach different conclusions is that gap exists between theory and practice. Quah concludes that consumption is not too smooth by specifying the very large ARIMA processes (about 200 MA lags) whereas Campbell and Deaton conclude that consumption is too smooth by including one or five MA lags. Even if Campbell and Deaton can recover the omitted information in theory, they cannot recover 200 MA lags in practice since they can always find shorter MA lags which approximate the true processes. That is why Campbell and Deaton fail to recover the true model.

Our approach to testing the excess smoothness of consumption has three advantages over the past literature. First, the data need not be detrended linearly or exponentially since all the variables used are already stationary. For the representative agent model to make any sense, the data has to be detrended since the representative agent model does not address the issue of growth when the discount rate is equal to the interest rate. Even with growth in labor income, consumption should be a random walk without a drift under the PIH. The consumption data simply rejects a random walk without a drift if we do not detrend it. Detrending, however, may cause spurious inferences. This is the dilemma faced by past tests. Our approach escapes the dilemma of whether or not to detrend the data. Second, we show that nondurable consumption is too smooth regardless of whether or not the labor income process has a unit root, which is the main criticism of the excess smoothness of consumption literature. If labor income has a unit root, shocks to current labor income become permanent, and therefore, permanent income becomes more volatile. If labor income has no unit roots, on the other hand, shocks to current labor income are only temporary, and therefore, permanent income becomes less volatile. Thus, consumption may appear too smooth compared to permanent income if we assume a unit root in labor income. Third, the omitted information problem may not cause any problem in our approach. Flavin (1993) criticizes Campbell and Deaton (1989) in that the omitted information may arise under some alternative hypothesis. Flavin shows that the omitted information problem disappears only with the excess sensitivity of consumption as a possible alternative hypothesis. Without restricting the class of possible alternative hypotheses, our approach is free from the omitted information problem.

Under the PIH, agents consume:

$$C_t = y_t^k + \left(\frac{r}{1+r}\right) \sum_{i=0}^{\infty} \left(\frac{1}{1+r}\right)^i E(y_{t+i}^l \mid \Phi_t) \tag{3.1}$$

where the right hand side is the permanent income. Also, the budget constraint for the agent is[8] :

$$y_{t+1}^k = (1+r)y_t^k + r(y_t^l - C_t).$$

Define

$$c_t^* \equiv y_t^k + \left(\frac{r}{1+r}\right) \sum_{i=0}^{\infty} \left(\frac{1}{1+r}\right)^i y_{t+i}^l. \tag{3.2}$$

Note c_t^* is *not* the perfect foresight consumption as P_t^* in LeRoy and Porter (1981) and Shiller (1981) since consumption should be constant in perfect foresight due to the first order condition: $E(C_{t+1} \mid \Phi_t) = C_t$. c_t^* is simply the *ex post* realization of C_t. $c_t^* - y_t^k$ would be a correct analogy to P_t^*. Under the PIH,

$$c_t^* = C_t + v_t \tag{3.3}$$

where v_t is the representative agent's forecast error: $v_{t\mid t}(\Phi_t) = 0$.

Divide (3.3) by total income y_t to transform nonstationary series into station-

[8] See Kim (1996) for a full description of the model.

ary series:

$$\frac{c_t^*}{y_t} = \frac{C_t}{y_t} + \frac{v_t}{y_t},$$

where $0 \leq \frac{C_t}{y_t} \leq 1$.

Proposition:

If the coefficient of the regression of $\frac{C_t}{y_t} - \frac{c_t^*}{y_t}$ onto $\frac{C_t}{y_t}$ is less than 0, consumption is too smooth under the PIH.

Proof:

$$\frac{C_t}{y_t} - \frac{c_t^*}{y_t} = -\frac{v_t}{y_t}.$$

Under the PIH,

$$Var\left(\frac{c_t^*}{y_t}\right) = Var\left(\frac{C_t}{y_t} + \frac{v_t}{y_t}\right) = Var\left(\frac{C_t}{y_t}\right) + Var\left(\frac{v_t}{y_t}\right),$$

since v_t is orthogonal to all the information at time t. Under the PIH, $Var\left(\frac{v_t}{y_t}\right)$ can be computed:

$$Var\left(\frac{v_t}{y_t}\right) = Var\left(\frac{c_t^*}{y_t} - \frac{C_t}{y_t}\right).$$

Therefore,

$$Var\left(\frac{C_t}{y_t}\right) = Var\left(\frac{c_t^*}{y_t}\right) - Var\left(\frac{c_t^*}{y_t} - \frac{C_t}{y_t}\right).$$

Consumption is too smooth if

$$Var\left(\frac{C_t}{y_t}\right) < Var\left(\frac{c_t^*}{y_t}\right) - Var\left(\frac{c_t^*}{y_t} - \frac{C_t}{y_t}\right). \tag{3.4}$$

(3.4) is equivalent to

$$1 < \frac{Cov\left(\frac{c_t^*}{y_t}, \frac{C_t}{y_t}\right)}{Var\left(\frac{C_t}{y_t}\right)}$$

The above ratio is a simply $1 - \gamma$ in the following regression.

$$\frac{C_t}{y_t} - \frac{c_t^*}{y_t} = \alpha + \gamma \frac{C_t}{y_t} + u_t.$$

Thus, we can instead test if $\gamma = 0$. If $\gamma < 0$, then the average propensity to consume is too smooth. *Q. E. D.*

$\{c_t^*\}_{t=0}^{T}$ is constructed as follows : (3.2) implies

$$c_t^* = \frac{1}{1+r} c_{t+1}^* + y_t^k - \frac{1}{1+r} y_{t+1}^k + \frac{r}{1+r} y_t^l,$$

where $c_T^* = C_T$. Table 2 rejects the null hypothesis of $\gamma = 0$ against the alternative hypothesis $\gamma < 0$ at 1 percent. Thus, this paper concludes that nondurable consumption seems to be too smooth. $\alpha = 0$ is also rejected, which suggests that $\frac{C_t}{y_t}$ is not the optimal forecast of $\frac{c_t^*}{y_t}$, and therefore the PIH does not seem to hold. This paper finds that the average propensity of consumption is smoother than the PIH dictates, without the unit root assumption in labor income.

[Table 2] Excess Smoothness Test

α	γ
5.0440 (0.3164)	−5.6350 (0.3397)

Standard errors are in parentheses.

Ⅳ. CONCLUSION

We characterize the failure of the PIH in terms of excess sensitivity and excess smoothness. This paper relies on the signal extraction approach to compute all the testable implications of the excess sensitivity of consumption. The signal extraction approach examines whether the model is reasonable, not whether the model is exactly correctly specified. The hypothesis of excess sensitivity of consumption to current income is tested as a null hypothesis as opposed to an alternative hypothesis. We indeed find that consumption is too sensitive to current income. Also, this paper finds that consumption is too smooth whether or not the labor income series has a unit root. This seemingly contradictory characterizations in fact point to the same phenomenon as Campbell and Deaton (1989) first pointed out. They find little evidence to support that permanent income is smoother than current income. Consumption is smooth not because the PIH holds, but because consumption adjusts slowly to new information in permanent income. Since agents do not adjust to new information instantaneously, consumption becomes more sensitive to current income. Therefore, the excess sensitivity and the excess smoothness are two sides of the identical phenomenon.

The PIH and the Keynesian consumption hypothesis are not as far apart as they are perceived. If labor income is a random walk, the PIH (3.1) becomes the Keynesian consumption hypothesis: $C_t = y_t$. If the liquidity constraint is very severe, the PIH becomes very similar to the Keynesian consumption hypothesis. Thus, even if agents are rational and forward looking, consumption may well follow the Keynesian consumption hypothesis.

APPENDIX

This appendix shows that if econometricians' information set includes $\{C_t, y_t^l, y_t^k\}$ or $\{l_t, y_t^l\}$, they could recover the omitted information.

Consumption may appear too smooth with the omitted information problem whether or not the labor income series has a unit root. Let ε_t^ϕ be the orthogonal information innovations to agents at time t, and ε_t^x be the orthogonal information innovations to econometricians at time t:

$$\varepsilon_t^\phi = L_\phi(t) - L_\phi(t-1),$$

$$\varepsilon_t^x = L_x(t) - L_x(t-1).$$

The true model is

$$\Delta C_t = \varepsilon_t^\phi \tag{A.1}$$

and implies $Var(\Delta C_t) = Var(\varepsilon_t^\phi)$. $Var(\varepsilon_t^\phi)$ is, however, unknown, and $Var(\varepsilon_t^x)$ is substituted instead. Since West (1988) shows that $Var(\varepsilon_t^x) > Var(\varepsilon_t^\phi)$, consumption may appear too smooth with the omitted information: $Var(\Delta C_t) < Var(\varepsilon_t^x)$.

Campbell and Deaton (1989), noting the omitted information problem, show that (A.1) implies with projection argument,

$$\Delta C_t = \varepsilon_t^x \tag{A.2}$$

Let $L_x(t)$ include C_t, y_t^l and y_t^k. Then, project (3.1) onto $L_x(t)$, and we can get

$$C_t = y_t^k + \left(\frac{r}{1+r}\right) \sum_{i=0}^{\infty} \left(\frac{1}{1+r}\right)^i E(y_{t+i}^l \mid X_t). \tag{A.3}$$

(A.3) implies (A.2). In other words, under the null hypothesis, ΔC_t is a signalling variable which correctly encapsulates the agents's information set if econometricians observe C_t, y_t^l and y_t^k.[9] Thus, they conclude that consumption is, in fact, too smooth.

Quah (1990) formulates the first difference of labor income as:

$$\Delta y_t^l = A_1(L)\varepsilon_t^1 + (1-L)A_0(L)\varepsilon_t^0,$$

where $A_i(z) \equiv \sum_{k=0}^{\infty} a_i^k z^k \neq 0$ for $|z| \leq 1$, $a_i^0 = 1$, ε_t^i is serially uncorrelated, and ε_t^1 is the permanent component and ε_t^0 is the transitory component. It implies

$$\Delta C_t = A_1\left(\frac{1}{1+r}\right)\varepsilon_t^1 + \left(1 - \frac{1}{1+r}\right) A_0\left(\frac{1}{1+r}\right)\varepsilon_t^0,$$

and

$$\Delta y_t^k = \frac{r}{1+r} L\left[\frac{A_1(L) - A_1\left(\frac{1}{1+r}\right)}{\frac{1}{1+r} - L}\right]\varepsilon_t^1$$

$$+ \frac{r}{1+r} L\left[\frac{(1-L)A_0(L) - \left(1 - \frac{1}{1+r}\right) A_0\left(\frac{1}{1+r}\right)}{\frac{1}{1+r} - L}\right]\varepsilon_t^0.$$

---

[9] Campbell and Deaton use saving and labor income instead. As Campbell (1987) shows, (3.1) implies *saving for a rainy day*:

$$l_t = -\sum_{i=0}^{\infty}\left(\frac{1}{1+r}\right)^i E(\Delta y_{t+i}^l \mid \Phi_t).$$

Project the above equation onto $L_x(t)$ where $L_x(t)$ includes l_t and y_t^l. Then it becomes

$$l_t = -\sum_{i=0}^{\infty}\left(\frac{1}{1+r}\right)^i E(\Delta y_{t+i}^l \mid X_t).$$

That implies

$$l_t - (1+r)l_{t-1} - \Delta y_t^l = \left(\frac{r}{1+r}\right)\sum_{i=0}^{\infty}\left(\frac{1}{1+r}\right)^i [E(y_{t+i}^l \mid X_t) - E(y_{t+i}^l \mid X_{t-1})].$$

Thus, l_t and y_t^l are signalling variables if econometricians observe l_t and y_t^l.

The determinant of the matrix moving average polynomial is

$$Det(z) =$$

$$\frac{r}{1+r} z \left[\frac{A_0(z) A_1\left(\frac{1}{1+r}\right)(1-z) - A_0\left(\frac{1}{1+r}\right) A_1(z)\left(1-\frac{1}{1+r}\right)}{\frac{1}{1+r} - z} \right],$$

and the inside of the bracket is $\frac{0}{0}$ when $z = \frac{1}{1+r}$. Thus, using the L'Hopital's rule, the inside of the bracket becomes

$$-A_0'(z) A_1\left(\frac{1}{1+r}\right)(1-z) + A_0(z) A_1\left(\frac{1}{1+r}\right) + A_0\left(\frac{1}{1+r}\right) A_1'(z)\left(1-\frac{1}{1+r}\right),$$

and does not seem to vanish at $z = \frac{1}{1+r}$. Thus, VAR with consumption and capital income is invertible and, in theory, the underlying innovation can be recovered from consumption and capital income. Campbell and Deaton are correct to assert that consumption or saving is the signalling variable for the omitted information.

REFERENCES

Blinder, A.S., and A. Deaton. 1985. "The time series consumption function revisited," *Brookings Papers on Economic Activity*, Vol. 2, pp.456-511.

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

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

Campbell, J.Y., and N.G. Mankiw. 1989. "Consumption, income, and interest rates : Reinterpreting the time series evidence," *NBER Macroeconomics Annual*, edited by O.J. Blanchard and S. Fischer, pp.185-216. Cambridge, MA:MIT Press.

Campbell, J.Y., and N.G. Mankiw. 1990. "Permanent income, current income, and consumption," *Journal of Business and Economic Statistics* 8(3), pp. 265-279.

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

Deaton, A.S. 1987. "Life-cycle models of consumption : Is the evidence consistent with the theory?" in Bewley, T. F., editor, *Advances in Econometrics,* Fifth World Congress, Vol. II, pp. 121-148.

Durlauf, S.N., and R.E. Hall. 1988. "Bounds on the variances of specification errors in models with expectations," Working Paper, Stanford University.

Durlauf, S.N., and R.E. Hall. 1989a. "A signal extraction approach to recovering noise in expectations based models," Working Paper, Stanford University.

Durlauf, S.N., and R.E. Hall. 1989b. "Measuring noise in stock prices," Working Paper, Stanford University.

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

Flavin, M.A. 1983. "Excess volatility in the financial markets: a reassessment of the empirical evidence," *Journal of Political Economy* 91, pp.929-956.

Flavin, M.A. 1985. "The excess sensitivity of consumption to current income - Liquidity constraints or myopia ?" *Canadian Journal of Economics*, pp. 117-136.

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

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

Hall, R.E. and F.S. Mishkin. 1982. "The sensitivity of consumption to transitory income : estimates from panel data on households," *Econometrica* 50, pp.

461-481.

Hayashi, F. 1982. "The permanent income hypothesis: estimation and testing by instrumental variables," *Journal of Political Economy* 90, pp. 895-916.

Kim, C. 1996. "Measuring deviations from the permanent income hypothesis," *International Economic Review* 37(1), pp.205-225.

Kim, C. 1995. "Asymmetry in Consumption," mimeo, Rutgers University.

Kleidon, A. 1986. "Variance Bounds Tests and Stock Price Volatility," *Journal of Political Economy* 94, pp.953-1001.

LeRoy, S.F., and R.D. Porter. 1981. "The present-value relation: Tests based on implied variance bounds," *Econometrica* 49(3), pp.555-574.

Newey, W.K., and K.D. West. 1987. "A simple positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix," *Econometrica* 55(3), pp.703-708.

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

Shiller, R.J. 1981. "Do stock prices move too much to be justified by subsequent changes in dividends ?" *American Economic Review* LXXI, pp.421-436.

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