Budget Constraints and Time-Series Evidence on Consumption

By JORDI GALÍ*

Is consumption more or less variable than predicted by the permanent-income hypothesis (PIH)? To answer that question, I develop a procedure based on a long-run restriction implied by the consumer's intertemporal budget constraint. In contrast to previous work, the approach here (i) does not require any assumptions on the stochastic properties of labor income, (ii) does not impose restrictions on the consumer's information set, and (iii) is robust to departures from the PIH model. The application of the procedure to postwar U.S. data suggests that consumption is smoother than the PIH model predicts. (*JEL* 131)

Recent work on consumption has focused on econometric tests of the restrictions implied by the permanent-income hypothesis (PIH). According to that hypothesis, the observed time-series for aggregate consumption is the result of intertemporal optimization by a forward-looking representative consumer who can borrow and lend without constraints at the market interest rate.

More specifically, under the assumptions[1] found in standard versions of the PIH model (henceforth, the "standard PIH model"), expected utility maximization implies that consumption should equal permanent income, the latter being defined as the annuity value of the sum of nonhuman wealth and the expected present value of future labor income (see e.g., Marjorie Flavin, 1981). Formally,

$$(1)\quad c_t = y_t^{\mathrm{P}} \equiv r\left[W_t + (1+r)^{-1}\sum_{j=0}^{\infty}(1+r)^{-j}E_t y_{t+j}\right]$$

where c is consumption, y^{P} is permanent income, r is the (constant) return on nonhuman wealth, W is nonhuman wealth, and y is labor income. E_t is the expectational operator conditional on all the information available to the representative consumer at time t.

A great deal of empirical work has been devoted to assessing the variability of consumption relative to the variability implied by the PIH model. The analysis typically focuses on variance ratios of the form

$$\psi \equiv \left[\frac{\mathrm{Var}(\Delta c)}{\mathrm{Var}(\xi)}\right]^{1/2}$$

where Δ is the usual first-difference operator and ξ denotes the innovation in permanent income, defined by

$$\xi_t \equiv y_t^{\mathrm{P}} - E_{t-1}y_t^{\mathrm{P}} = r(1+r)^{-1}\sum_{j=0}^{\infty}(1+r)^{-j}(E_t - E_{t-1})y_{t+j}.$$

*Graduate School of Business, 607 Uris Hall, Columbia University, New York, NY 10027. I am grateful to Ricardo Caballero, two anonymous referees, and workshop participants at Columbia, Princeton, and the Sixth World Congress of the Econometric Society for helpful comments. I am solely responsible for all errors and misinterpretations.

[1]Namely, an infinite-lived representative consumer, quadratic and time-separable preferences, nondurability, a constant return on nonhuman wealth equal to the time discount rate, and absence of shocks to preferences.

As is well known (Flavin, 1981; Angus Deaton, 1987), the standard PIH model implies $\Delta c_t = \Delta y_t^p = \xi_t$. Accordingly, ψ should be equal to 1. Thus, estimates of ψ significantly different from 1 should be interpreted as evidence against the standard PIH model. Furthermore, a value of ψ smaller (greater) than 1 would suggest that consumption is too smooth (too volatile) relative to permanent income.

Unfortunately, estimation of the variance ratio ψ is not trivial, since neither permanent income y^p nor its innovation ξ is observable. The main purpose of this paper is to suggest a new approach to identification and estimation of $\mathrm{Var}(\xi)$ and ψ. Even though different procedures aiming at a similar goal have been developed in the literature, those procedures are correct only (i) under the assumption that no variable has predictive power for future labor income, beyond that of its current and past values (e.g., Flavin, 1981; Deaton, 1987; Francis Diebold and Glenn Rudebusch, 1991) or, alternatively, (ii) under the null hypothesis that the standard PIH model holds (e.g., Kenneth West, 1988; John Campbell and Deaton, 1989). Unfortunately, both assumptions have been rejected econometrically using postwar U.S. data, so no method seems to be available to estimate variance ratios like ψ.

The approach to identification and estimation of ψ developed here relies neither on (i) nor (ii). Instead, it is based on the restrictions imposed by the budget constraint on the low-frequency properties of aggregate consumption. Furthermore, the approach of the present paper does not impose any restrictions on the order of integration or other stochastic properties of aggregate labor income.

Before introducing the method, I will briefly discuss the two basic approaches found in the literature.

A. *The Univariate Approach*

This approach relies on the assumption that some transformation of labor income follows a univariate ARMA process driven by a disturbance ε, with the latter being interpreted as the true innovation in labor income. Formally, a model of the following form is assumed:

$$\phi(L)\, y_t^* = \theta(L)\varepsilon_t$$

with

$$\begin{aligned}\varepsilon_t &\equiv y_t^* - P[\,y_t^* | y_{t-1}^*, y_{t-2}^*, \dots] \\ &= y_t^* - E_{t-1} y_t^*\end{aligned}$$

where P is the orthogonal projection operator and where y^* denotes a "suitable transformation" of aggregate labor income such that $\{y_t^*\}$ is a stationary stochastic process. As is made explicit by the above equality, under this approach, expectations are assumed to coincide with orthogonal projections on lagged values of y^*.

In practice, alternative transformations have been used by different authors: deterministic detrending (Flavin, 1981), first differences (Deaton, 1987), log first differences (Campbell and Deaton, 1989 part I), and fractional differences (Diebold and Rudebusch, 1991). In all cases, however, an expression of the form $\mathrm{Var}(\xi) = F(\phi, \theta, r)\mathrm{Var}(\varepsilon)$ holds. In words, the variance of the permanent-income innovation is proportional to the variance of the univariate process residual, with the factor of proportionality F being a well-defined function[2] of the interest rate and the AR and/or MA coefficients vectors (ϕ, θ) of that univariate process. Given $F(\cdot)$, a value for r, and consistent estimates $\hat{\phi}$, $\hat{\theta}$, and $\widehat{\mathrm{Var}(\varepsilon)}$, a consistent estimate of $\mathrm{Var}(\xi)$ is given by $F(\hat{\phi}, \hat{\theta}, r)\widehat{\mathrm{Var}(\varepsilon)}$. From there, estimation of ψ follows trivially.

In practice, however, the estimates of $\mathrm{Var}(\xi)$ and the corresponding estimates of ψ thus obtained are very sensitive to the assumption made on the order of integration of labor income (i.e., on the particular

[2] Of course, the form of $F(\cdot)$ depends on the specific transformation made.

transformation of that variable used). Most noticeably, the use of detrended labor income leads to estimates of ψ greater than 1 ("excess variability" in consumption), whereas the assumption of a unit root generates estimates of the same statistic that are smaller than 1 ("excess smoothness" in consumption); see Deaton (1987) for a comparison of the two results.

A more fundamental weakness of the univariate approach, however, lies in its implicit assumption that no variable other than current and past values of labor income can help predict future labor income.[3] Alternatively, consumers are assumed to form their expectations of future labor income on the basis of a restricted information set, containing only current and past values of that variable. Both assumptions are, to begin with, intuitively unappealing. More importantly, the evidence that savings, among other variables, Granger-causes labor income (Campbell, 1987) makes both assumptions clearly untenable.[4]

B. *The West and Campbell-Deaton Approaches*

West (1988) and Campbell and Deaton (1989) developed alternative procedures for estimating $\text{Var}(\xi)$ without imposing any restriction on the consumer's information set or the predictability of labor income. However, as is made clear in their papers (and particularly stressed by Flavin [1988]), the estimates of $\text{Var}(\xi)$ thus obtained are only correct under the null hypothesis that the standard PIH model holds. Consequently, they should only be used for the purpose of testing that hypothesis. The outcome of those tests, carried out by the above-mentioned authors, systematically rejects the PIH null hypothesis, as a result of a ψ estimate significantly smaller than 1. Interestingly, that very outcome implies that the estimates of $\text{Var}(\xi)$ and ψ obtained by both the West and the Campbell-Deaton procedures, though sufficient to reject the PIH null hypothesis, are no longer "admissible" estimates of the true values of $\text{Var}(\xi)$ and ψ. In other words, those estimates do not contain any information on the behavior of consumption, other than the observation that the latter does not satisfy the PIH model. In particular, estimates of ψ less than 1 are not (necessarily) evidence of the presence of excess smoothness in consumption nor are they good measures of that eventual excess smoothness (i.e., good measures of the true ψ).

C. *This Paper's Approach: Basic Features and Outline*

As mentioned above, the present paper develops an alternative approach to identification and estimation of the variance ratio ψ. As in West (1988) and Campbell and Deaton (1989), my approach does not require any restriction on the (unobservable) information set used by consumers to form their expectations about future income. Its main advantage relative to the West and Campbell-Deaton approaches lies in its robustness to a variety of departures from the standard PIH model. As is shown below, the variance ratio ψ is identified using a restriction imposed by the consumer's intertemporal budget constraint on the low-frequency properties of the consumption time-series. To the extent that the budget constraint is met regardless of whether the PIH holds or not[5] and, in the latter case, regardless of the nature of the departures from the PIH, identification of $\text{Var}(\xi)$ and ψ can be achieved outside the PIH null.

The approach of this paper shows an additional advantage over previous methods: it does not rely on any assumption

[3] In other words, labor income is assumed to be "Granger-causally prior" to any other variable.

[4] West (1988) shows that the estimate of $\text{Var}(\xi)$ based on the assumption of a univariate labor-income process will systematically overstate the variance of true innovations in permanent income and thus introduce a downward bias in the estimates of ψ.

[5] In other words, the present paper uses the assumption that the intertemporal budget constraint is met as an identifying restriction. This contrasts with Lars Hansen et al. (1990), which focuses on possible overidentifying (and thus testable) restrictions implied by that constraint.

about the stochastic properties of the labor-income process. That overcomes two problems. First, it is not necessary to make any assumption on the order of integration of labor income. This is particularly important given the difficulties in determining the appropriate "transformation" of that variable (i.e., its true order of integration) and the large differences in terms of the variability of consumption predicted by the PIH across empirical models based on alternative "transformations." Second, the approach here allows for a labor-income process with different components whose innovations are observable to consumers but not to the econometrician, thus overcoming the "nonfundamentalness" problem pointed out by Danny Quah (1990).

The outline of the paper is as follows. Section I develops a basic framework, which is capable of nesting a variety of departures from the PIH. Section II shows how the intertemporal budget constraint can be used to identify both the variance of permanent-income innovations and the variance ratio ψ, suggesting at the same time alternative ways to proceed with estimation of those parameters. In the same section, I apply those estimation procedures to postwar U.S. data and suggest a "spectral" interpretation of the results. Section III uses the consumption framework developed in Section I to address an additional issue: the relationship between measures of consumption variability like ψ and the evidence of predictability of consumption changes. Section IV summarizes and concludes.

I. A Simple Framework for the Time-Series Analysis of Consumption

The model assumes an infinite-lived representative consumer facing, as of time 0, a sequence of dynamic budget constraints of the form

$$W_{t+1} = W_t(1+r) + y_t - c_t \tag{2}$$

where the different variables are defined as in the Introduction. Unless it is otherwise specified, all the equations in this section hold for $t = 0, 1, 2, \dots \infty$.

The consumer's transversality condition takes the form

$$\lim_{T \to \infty} W_T(1+r)^{-T} \geq 0. \tag{3}$$

Assuming nonsatiation, (2) and (3) make it possible to obtain the intertemporal budget constraint:

$$r(1+r)^{-1} \sum_{j=0}^{\infty} (1+r)^{-j} E_t c_{t+j} = y_t^{\mathrm{p}}$$

where y^{p} is defined as in (1).

Applying the law of iterated expectations, one obtains

$$r(1+r)^{-1} \sum_{j=0}^{\infty} (1+r)^{-j} (E_t - E_{t-1}) c_{t+j} = \xi_t \tag{4}$$

where $\xi_t \equiv y_t^{\mathrm{p}} - E_{t-1} y_t^{\mathrm{p}}$ is, as above, the innovation in permanent income. In words, (4) states that the present discounted value of revisions in expectations of future consumption is equal to the innovation in total wealth, ξ / r. Note that this result follows from the representative consumer's intertemporal budget constraint, and it involves no assumptions such as the nature of preferences, goods durability, the relationship between r and the time-discount rate, presence or absence of liquidity constraints, shocks to preferences, or the stochastic properties of the labor-income process.

A (rather weak) assumption is made on the nature of the time-series for aggregate consumption: aggregate consumption, on a per capita basis, is assumed to be a difference-stationary [i.e., $I(1)$] process. That assumption is consistent with any model in which households adjust their expected level of consumption at all future horizons in response to changes in permanent income. In most models, that feature results from households' willingness to smooth consumption over time and is independent of the stochastic properties of labor income (and, in particular, of its order of integration).[6] Empirically, the hypothesis of a unit root in the consumption time-series corresponding to the postwar United States is not rejected by standard unit-root tests, even at high significance levels (e.g., Campbell, 1987; Galí, 1990).

Consumption change (Δc) is modeled as a stationary stochastic process with two components: a "fundamental" component (Δc^*), which depends on current and past innovations in permanent income, and a "noise" component (n), unrelated to those innovations. Formally,

$$(5) \qquad \Delta c_t = \Delta c_t^* + n_t$$

where $\Delta c_t^* = \alpha + \sum_{j=0}^{\infty}\beta_j \xi_{t-j}$ and $n_t = \sum_{j=0}^{\infty}\varphi_j \eta_{t-j}$.

Thus, two different sources of consumption movements are allowed for: shocks to permanent income (ξ) and shocks to the noise component (η). By construction $E_t \xi_{t+j} = 0$ for any $j > 0$, which in turn implies that the ξ's are serially uncorrelated ($E\xi_i \xi_j = 0$ for $i \neq j$) and uncorrelated with lagged η's ($E\xi_t \eta_{t-j}$ for $j > 0$ and all t). I further assume that shocks to the noise component are serially uncorrelated ($E\eta_i \eta_j = 0$ for $i \neq j$) and uncorrelated with contemporaneous or past innovations in permanent income ($E\eta_t \xi_{t-j}$ for $j \geq 0$ and all t).[7]

The model above allows for two general types of deviations from the PIH: (i) the presence of lags in the response of consumption to innovations in permanent income and (ii) the presence of changes in consumption unrelated to news about current or future income. As is shown below, (i) nests a variety of specific consumption models that allow for durability, time nonseparabilities, and so on. On the other hand, (ii) can be seen as a flexible way of capturing shocks to preferences or, simply, suboptimal or myopic behavior.

Under the consumption model given by (5), equation (4) can be rewritten (after some manipulation) as follows:

$$\xi_t \sum_{j=0}^{\infty} (1+r)^{-j}\beta_j + \eta_t \sum_{j=0}^{\infty} (1+r)^{-j}\varphi_j = \xi_t.$$

Since the previous condition must hold for any realizations of ξ_t and η_t, it follows that

$$(6) \qquad \sum_{j=0}^{\infty} (1+r)^{-j}\beta_j = 1$$

$$(7) \qquad \sum_{j=0}^{\infty} (1+r)^{-j}\varphi_j = 0.$$

The previous constraints have an intuitive interpretation: if the intertemporal budget constraint is to be satisfied, then sooner or later (i) the level of consumption must be adjusted to the new level of permanent income, and (ii) changes in consumption unbacked by corresponding changes in permanent income must be undone. The interest rate, together with the time pattern of the adjustment, determines in each case the size of the permanent effects on the level of consumption.

Equations (5), (6), and (7) yield a rather general characterization of consumption, which conforms to a variety of models. This is illustrated next, by means of some examples.

Example 1: The Standard PIH Model.—This model, as originally developed in Robert Hall (1978), is characterized by quadratic utility, an interest rate equal to the discount rate, and an absence of preference shocks. It corresponds to the model above with $\beta_0 = 1$, $\alpha = \beta_j = \varphi_i = 0$ for $i = 0, 1, 2, \ldots$ and $j = 1, 2, \ldots$.

Example 2: Durable Goods.—This model was studied by Gregory Mankiw (1982). It

[6]Some authors have argued that the logarithm of consumption—and not consumption—is difference-stationary (e.g., Campbell and Deaton, 1989). This issue is not likely to be important in practice: both $\Delta \log c$ and Δc show very similar features, at least in postwar U.S. data. Intertemporal budget constraints and most models' predictions are, however, expressed in terms of levels, not in terms of logs.

[7]Notice that the orthogonality assumption "defines" the noise component as a residual component (i.e., as the component of consumption changes uncorrelated with permanent income innovations at all leads and lags).

can be easily shown to correspond to $\beta_0 = (1+r)/(r+\delta)$, $\beta_1 = -(1-\delta)(1+r)/(r+\delta)$, and $\alpha = \beta_j = \varphi_i = 0$ for $i = 0,1,2,\ldots$ and $j = 1,2,\ldots$, where δ denotes the rate of depreciation of consumption goods.

Example 3: Habit Formation.—A simple habit-formation model is obtained by assuming a form of time nonseparability given by a (quadratic) utility function $U(c_t - \tau c_{t-1})$, with $|\tau| < 1$. Such a model yields a process for consumption changes as in (5), with $\beta_j = [1-(\tau/1+r)]\tau^j$, $\alpha = \varphi_j = 0$ for $j = 0,1,2,\ldots$. A version of this model is exposited in Deaton (1987).

Example 4: "λ" Model.—Following Campbell and Mankiw (1989), it is assumed that a fraction λ of consumers set consumption equal to labor income, maybe as a result of binding liquidity constraints. The remaining fraction $(1-\lambda)$ of consumers behave according to the PIH model. In addition, it is assumed that Δy follows an AR(1) model with autoregressive coefficient ρ and unconditional mean μ. As r approaches 0, one can show that aggregate consumption in this case will be characterized by the model above with $\alpha = \lambda\mu$, $\beta_0 = (1-\lambda\rho)$, and $\beta_j = \lambda(1-\rho)\rho^j$, $j \geq 1$.

Obviously, any of the previous models can be "augmented" with a noise component n, in order to capture shocks to preferences or, more generally, consumption changes caused by factors other than changes in permanent income. Such a noise component can take an infinite number of forms; yet the coefficients of its moving-average representation must satisfy restriction (7). An example often found in the literature (e.g., Flavin, 1981) is given by $\varphi_0 = 1$, $\varphi_1 = -(1+r)$, and $\varphi_j = 0$ for $j = 2,3,\ldots$.

Similarly, a variety of departures from the standard PIH model can generate a nonzero (and presumably positive) value for α, the drift term in consumption. Among them are models with an interest rate different from the time-discount rate, models allowing for precautionary savings (e.g., Ricardo Caballero, 1990), and models with finite horizons (e.g., Richard Clarida, 1988; Galí, 1990).

In summary, the model of consumption given by (5), (6), and (7) is quite general, its main restrictions being the assumptions of linearity and time-invariance and the requirement that the consumer meet his intertemporal budget constraint. As will be seen below, that model will be useful in the identification and estimation of ψ.

II. Budget Constraints and Consumption Variability

A. *Budget Constraints and Identification of Variance Ratios*

In this section, I show how the restrictions implied by the budget constraint on the consumption time-series allow one to identify the variance ratio ψ and to construct a consistent estimator for it. As will become clear below, it is useful to introduce a second variance ratio γ, defined by the following transformation of ψ:

$$\gamma \equiv \left[\frac{\mathrm{Var}(\xi)}{\mathrm{Var}(\Delta c)}\right] \equiv \left[\frac{1}{\psi}\right]^2.$$

The discussion is simplified by working in the frequency domain. Let $h_x(\omega) \equiv (2\pi)^{-1}\sum_{s=-\infty}^{\infty} R_x(s)\exp(-i\omega s)$ be the spectral density of a stationary random sequence $\{x_t\}_{t=-\infty}^{\infty}$, with $R_x(s) \equiv E[(x-Ex)(x_{-s}-Ex)]$ being the autocovariance of x at lag s. The next lemma is the basis of much of the analysis below.

LEMMA: *Under the consumption model given by* (5), (6), *and* (7), *the following results obtain*:

$$\text{(i)} \quad \lim_{r \to 0}\left[2\pi h_{\Delta c^*}(0) - \mathrm{Var}(\xi)\right] = 0$$

$$\text{(ii)} \quad \lim_{r \to 0} h_n(0) = 0.$$

(See Appendix for the proof.)

By applying a simple continuity argument to the proof, it can easily be seen that $2\pi h_{\Delta c^*}(0) \cong \mathrm{Var}(\xi)$ and $h_n(0) \cong 0$ will be "good" approximations, given the relevant value ranges for r and assuming that the β's and the φ's converge to zero "fast enough." The Appendix shows that for plausible parameter values the magnitude of the error associated with these approximations will generally be small.

Orthogonality at all leads and lags between Δc and n implies $h_{\Delta c}(\omega) = h_{\Delta c^*}(\omega) + h_n(\omega)$ for all $0 < |\omega| < \pi$, and the following result follows directly from the Lemma:

$$\lim_{r \to 0} \left[2\pi h_{\Delta c}(0) - \mathrm{Var}(\xi)\right] = 0. \tag{8}$$

Let $f_{\Delta c}(\omega) \equiv h_{\Delta c}(\omega)/\mathrm{Var}(\Delta c)$ be the normalized spectral density of Δc. Applying the above definition of γ to (8), one obtains

$$\lim_{r \to 0} \left[2\pi f_{\Delta c}(0) - \gamma\right] = 0 \tag{9}$$

where $2\pi f_{\Delta c}(0) = [1 + 2\sum_{s=1}^{\infty} \rho_{\Delta c}(s)]$, with $\rho_{\Delta c}(s)$ $[\equiv R_{\Delta c}(s)/\mathrm{Var}(\Delta c)]$ denoting the autocorrelation of consumption changes at lag s. Thus, as the interest rate approaches zero, the variance ratio γ converges to the normalized spectral density of consumption changes at frequency zero.

B. *Estimation of Variance Ratios γ and ψ*

Given plausibly low interest rates, it follows from (9) that consistent estimators for $2\pi f_{\Delta c}(0)$ qualify as consistent estimators for the variance ratio γ. The variance ratio ψ can be consistently estimated using the relationship $\psi \equiv \gamma^{-1/2}$. As argued above, this class of estimators is robust to alternative assumptions on the stochastic properties of labor income, as well as to any departures from the standard PIH model that can be accommodated by the general consumption framework given by (5), (6), and (7).

The strategy pursued here involves, in a first stage, constructing an estimator for $f_{\Delta c}(\omega)$.[8] As is well known (e.g., Priestley, 1981), given N observations $\Delta c_1, \Delta c_2, \ldots, \Delta c_N$, the (normalized) spectral density $f_{\Delta c}(\omega)$ can be consistently estimated using

$$\hat{f}_{\Delta c}(\omega) = (2\pi)^{-1} \sum_{s=-(N-1)}^{N-1} \lambda(s)\hat{\rho}(s)\cos(s\omega) \tag{10}$$

where

$$\hat{\rho}_{\Delta c}(s) = \frac{\sum_{t=s+1}^{N} (\Delta c_t - \overline{\Delta c})(\Delta c_{t-s} - \overline{\Delta c})}{\sum_{t=1}^{N} (\Delta c_t - \overline{\Delta c})^2}$$

$$\overline{\Delta c} = \left(\frac{1}{N}\right) \sum_{t=1}^{N} \Delta c_t$$

and where $\{\lambda(s)\}$ is a real even function of s, usually referred to as the "lag window." Alternative "lag windows" can be used for estimation purposes, each having sample properties which usually depend on the "true" spectral density. In this paper, I use two types of windows:

(i) the "Bartlett window" $\{\lambda_B(s)\}$, given by

$$\lambda_B(s) = \begin{cases} 1 - |s|/M & |s| \le M \\ 0 & |s| > M \end{cases}$$

(ii) the "Bartlett-Priestley window" $\{\lambda_P(s)\}$, given by

$$\lambda_P(s) = \begin{cases} 1 & |s| = 0, 2M, 4M, \ldots \\ \dfrac{3M^2}{(\pi s)^2}\left[\dfrac{\sin(\pi s/M)}{\pi s/M} - \cos(\pi s/M)\right] & \text{otherwise,} \end{cases}$$

The choice of the Bartlett window was motivated by its simplicity as well as its

[8]The statistical issues involved in the estimation of the spectrum are discussed in M. B. Priestley (1981), among others. Recent applications of spectral-density estimation to macroeconomics can be found in John Cochrane (1988) and Campbell and Deaton (1989). The latter paper uses a similar spectrum estimator to evaluate the spectral density of labor-income changes at frequency zero. As is clear from their derivation, that measure can be interpreted as an estimate of $\mathrm{Var}(\xi)$ only under the PIH null hypothesis (which their procedure itself rejects).

widespread use among macroeconomists (e.g., Whitney Newey and West, 1987; Cochrane, 1988; Campbell and Deaton, 1989). The Bartlett-Priestley window, which corresponds to a weighted integral of the periodogram with weights given by a quadratic function, can be shown to minimize the relative mean-square error of the estimated spectrum (see Priestley, 1981 section 6.2.3).

Both windows are controlled by parameter M. Consistency of the estimator in (10) requires that $M \to \infty$ and $(M/N) \to 0$, as $N \to \infty$. In practice, however, a finite value of M has to be chosen, which involves a trade-off between the bias and variance of $\hat{f}_{\Delta c}(\omega)$ [of order $O(1/M)$ and $O(M/N)$, respectively]. In the Appendix, I show that the value of M that minimizes the relative mean-square error of $\hat{f}_{\Delta c}(0)$ in the case at hand is likely to be relatively small for both windows, probably less than 10. Nevertheless, below I report estimates based on a wide range of M values, in order to guarantee the robustness of the results.

C. *Empirical Results*

The data were obtained from the Citibase tape and correspond to the U.S. National Income and Product Accounts measure of quarterly consumption of nondurables and services, expressed in 1982 dollars, seasonally adjusted, and on a per capita basis. The sample period is 1947:1–1988:3.

Table 1 reports the estimates of γ and ψ (denoted by $\hat{\gamma}$ and $\hat{\psi}$) based on both the Bartlett and the Bartlett-Priestley estimators of $\hat{f}_{\Delta c}(0)$ for different M values. Standard errors for $\hat{\gamma}$ (in parentheses) are based on estimates of the asymptotic variance of $\hat{f}_{\Delta c}(0)$.[9] Standard errors for $\hat{\psi}$ (also in parentheses) were computed using the delta method. I also report asymptotic P values corresponding to one-sided tests of the null hypothesis H_0: $\gamma = \psi = 1$, against the alternative H_A: $\gamma > 1$ (or equivalently, $\psi < 1$). The P values are computed under the assumption of asymptotic normality of $\hat{\gamma}$.[10]

The point estimates of ψ thus obtained are systematically less than 1. Their order of magnitude is robust to the type of window used and the values of the window parameter M considered. The asymptotic P values for values of M between 2 and 10 point toward a rejection of the PIH null hypothesis (against the excess-smoothness alternative) at conventional significance levels. Given that the number of observations is fixed, increases in the value of M necessarily increase the standard errors of the estimates, thus raising the significance level at which the PIH null can be rejected. Interestingly, the point estimates are not significantly altered by the size of the window: most of them lie within a narrow range, around a value of 0.6. That result is also robust to the type of window used. As argued above, and in contrast with the estimates found in the existing literature, the $\hat{\psi}$ value obtained here can be interpreted as a consistent estimate of the true underlying variance ratio ψ, even if the PIH null hypothesis is rejected.

Though consistent, both the Bartlett and Bartlett-Priestley spectrum estimates are generally biased in small samples. Accordingly, the estimates of γ and ψ are also likely to be biased, since they are based on those spectrum estimates. Table 1 reports the expected value of $\hat{\gamma}$ and $\hat{\psi}$ under the assumption of normal, serially uncorrelated consumption changes, as approximated by a Monte Carlo simulation.[11] Interestingly, un-

[9]The asymptotic variance of $\hat{f}_{\Delta c}(0)$ is given by $(4/3)(M/N)f_{\Delta c}^2(0)$ for the "Bartlett" estimator and $(12/5)(M/N)f_{\Delta c}^2(0)$ for the "Bartlett-Priestley" estimator. Estimates were obtained by replacing $f_{\Delta c}(0)$ with its estimate $\hat{f}_{\Delta c}(0)$ (see Priestley, 1981 section 6.2.4.).

[10]The normality assumption is based on the result of asymptotic normality of $\hat{f}(\omega)$, under some regularity conditions. In small samples, the assumption of normality must be seen only as a convenient approximation. Notice also that the normality of $\hat{\gamma}$ implies that $\hat{\psi}$ cannot be normally distributed; accordingly, the latter's standard errors cannot be used to construct the usual t tests.

[11]The Monte Carlo results are based on 300 simulated series of normal, independent and identically distributed consumption changes, with 167 observations each and with mean and variance given by their sample counterparts.

TABLE 1—ESTIMATES OF γ AND ψ

Window size (M)	$\hat{\gamma}$	$\hat{\psi}$	P value	Monte Carlo results (PIH) $E(\hat{\gamma})$	$E(\hat{\psi})$	$\hat{\psi}_{0.05}$
A. *Bartlett Window*:						
2	1.241	0.897	0.125	0.989	1.008	0.946
	(0.157)	(0.057)				
3	1.415	0.840	0.059	0.983	1.013	0.928
	(0.220)	(0.065)				
4	1.619	0.785	0.033	0.975	1.020	0.912
	(0.291)	(0.070)				
5	1.784	0.748	0.028	0.965	1.028	0.906
	(0.358)	(0.075)				
6	1.876	0.730	0.033	0.958	1.034	0.897
	(0.413)	(0.080)				
7	1.962	0.713	0.039	0.952	1.040	0.890
	(0.666)	(0.084)				
8	2.053	0.697	0.043	0.944	1.046	0.882
	(0.522)	(0.088)				
9	2.104	0.689	0.051	0.936	1.054	0.868
	(0.567)	(0.092)				
10	2.151	0.681	0.059	0.929	1.061	0.866
	(0.611)	(0.096)				
20	2.369	0.649	0.150	0.869	1.128	0.838
	(0.952)	(0.130)				
30	2.355	0.651	0.342	0.815	1.197	0.823
	(1.951)	(0.160)				
40	2.503	0.632	0.290	0.758	1.266	0.814
	(1.423)	(0.179)				
50	2.638	0.615	0.328	0.699	1.353	0.812
	(1.677)	(0.195)				
B. *Bartlett-Priestley Window*:						
2	1.459	0.827	0.064	0.985	1.014	0.917
	(0.248)	(0.070)				
3	1.749	0.755	0.040	0.976	1.024	0.904
	(0.365)	(0.078)				
4	1.990	0.708	0.039	0.965	1.035	0.883
	(0.480)	(0.085)				
5	2.096	0.690	0.052	0.952	1.047	0.881
	(0.565)	(0.093)				
6	2.208	0.672	0.063	0.942	1.058	0.871
	(0.652)	(0.099)				
7	2.337	0.654	0.072	0.929	1.071	0.868
	(0.745)	(0.104)				
8	2.351	0.652	0.091	0.914	1.084	0.858
	(0.802)	(0.111)				
9	2.402	0.645	0.106	0.902	1.095	0.853
	(0.869)	(0.116)				
10	2.448	0.639	0.120	0.894	1.104	0.845
	(0.933)	(0.121)				
20	2.413	0.643	0.277	0.802	1.237	0.824
	(1.301)	(0.173)				
30	2.540	0.627	0.358	0.715	1.403	0.821
	(1.678)	(0.207)				
40	2.733	0.604	0.405	0.640	1.630	0.815
	(2.085)	(0.230)				
50	2.905	0.586	0.441	0.574	1.918	0.808
	(2.477)	(0.250)				

Notes: Numbers in parentheses are standard errors. The data were obtained from the Citibase tape and correspond to the U.S. National Income and Product Accounts measures of quarterly consumption of nondurables and services, expressed in 1982 dollars and on a per capita basis. The sample period is 1947:1–1988:3. See text for a description of the statistics.

der the PIH null hypothesis, the small sample bias tends to generate values of ψ greater than 1, (i.e., "excess volatility" of consumption).[12] The last column in Table 1 gives the simulated 5-percent critical value corresponding to a one-sided test of the PIH null hypothesis against the excess-smoothness alternative. Interestingly, when the small-sample adjusted critical values are used, the estimates reject the PIH null hypothesis at the 5-percent level for all window sizes considered. Of course, these results only strengthen the evidence of "excess smoothness" reported previously.

Most of the $\hat{\psi}$ estimates lie in the range between 0.6 and 0.8. Given the sign of the small-sample bias, that range should probably be seen as an upper-bound range for the true variance ratio ψ. That implies that the variability of consumption is less than 60–80 percent of the corresponding variability predicted by the PIH model. That figure can be compared to those obtained by other authors using different approaches. Thus, for instance, the estimate of ψ obtained under the assumption of a unit root in labor income and a univariate process for Δy is 0.34 (Deaton, 1987). Benchmark estimates of the same ratio obtained by West (1988) and Campbell and Deaton (1989) (also under the maintained hypothesis of a unit root in labor income or its log) are 0.36 and 0.59, respectively.[13] Under the assumption of an ARFIMA(p,d,q) model for labor income, the estimate of ψ ranges from 0.88 (when $d=0.6$) to 0.10 (when $d=1.2$) (Diebold and Rudebusch, 1991).[14] The estimate obtained under the assumption of a trend-stationary labor income is 2.86 (Deaton, 1987). Thus, the fact that $\hat{\psi}$ takes a value that is less than 1 seems to support the basic conclusion of excess smoothness in consumption previously drawn by some of the researchers working under the assumption that labor income is integrated of order 1 (or a fraction close to 1), even though their estimators are not robust to departures from the null hypothesis or unrestricted information sets.

D. *A Spectral Interpretation of Variance Ratio ψ*

The variance ratio ψ can be given a simple interpretation, related to the spectrum of Δc. As is well known, the variance of Δc is represented by the area below the spectrum $h_{\Delta c}(\omega)$ [i.e., $\text{Var}(\Delta c)=\int_{-\pi}^{\pi} h_{\Delta c}(\omega)\,d\omega$]. On the other hand, the Lemma showed that, for r values close to zero, $\text{Var}(\xi)\cong 2\pi h_{\Delta c}(0)$. Hence, $\text{Var}(\xi)$ can be represented by the area of a rectangle with height $h_{\Delta c}(0)$ and width 2π.

Under the standard PIH model (i.e., if $\Delta c_t=\xi_t$ for all t), the equality $h_{\Delta c}(\omega)=h_{\Delta c}(0)$ would obtain for all ω. Thus, a measure of "excess smoothness" is given by the difference between the area of the $[h_{\Delta c}(0)\times\pi]$ rectangle and the area under the spectrum $h_{\Delta c}(\omega)$, $0\le\omega\le\pi$. This basic feature is illustrated in Figures 1 and 2. Figure 1 plots the estimate of $h_{\Delta c}(\omega)$, $0\le\omega\le\pi$, obtained using a Bartlett window with $M=5$. Figure 2 shows a similar estimate, now obtained with a Bartlett-Priestley window (also with $M=5$). Both figures also show the "PIH spectrum" described above. In both cases, the gap between the two areas is apparent, in a way consistent with the finding of excess smoothness. More interestingly, the plots suggest that such excess

[12] Under the PIH null hypothesis, $\rho(s)=0$ for $|s|>0$. Thus, under that null hypothesis, the bias in $\hat{f}(0)$ is a consequence of the bias in $\hat{\rho}(s)$ for $|s|>0$, not of the use of a window. Under some regularity conditions, $E\hat{\rho}(s)<0$, $|s|>0$ for a serially uncorrelated process with unknown mean (e.g., Wayne Fuller, 1976) p. 242). Thus, under the PIH null hypothesis, $\hat{\gamma}$ will be biased downward, and the expected value of $\hat{\psi}$ will be greater than 1. The larger is the window size M, the more autocorrelation terms are included, thus increasing the size of the bias.

[13] The West value corresponds to an AR(1) model for Δy (see West, 1988 p. 25). The Campbell-Deaton value corresponds to their VAR-1 model (Campbell and Deaton, 1989 p. 366). In both cases, (scaled up) data on consumption of nondurables and services are used.

[14] An ARFIMA(p,d,q) model for y is just an ARMA(p,q) model for $\Delta^d y$. Estimates were constructed using Diebold and Rudebusch's (1991) estimates of F (see their table 3; denoted as κ in their paper) together with the estimates $s(\varepsilon)=25.2$ and $s(\Delta c)=15.8$ reported in Campbell and Deaton (1989).

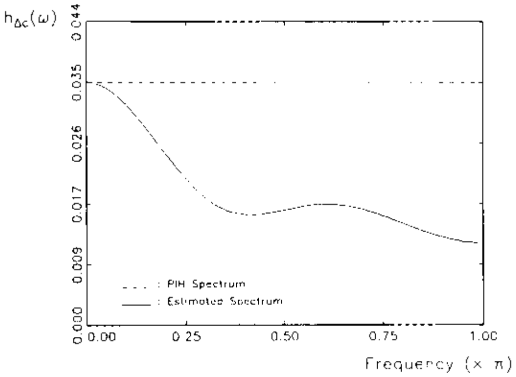

FIGURE 1. ESTIMATES OF $h_{\Delta c}(\omega)$ OBTAINED USING A BARTLETT WINDOW WITH $M=5$

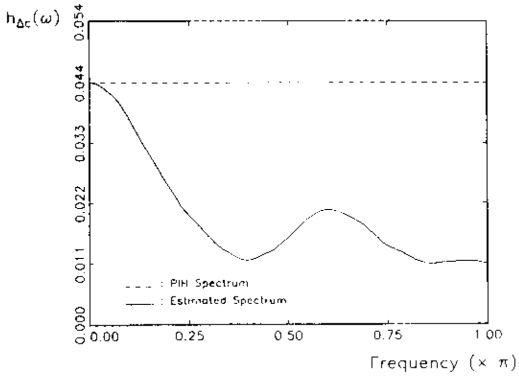

FIGURE 2. ESTIMATES OF $h_{\Delta c}(\omega)$ OBTAINED USING A BARTLETT-PRIESTLY WINDOW WITH $M=5$

smoothness holds "at all frequencies," that is, $h_{\Delta c}(\omega) \leq h_{\Delta c}(0)$, $0 < \omega \leq \pi$. The inequality becomes an equality at frequency zero, where the spectrum of Δc and the PIH spectrum must coincide if the intertemporal budget constraint is to be met.

The concept of excess smoothness "at all frequencies" is clearly stronger than the usual concept of excess smoothness defined as $\psi < 1$. In other words, nothing prevents the spectrum of Δc from being greater than $h_{\Delta c}(0)$ (the constant level of the PIH spectrum) at some frequency range, even if ψ is less than 1.

As an illustration of the previous point, consider the following model of consumption with both habit formation and noise:

$$\Delta c_t = \Delta c_t^* + n_t$$

$$(1-\tau L)\Delta c_t^* = [1-\tau/(1+r)]\xi_t$$

$$n_t = [1-(1+r)L]\eta_t.$$

As $r \to 0$, the spectrum corresponding to that process is given by

$$h_{\Delta c}(\omega) = (2\pi)^{-1}\mathrm{Var}(\xi) \times \left(\frac{(1-\tau)^2}{1-2\tau\cos\omega+\tau^2} + 2(1-\cos\omega)\Omega \right)$$

where $\Omega \equiv \mathrm{Var}(\eta)/\mathrm{Var}(\xi)$ measures the relative importance of noise innovations. The corresponding variance ratio ψ is given by

$$\psi = \left[\frac{1-\tau}{1+\tau} + 2\Omega \right]^{1/2}.$$

As can be easily checked, there exists a wide range of admissible values for τ and Ω for which $\psi < 1$, that is, for which consumption is smoother than the standard PIH model predicts and such that, for some range of frequencies, the spectrum of Δc is larger than $(2\pi)^{-1}\mathrm{Var}(\xi)$, the PIH spectrum. As an example, consider the previous model with $\tau = 0.6$ and $\Omega = 0.3$. The corresponding theoretical spectrum of Δc is shown in Figure 3. As is apparent, the area under the spectrum is smaller than the area under the PIH spectrum (also shown in Fig. 3). In fact, from the formula above, it is known that $\psi = 0.921$ for those parameter values. Yet, for a range of high frequencies, the spectrum of Δc lies above the PIH spectrum. Thus, in the previous sample, consumption is globally smoother than permanent income but shows, at the same time, "excess variability" at the highest frequencies.

As mentioned above, and in contrast with the previous example, the estimated spectrum of Δc suggests that aggregate consumption in the postwar United States shows excess smoothness "at all frequencies." Indeed, the shape of the estimated spectrum

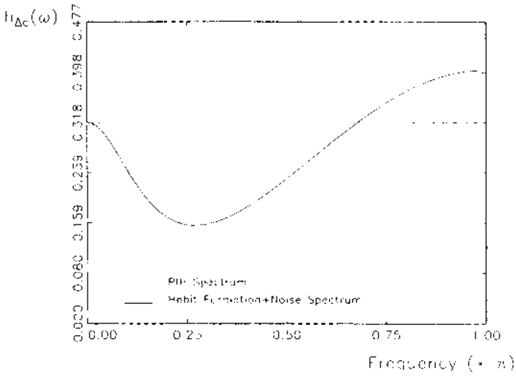

FIGURE 3. THE HABIT-FORMATION + NOISE SPECTRUM, WITH $\tau = 0.6$ AND $\Omega = 0.3$

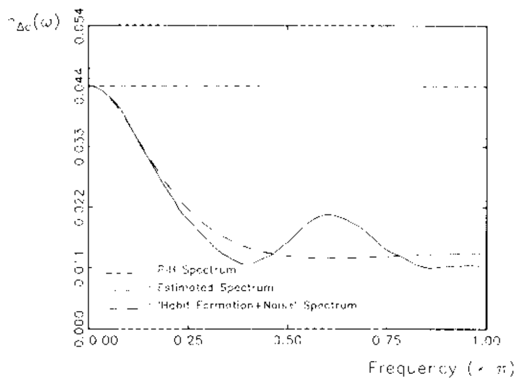

FIGURE 4. ESTIMATED SPECTRUM AND HABIT-FORMATION + NOISE SPECTRUM, WITH $\tau = 0.5$ AND $\Omega = 0.05$

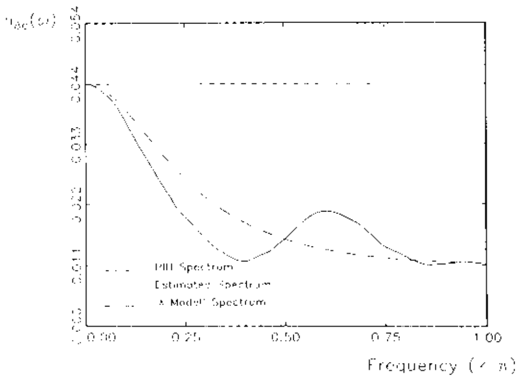

FIGURE 5. ESTIMATED SPECTRUM AND λ-MODEL SPECTRUM, WITH $\lambda = 0.8$ AND $\rho = 0.44$

matches quite well with the theoretical spectra of some simple models, properly calibrated. The latter point is illustrated in Figures 4 and 5. Figure 4 plots the estimated spectrum of Δc, together with the theoretical spectrum for Δc generated by the "habit formation + noise" model introduced above, with $\tau = 0.5$ and $\Omega = 0.05$. Figure 5 compares the estimated spectrum with the theoretical spectrum corresponding to the "λ" model found in Example 4 of Section I, with $\lambda = 0.8$ and $\rho = 0.44$.[15] Note that the value of λ used as a rough fit for the estimated spectrum is somewhat larger than estimates of that parameter found in the literature: Hall and Frederic Mishkin (1982) report estimates for λ around 0.2, while the estimates in Campbell and Mankiw (1989) range from 0.23 to 0.69.

The similarity observed in both cases suggests that either model can roughly match the spectral properties of Δc observed in the postwar U.S. data. Unfortunately, the tools developed in this paper cannot, by themselves, help determine which is the correct model. That task will have to involve looking at alternative models' predictions, other than those summarized in the spectrum of consumption changes.

III. Budget Constraints, Excess Smoothness, and Predictability of Consumption Changes

In this section, I use the framework developed above to show the relationship between (i) excess smoothness in consumption, defined here as a value of ψ less than 1, and (ii) predictability of consumption changes (i.e., a significant correlation between Δc_t and variables observed as of time $t-1$). Both phenomena, which seem to be present in postwar U.S. data,[16] are in con-

[15] The 0.44 value for ρ corresponds to the estimate of the first-order autoregressive coefficient in an AR(1) model for labor-income changes, as reported in Deaton (1987).

[16] Predictability of consumption changes was originally established by Flavin (1981). For the excess-

flict with the PIH model. Does one necessarily imply the other?[17] To the extent that aggregate consumption can be characterized by the general model given by (5), (6), and (7), two basic results arise, which are stated in the two following propositions.

PROPOSITION 1: *Predictability of consumption changes has no bearing on the value of the variance ratio ψ (i.e., on the variability of consumption relative to the variability implied by the PIH model.*

Proposition 1 can be proved "by illustration." Consider the following models of consumption, all of them consistent with (5), (6), and (7):

(i) $\Delta c_t = [(1+r)/(r+\delta)][1-(1-\delta)L]\xi_t$

(durable goods)

(ii) $\Delta c_t = (1+r)\xi_{t-1}$

(lagged response)

(iii) $(1-\tau L)\Delta c_t = [1-\tau/(1+r)]\xi_t$

(habit formation).

The three previous models imply that consumption changes should be predictable by any variable correlated with lagged ξ's (e.g., lagged labor-income changes). Yet, as r approaches zero, the models above are easily shown to be associated with values of ψ greater, equal, and smaller than 1, respectively. In other words, (i) generates "excess variability" of consumption, in (ii) consumption is as variable as predicted by the PIH model, and (iii) leads to "excess smoothness."

smoothness finding, see all the references cited in previous sections.

[17]A related question was addressed by Campbell (1987) and Campbell and Deaton (1989). Those papers examine the relationship between predictability of consumption changes and violations of the PIH condition $\Delta c_t = \xi_t$. The approach here allows one to examine the link between predictability of consumption changes and the value of the variance ratio ψ, the measure of relative consumption variability used in the literature.

PROPOSITION 2: *Excess smoothness in consumption (i.e., $\psi < 1$) implies that consumption changes should be predictable by any lagged variable correlated with lagged innovations in permanent income.*

Proposition 2 can be proved by showing that unpredictability of consumption changes by variables correlated with lagged permanent-income innovations cannot coexist with excess smoothness. Suppose Δc is not predictable by variables correlated with lagged ξ's. In that case, the "fundamental" component of consumption changes cannot depend on lagged ξ's, so it must be that $\Delta c_t^* = \sum_{j=0}^{\infty}\beta_j\xi_{t-j} = \beta_0\xi_t$. From (6), it is known that, as $r \to 0$, the intertemporal budget constraint requires that $\beta_0 \to 1$. Thus, for low values of r, $\mathrm{Var}(\Delta c^*) = \mathrm{Var}(\xi)$ must (approximately) hold. By construction, $\mathrm{Var}(\Delta c) \geq \mathrm{Var}(\Delta c^*)$, so it must be the case that $\psi \geq 1$, that is, there cannot be excess smoothness. By contraposition, excess smoothness ($\psi < 1$) necessarily implies $\Delta c_t^* = \sum_{j=0}^{\infty}\beta_j\xi_{t-j}$, with $\beta_j \neq 0$ for some $j > 0$, which guarantees that consumption changes are predictable by some lagged ξ's and/or variables correlated with some lagged ξ's.

Summarizing, for any model of consumption satisfying the general framework given by (5), (6), and (7), two basic results obtain: (i) excess smoothness in consumption ($\psi < 1$) implies that consumption changes are predictable; but, (ii) the finding of predictability of consumption changes has no bearing on the variability of consumption.

IV. Summary and Conclusions

The present paper has developed a simple framework for the time-series analysis of consumption, based on the restrictions implied by the intertemporal budget constraint on the low-frequency properties of consumption. An estimator of the variance ratio measuring the variability of consumption relative to permanent income is derived. That estimator shows two advantages over estimators found in the literature: (i) it imposes no restrictions on the order of integration, predictability, and other stochastic properties of the income process or on the information set used by consumers to form their expectations; and (ii) it is robust to a wide range of departures from the standard PIH model.

Empirical results using postwar U.S. data suggest that the variability of consumption is less than 80 percent of the variability predicted by the standard PIH model. That result supports the finding of "excess smoothness" obtained by researchers working under the assumption of a unit root in the income process (e.g., Deaton, 1987; West, 1988). In addition, the paper shows that such "excess smoothness" seems to hold at *all frequencies*—a stronger result. The results are shown to be consistent with simple models of habit-formation and liquidity constraints, among others.

Finally, the paper has shown how the intertemporal budget constraint implies a certain relationship between measures of relative variability of consumption and predictability of consumption changes. In particular, I have shown that excess smoothness implies predictability of consumption changes, though the converse is not true.

APPENDIX

PROOF OF LEMMA:

From the properties of spectral densities and the specification of the Δc, Δc^*, and n processes above, it is known that

$$h_{\Delta c^*}(\omega) = (2\pi)^{-1}|\beta(e^{-i\omega})|^2 \operatorname{Var}(\xi)$$

$$h_n(\omega) = (2\pi)^{-1}|\varphi(e^{-i\omega})|^2 \operatorname{Var}(\eta)$$

where $|\cdot|^2$ is the usual "product by complex conjugate" operator, $\beta(z) \equiv \sum_{j=0}^{\infty}\beta_j z^j$, and $\varphi(z) \equiv \sum_{j=0}^{\infty}\varphi_j z^j$.

The spectral density of Δc^* and n at frequency zero is thus given by $h_{\Delta c^*}(0) = (2\pi)^{-1}(\sum_{j=0}^{\infty}\beta_j)^2 \operatorname{Var}(\xi)$. Clearly,

$$\lim_{r\to 0}\left[2\pi h_{\Delta c^*}(0) - \left(\sum_{j=0}^{\infty}(1+r)^{-j}\beta_j\right)^2 \operatorname{Var}(\xi)\right] = 0$$

and (i) in the Lemma follows from applying (6) to the previous expression. Similarly,

TABLE A1—VALUES OF $\sum_{j=0}^{\infty}(1+r)^{-j}\beta_j$ AND $\sum_{j=0}^{\infty}\beta_j$ FOR DIFFERENT VALUES OF r AND τ

τ	$\sum_{j=0}^{\infty}\beta_j$	r	$\sum_{j=0}^{\infty}(1+r)^{-j}\beta_j$
0.3	1.428	0.0025	1.427
0.3	1.428	0.01	1.422
0.3	1.428	0.02	1.416
0.7	3.333	0.0025	3.314
0.7	3.333	0.01	3.258
0.7	3.333	0.02	3.187
0.9	10	0.0025	9.780
0.9	10	0.01	9.181
0.9	10	0.02	8.500

$h_n(0) = (2\pi)^{-1}(\sum_{j=0}^{\infty}\varphi_j)^2 \operatorname{Var}(\eta)$, and

$$\lim_{r\to 0}\left[2\pi h_n(0) - \left(\sum_{j=0}^{\infty}(1+r)^{-j}\varphi_j\right)^2 \operatorname{Var}(\eta)\right] = 0.$$

Using (7), result (ii) in the Lemma follows.

Note that $r = 0$ is not assumed, since y^P is not well-defined in that case. Instead, r is assumed to be small enough, and the β's and φ's are assumed to converge to zero fast enough, to make $\sum_{j=0}^{\infty}(1+r)^{-j}\beta_j \cong \sum_{j=0}^{\infty}\beta_j$ and $\sum_{j=0}^{\infty}(1+r)^{-j}\varphi_j \cong \sum_{j=0}^{\infty}\varphi_j$ good approximations. As an illustration of the "goodness" of such an approximation, consider a model with $\beta_j = \tau^j$. Table A1 compares the values of $\sum_{j=0}^{\infty}(1+r)^{-j}\beta_j$ and $\sum_{j=0}^{\infty}\beta_j$ for different (plausible) values for r and τ (keep in mind that the empirical analysis in the paper uses quarterly data).

Even in the worst possible case considered, with slowly decaying β's ($\tau = 0.9$) and the highest interest rate (2 percent quarterly), the approximation error is relatively small (15 percent). More plausible settings yield an approximation error in the neighborhood of 1 percent.

Choice of Window Parameters[18]

For any type of window, and given a finite number of available observations, the choice

[18]This section draws heavily on Priestley (1981 section 7.3).

of the window-parameter M involves a basic trade-off between the bias and the variance of the spectrum estimator $\hat{h}_{\Delta c}(\omega)$. A useful criterion consists of choosing the value of M that minimizes the relative mean-square error of the estimator, defined by

$$\mathrm{RMS}^2(\omega) \equiv \left[E\left(\hat{h}_{\Delta c}(\omega) - h_{\Delta c}(\omega)\right)^2\right] \Big/ h^2_{\Delta c}(\omega)$$

$$= \left\{\mathrm{Var}\left[\hat{h}_{\Delta c}(\omega)\right] + b^2\left[\hat{h}_{\Delta c}(\omega)\right]\right\} \Big/ h^2_{\Delta c}(\omega)$$

where $b(\hat{h}_{\Delta c}(\omega)) \equiv E\hat{h}_{\Delta c}(\omega) - h_{\Delta c}(\omega)$ denotes the estimator bias.

Consider first the Bartlett estimator introduced in the text. It is possible to show that, at frequency zero, the following expression for the relative mean-square error holds:

$$\mathrm{RMS}^2_{\mathrm{B}}(0) = (4/3)(M/N) + (1/M)^2 A(\Delta c)$$

where

$$A(\Delta c) \equiv \left\{\left[\sum_{s=-\infty}^{\infty} |s| \rho_{\Delta c}(s)\right] \Big/ \left[\sum_{s=-\infty}^{\infty} \rho_{\Delta c}(s)\right]\right\}^2.$$

The value of M that minimizes $\mathrm{RMS}^2_{\mathrm{B}}(0)$ is given by $M^* = [1.5NA(\Delta c)]^{1/3}$. An approximation for $A(\Delta c)$ can be obtained from its (truncated) sample counterpart

$$\hat{A}(\Delta c) = \left\{\left[\sum_{s=-k}^{k} |s| \rho_{\Delta c}(s)\right] \Big/ \left[\sum_{s=-k}^{k} \rho_{\Delta c}(s)\right]\right\}^2.$$

Table A2 reports the values of $\hat{A}(\Delta c)$ and M^* for a range of k values (different truncation points), given $N = 165$. Thus, for a variety of approximations of $A(\Delta c)$, the corresponding values of M that minimize the relative mean-square error are, for the cases considered, less than 10 (i.e., rather small).

A similar analysis can be carried out for the Bartlett-Priestley estimator. The corresponding expression for the relative mean-square error at $\omega = 0$ is given by

TABLE A2—VALUES OF $\hat{A}(\Delta c)$ AND M^* FOR A RANGE OF k VALUES

Truncation lag	$\hat{A}(\Delta c)$	M^*
1	0.10	2.97
2	0.35	4.42
3	1.20	6.67
5	1.38	7.00
10	3.24	9.29
15	2.34	8.34

$$\mathrm{RMS}^2_{\mathrm{P}}(0) = (12/5)(M/N) + 0.16(\pi/MB_{\mathrm{h}})^4$$

where B_{h} is the "spectral bandwidth," as defined in Priestley (1981). The value of M that minimizes $\mathrm{RMS}^2_{\mathrm{P}}(0)$ is approximately given by $M^* \cong [0.26N(\pi/B_h)^4]^{1/5}$. If one takes $\pi/6$ to be a "safe" lower bound for B_{h}—this would accommodate, for example, any AR(1) process with first-order autocorrelation between 0 and 0.77—an upper bound for M^* is given by 9, again a relatively small value.

Intuitively, the desirability of small values for the window-parameter M is related to a key feature of the Δc time-series: the quick rate of decay in its autocovariogram. That feature is typically associated with a slowly decaying spectrum (i.e., a spectrum with a large bandwidth). Under those circumstances, the bias tends to be small and relatively unaffected by changes in the window width. Loosely speaking, in order to minimize the mean-square error one gives a relatively large "weight" to the variance, which results in a small value of M.

REFERENCES

Caballero, Ricardo J., "Consumption Puzzles and Precautionary Savings," *Journal of Monetary Economics*, January 1990, *25*, 113–36.

Campbell, John Y., "Does Saving Anticipate Declining Labor Income? An Alternative Test of the Permanent Income Hypothesis," *Econometrica*, November 1987, *55*, 1249–73.

______ and **Deaton, Angus**, "Why Is Consumption So Smooth?" *Review of Economic Studies*, July 1989, *56*, 357–74.

______ and **Mankiw, N. Gregory**, "Consumption, Income and Interest Rates: Reinterpreting the Time Series Evidence," in Olivier J. Blanchard and Stanley Fischer, eds., *NBER Macroeconomics Annual 1989*, Cambridge, MA: MIT Press, 1989, pp. 185–215.

Clarida, Richard H., "Smooth Consumption and Noisy Permanent Income: Is There A Puzzle?" mimeo, Columbia University, June 1988.

Cochrane, John H., "How Big Is the Random Walk in GNP?" *Journal of Political Economy*, October 1988, *96*, 893–920.

Deaton, Angus, "Life-Cycle Models of Consumption: Is the Evidence Consistent with the Theory?" in T. Bewley, ed., *Advances in Econometrics, Fifth World Congress*, Vol. 2, Cambridge: Cambridge University Press, 1987.

Diebold, Francis X. and Rudebusch, Glenn D., "Is Consumption Too Smooth? Long Memory and the Deaton Paradox," *Review of Economics and Statistics*, February 1991, *73*, 1–17.

Flavin, Marjorie, "The Adjustment of Consumption to Changing Expectations About Future Income," *Journal of Political Economy*, October 1981, *89*, 974–1009.

______, "The Excess Smoothness of Consumption: Identification and Interpretation," NBER (Cambridge, MA) Working Paper No. 2807, December 1988.

Fuller, Wayne A., *Introduction to Statistical Time Series*, New York: Wiley, 1976.

Galí, Jordi, "Finite Horizons, Life Cycle Savings and Time Series Evidence on Consumption," *Journal of Monetary Economics*, November 1990, *26*, 433–52.

Hall, Robert E., "Stochastic Implications of the Life-Cycle–Permanent-Income Hypothesis: Theory and Evidence," *Journal of Political Economy*, December 1978, *86*, 971–87.

______ and **Mishkin, Frederic S.**, "The Sensitivity of Consumption to Transitory Income: Estimates from Panel Data on Households," *Econometrica*, March 1982, *50*, 461–81.

Hansen, Lars Peter, Roberds, William and Sargent, Thomas J., "Time Series Implications of Present Value Budget Balance and of Martingale Models of Consumption or Taxes," mimeo, University of Chicago, 1990.

Mankiw, N. Gregory, "Hall's Consumption Hypothesis and Durable Goods," *Journal of Monetary Economics*, November 1982, *10*, 417–25.

Newey, Whitney K. and West, Kenneth D., "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, May 1987, *55*, 703–8.

Priestley, M. B., *Spectral Analysis and Time Series*, San Diego: Academic Press, 1981.

Quah, Danny, "Permanent and Transitory Movements in Labor Income: An Explanation for Excess Smoothness in Consumption," *Journal of Political Economy*, June 1990, *98*, 449–75.

West, Kenneth D., "The Insensitivity of Consumption to News about Income," *Journal of Monetary Economics*, January 1988, *21*, 17–34.