Federal Reserve Bank of Minneapolis
Research Department

historical

THE PERMANENT INCOME HYPOTHESIS REVISITED

Lawrence J. Christiano, Martin Eichenbaum,
and David Marshall*

Working Paper 335

Revised November 1987

ABSTRACT

This paper investigates whether there are simple versions of the permanent income hypothesis which are consistent with the aggregate U.S. consumption and output data. Our analysis is conducted within the confines of a simple dynamic general equilibrium model of aggregate real output, investment, hours of work and consumption. We study the quantitative importance of two perturbations to the version of our model which predicts that observed consumption follows a random walk: (i) changing the production technology specification which rationalizes the random walk result, and (ii) replacing the assumption that agents' decision intervals coincide with the data sampling interval with the assumption that agents make decisions on a continuous time basis. We find substantially less evidence against the continuous time models than against their discrete time counterparts. In fact neither of the two continuous time models can be rejected at conventional significance levels. The continuous time models outperform their discrete time counterparts primarily because they explicitly account for the fact that the data used to test the models are time averaged measures of the underlying unobserved point-in-time variables. The net result is that they are better able to accommodate the degree of serial correlation present in the first difference of observed per capita U.S. consumption.

*Christiano, Federal Reserve Bank of Minneapolis; Eichenbaum, Carnegie-Mellon University and the National Bureau of Economic Research; Marshall, Carnegie-Mellon University. We have benefited from helpful discussions with S. Rao Aiyagari, Lars Hansen, and Thomas Sargent. The first two authors acknowledge the financial support of the National Science Foundation. The third author was an Alfred P. Sloan doctoral dissertation Fellow while this paper was being written.

1. Introduction

Few subjects in macroeconomics have received as much attention as the relationship between aggregate consumption and output. This attention reflects, at least in part, the belief that an understanding of the structural determinants of aggregate consumption is central to resolving many of the outstanding issues in business cycle theory. During the past decade much of the empirical literature on aggregate consumption has centered on Hall's [1978] demonstration that, under certain conditions, the permanent income hypothesis (PIH) implies that consumption is a random walk. Under this random walk hypothesis (RWH) no variable apart from current consumption should be of value in predicting future consumption.

In fact, a number of authors, including Flavin [1981] and Hayashi [1982], report statistically significant correlations between the change in consumption and lagged consumption and income. The response to these findings has generally fallen into one of two categories. First, some researchers have attributed the "excess sensitivity" of consumption to current and lagged income to the presence of a substantial number of consumers who are liquidity constrained. Under this interpretation, the PIH is fundamentally flawed as a principle for organizing the aggregate time series data (see for example Hall and Mishkin [1982] and Zeldes [1985]).

A second view of the empirical shortcomings of the RWH is that they do not reflect the failure of the PIH per se. Instead they reflect the failure of the auxiliary assumptions required to derive the RWH from the PIH. This view underlies both intertemporal capital asset pricing models (eg., Hansen and Singleton [1982, 1983], Dunn and Singleton [1986] and Eichenbaum and Hansen [1986]) and real business cycle theories (eg., Kydland and Prescott [1982], Long and Plosser [1983], and Michener [1984]) which abstract from liquidity constraints and other market imperfections which would prevent consumers from optimally adjusting consumption to permanent income. This view also underlies Lucas' [1985] argument that the welfare gains associated with countercyclical government policies would, at the very best, be small. Given the radically different policy implications of the two types of responses it is not surprising that the relationship between aggregate consumption and income continues to command widespread interest.

This paper pursues the second of the two responses discussed above. We investigate whether there are perturbations of the random walk version of the PIH, as implemented by Hall [1978] and Flavin [1981], which are consistent with the aggregate consumption and output data. The two perturbations we consider are: (i) changing the production technology to a specification which no longer implies the RWH, and (ii) replacing the assumption that agents' decision intervals coincide with the data sampling interval with the assumption that agents make decisions on a continuous time basis.

Our analysis follows Hansen [1986] and Sargent [1986] in interpreting the PIH as a simple version of the Brock-Mirman growth model in which the equilibrium law of motion for consumption and output takes the form of a constrained vector ARMA. Consumers' preferences are defined over consumption and leisure in a way that nests the specification considered by Hall [1978] and Flavin [1981]. Output is produced using both labor and capital according to a Leontieff type production function in which the labor requirement per unit of capital is allowed to be stochastic. When this labor requirement is nonstochastic our model satisfies the RWH. When agents derive disutility from working and the labor requirement per unit of capital is a nontrivial stochastic process, consumption does not follow a random walk. Aggregate income will Granger cause the first difference of consumption and current and lagged changes in consumption will be of value for predicting future changes in consumption. Consequently, this version of our model can, in principle, explain Flavin's rejection of the RWH.

A second possible explanation of these rejections is the impact of temporal aggregation bias. Sims [1971], Geweke [1978], Marcet [1986] and Christiano [1982, 1985] have shown that temporal aggregation bias can induce spurious serial correlation and Granger causality. In fact, much of the empirical evidence against different versions of the PIH consists of findings that the first difference of aggregate consumption is serially correlated and is Granger caused by a variety of other variables. If agents make economic decisions at intervals of time that are finer than the data sampling interval these serial correlation and Granger causality findings could be spurious in the sense that they reflect only the effects of temporal aggregation bias.

In order to investigate this possibility we analyze continuous time versions of our discrete time model. These models are estimated using techniques developed by Hansen and Sargent [1980, 1981] for estimating continuous time models from discrete time data. This strategy allows us to directly address the possibility of temporal aggregation bias and to explicitly account for the fact that consumption and income data are not point-in-time sampled.

The remainder of this paper is organized as follows. In section 2 we present the discrete time versions of our model. Empirical results for the discrete time models are presented in section 3. In section 4 we present the continuous time analogue to the models discussed in section 2. Empirical results for the continuous time models are discussed in section 5. Section 6 concludes the paper.

2. The Discrete Time Permanent Income Hypothesis

2.A The Model

We suppose that the time series on economy-wide consumption, the stock of capital, and output correspond to the solution of an optimal resource allocation problem which can be decentralized as a competitive equilibrium. The social planning problem that we consider has more than one interpretation in terms of consumers' preferences and the technology for producing new consumption goods. For pedagogical reasons we find it convenient to proceed in terms of one of these interpretations. Other interpretations are discussed after the optimal resource allocation problem has been stated. There we indicate that, given observations only on aggregate consumption and output and abstracting from growth considerations, our model is observationally equivalent to a version of the model considered by Sargent [1986].

A representative consumer ranks alternative streams of consumption and leisure according to the preference specification,

$$(2.1) \qquad E_0 \sum_{t=0}^{\infty} \beta^t \{-\tfrac{1}{2}(c_t-b_t)^2-\alpha_t h_t\}$$

where $0 < \beta < 1$ is the subjective discount rate, b_t denotes the consumer's bliss point for consumption at time t, c_t denotes consumption at time t, h_t denotes work effort at time t, α_t is the marginal disutility of work in period t and E_t is the expectations operator conditioned on the information set I_t, $t \geq 0$. The set I_t contains observations on all model variables dated t and earlier. Throughout this paper we assume that b_t and α_t are deterministic functions of time.

There is a technology that converts time t consumption goods and labor effort into time t + 1 consumption goods. This technology is given by

$$(2.2) \qquad \tilde{y}_t = \min\{\tilde{\delta}k_{t-1}, \tau_{t-1}h_{t-1}\} + e_t.$$

Here, $\tilde{y}_t$ denotes output, k_{t-1} is the capital stock at the <u>end</u> of time t - 1. We think of the variable e_t either as the endowment of consumption at time t or as an aggregate shock to the production function at time t which affects only the average productivity of labor and capital. The variable $\tilde{\delta}/\tau_{t-1}$ represents the labor requirement per unit of capital.

The economy-wide resource constraint is given by:[2.1]

$$(2.3) \qquad c_t + k_t - (1-d)k_{t-1} = \tilde{y}_t$$

where d is the rate at which a unit of capital depreciates, $0 \leq d < 1$ and $\tilde{\delta} \geq d$. We impose the condition

$$(2.4) \qquad \beta[\tilde{\delta}+(1-d)] = \beta\delta = 1, \text{ where } \delta \equiv \tilde{\delta} + 1 - d.$$

Condition (2.4) results in a unit autoregressive root in the consumption process which is a necessary (but not sufficient) condition for the RWH.

As in Hansen [1986] and Sargent [1986], we do not impose a nonnegativity constraint on the choice variables of the model. Imposition of this constraint makes it difficult if not impossible to solve the model analytically. Instead we follow Hansen [1986] in imposing the requirement that[2.2]

$$(2.5)\qquad E_0 \sum_{t=0}^{\infty} \beta^t k_t^2 < \infty.$$

This condition emerges from viewing our infinite horizon economy as the limit of a sequence of finite horizon economies in which we impose the constraint that the terminal capital stock is zero (for a discussion, see Hansen, Roberds, and Sargent [1987]). In deriving the solution to the optimal resource allocation problem, it is convenient to impose the restriction that capital and labor are always fully utilized:

$$(2.6)\qquad \tilde{\delta} k_t = \tau_t h_t \quad \text{for all } t.$$

Christiano, Eichenbaum and Marshall [1987] discuss conditions under which this restriction is nonbinding.

Relations (2.1)-(2.3), (2.6) and the definition $H_t \equiv \tilde{\delta}\alpha_t/\tau_t$, imply that the social planner's problem is to maximize

$$(2.7)\qquad E_0 \sum_{t=0}^{\infty} \beta^t \{-\tfrac{1}{2}\,[\delta k_{t-1} - k_t + e_t - b_t]^2 - H_t k_t\}$$

by choice of a contingency plan for k_t subject to (2.5).

Planning problems observationally equivalent to (2.7) can arise from a variety of other models. For example, we could replace (2.2) with the assumption that no labor is required to produce new consumption goods,

$$(2.2)'\qquad \tilde{y}_t = \tilde{\delta} k_{t-1} + \varepsilon_t,$$

and assume that there is a stochastic labor requirement for maintaining the capital stock given by (2.6). In addition, Christiano, Eichenbaum and Marshall [1987] show that, given only observations on consumption and output, and assuming τ_t is deterministic, another observationally equivalent model can be obtained by allowing b_t to be stochastic. An example of such a model is contained in Sargent [1986], where no labor is needed to produce output (ie., $\tau_t \equiv \infty$) and b_t is a nontrivial random variable.

We use the following notation:

$$(2.8) \qquad x_{pt} \equiv (1-\beta)E_t \sum_{j=0}^{\infty} \beta^j x_{t+j}.$$

$$(2.9) \qquad \mu_{x_p,t} \equiv x_{pt} - E_{t-1}x_{pt}.$$

Definitions (2.8) and (2.9) apply to any random variable x_t for which the indicated conditional expectation exists. Below, we refer to objects like x_{pt} as the "permanent" value of x_t and to $\mu_{x_p,t}$ as the innovation to the permanent value of x_t. 2.3/

In Appendix A, we show that the equilibrium laws of motion for k_t and c_t are:

$$(2.10) \qquad k_t - k_{t-1} = (e_t - e_{pt}) - (b_t - b_{pt}) - \beta H_{pt}/(1-\beta).$$

$$(2.11) \qquad c_t = e_{pt} + (b_t - b_{pt}) + [\beta/(1-\beta)]H_{pt} + (\delta-1)k_{t-1}.$$

Since net output, y_t, is equal to consumption plus net investment, we see that

$$(2.12) \qquad c_t - y_t = (e_{pt} - e_t) + (b_t - b_{pt}) + \beta H_{pt}/(1-\beta).$$

Relation (2.10) implies that investment increases when the current value of the productivity shock exceeds its permanent value and decreases when the utility associated with a given amount of consumption is unusually high ($b_t > b_{pt}$). In addition, net investment depends negatively on H_{pt}, reflecting the utility cost of the labor input needed to make additions to the capital stock productive in the future.

Relation (2.11) implies that consumption increases when the utility associated with consumption is unusually high and depends positively on permanent endowment income and the capital stock. In addition consumption depends positively on H_{pt}. This is because high values of H_{pt} signify a low opportunity cost of consuming goods at time t as opposed to combining them with labor in order to produce future consumption goods. According to (2.12) unusually high levels of utility associated with consumption, high levels of H_{pt} or unusually low levels of endowment income ($e_t < e_{pt}$) cause consumption to exceed current period income.

Definition: We say that consumption satisfies the **random walk hypothesis** if and only if consumption is a martingale possibly with deterministic, time varying drift, i.e.,

$$(2.13) \qquad E_{t-1}c_t = c_{t-1} + f_t$$

where f_t is a deterministic, but possibly trivial, function of time.2.4/

From relations (2.10) and (2.11) we see that

$$(2.14) \qquad \Delta c_t = \mu_{e_p,t} - \mu_{b_p,t} + \Delta b_t + [\beta/(1-\beta)]\mu_{H_p,t} - H_{t-1},$$

where $\Delta c_t \equiv c_t - c_{t-1}$. Since b_t is by assumption deterministic, the RWH will be satisfied if and only if H_t is deterministic. Since α_t is by assumption deterministic, we conclude that the RWH will hold if and only if the time t labor requirement per unit of capital, τ_t, is deterministic.2.5/

2.B Some Evidence on the Discrete Time RWH

The RWH has testable implications independent of the assumed probability structure of e_t. This section reports empirical evidence against these implications, confirming results reported in Flavin [1981] and Hayashi [1982]. We begin with a description of the data.

Our model divides total output into only two categories: consumption and investment (see (2.3)) In view of this we measure total consumption, c, as the sum of total government consumption, c_g, and private consumption. The latter was measured as real quarterly expenditures on nondurable consumption goods (c_{nd}) and services (c_s), plus an estimate of the service flow from the stock of consumer durables (c_{sd}). All of these measures except c_{sd} and c_g were taken from the National Income and Product Accounts (NIPA). The service flow from the stock of consumer durables was obtained from the data base documented in Brayton and Mauskopf [1985]. Government consumption was measured by NIPA real government purchases of goods and services (g) minus real government investment (i_g). A measure of i_g was provided to us by John Musgrave of the U.S. Department of Commerce's Bureau of Economic Analysis.[2.6] Gross output, $\tilde{y}$, was measured by per capita real quarterly GNP plus c_{sd}. All series cover the period 1950,2-1985,3 and are converted to per capita terms using a measure of the total population that includes armed forces overseas (data mnemonic NPT, obtained from the Wharton Econometrics data base.)

Table 2.1 provides an idea of the order of magnitude of the components of our measure of consumption. First, note that c_{sd} is only a small part of c, going from about two percent in the 1950's to about three percent in the 1980's. Second, c_g and c_{nd} are both gradually declining fractions of c whereas c_s and c_{sd} have an upward trend.[2.7] Finally, note that c_g is significantly smaller than the standard measure of government consumption, g, due to the fact that roughly 20 percent of government purchases represent investment activity.

The maintained hypothesis throughout this paper is that c and $\tilde{y}$ are made covariance stationary by detrending using a common geometric growth rate, ϕ.[2.8] This is an implication of all the fully parameterized versions of the model (see sections 2.C, 4.B, and 4.C). The value of ϕ used to construct the detrended data is $\exp(\theta)$, with $\theta = .004568$. This value of θ is the coefficient on time of the regression of $\log c_t$ and $\log \tilde{y}_t$ on a linear time trend, computed subject to the restriction that the growth rates in consumption and output are equal. When this restriction is not imposed the measured growth rates in consumption and output are .004582 and .004606.

According to the discrete time RWH, Δc_t is uncorrelated with information dated t - 1 and earlier. It is straightforward to test this implication using results in Hansen [1982]. Define the function:

$$(2.15) \quad \psi(C,\phi^{-t}\Delta c_t) = \phi^{-t}\Delta c_t - C.$$

To simplify notation, we write

$$(2.15)' \quad \psi_t \equiv \psi(C,\phi^{-t}\Delta c_t),\ \psi_t^0 \equiv \psi(C^0,\phi^{-t}\Delta c_t),$$

where $C^0 = E\phi^{-t}\Delta c_t$ is an unknown parameter. According to the RWH, $E_{t-1}\psi_t^0 = 0$, so that $E\psi_t^0 z_{it} = 0$ for all z_{it} contained in I_{t-1}. In practice we considered the vectors:

$$(2.16) \quad z_{1t} = [1,\phi^{-(t-1)}\Delta c_{t-1},\ldots,\phi^{-(t-M)}\Delta c_{t-M},\phi^{-(t-1)}\tilde{y}_{t-1},\ldots,\phi^{-(t-M)}\tilde{y}_{t-M}],$$

and,

$$(2.16)' \quad z_{2t} = [1,\phi^{-(t-1)}\Delta c_{t-1},\ldots,\phi^{-(t-M)}\Delta c_{t-M},\phi^{-(t-1)}(c_{t-1}-\tilde{y}_{t-1}),$$
$$\ldots,\phi^{-(t-M)}(c_{t-M}-\tilde{y}_{t-M})],$$

with $M = 4$. Under the assumption that $\phi^{-t}\Delta c_t$ and z_{it} are jointly stationary and ergodic the generalized method of moments (GMM) procedure described in Hansen [1982] can be used to estimate the parameter C and test the null hypotheses, $E\psi_t^0 z_{it} = 0$, $i = 1, 2$. Define the function $g_{iT} = (1/T)\sum_{t=1}^{T}\psi_t z_{it}$, $i = 1, 2$. Our estimator of C^0 is the argmax of $J_{iT} = g_{iT}'W_{iT}^{-1}g_{iT}$, where W_{iT} is the sample covariance matrix of $\psi_t z_{it}$, $i = 1, 2$. The minimized value of J_{iT}, $i = 1, 2$ is asymptotically distributed as a chi-square random variable with eight degrees of freedom.

The significance levels of our test statistics are reported in the "Lags 1 - 4" portion of Table 5.2. Tables 5.2a and 5.2b report results based on the instrument vectors z_{1t} and z_{2t}, respectively. As a check on robustness we report results for three sample periods. Moreover, we also present results for measures of consumption and income other than ours. In addition to providing further evidence on the robustness of our results, this also facilitates comparison with results in Hall [1978] and Flavin [1981]. The columns marked $(c,\tilde{y})$ report results for the concepts of consumption and income in our model. The columns marked (c_{nd},y_d) report results for the consumption and income concepts used in Flavin [1981] ($y_d \equiv$ per capita disposable income), while columns marked $(c_{nd}+c_{sd},y_d)$ report results for the consumption and income concepts used in Hall [1978]. Following Flavin, these measures of consumption and income were detrended using different, though constant, geometric growth rates.[2.9/] The results of Table 5.2 provide very strong evidence against the RWH. In particular we can, with only two exceptions, reject the discrete time RWH at the five percent significance level or higher. In light of these results, we now turn to a parameterization of the model which does not satisfy the RWH.

2.C The Discrete Time Stochastic Labor Requirement Model (DSLR)

In this subsection we describe a parameterization of our discrete time model in which the labor requirement per unit of capital, τ_t, is stochastic so that the RWH does not hold. We refer to this version of the model as the discrete time stochastic labor requirement (DSLR) model. The testable implications of the DSLR model are determined by our restrictions on $\{e_t,H_t\}$ in the sense that, absent any such restrictions, that model is not refutable. This follows from results in Christiano, Eichenbaum, and Marshall [1987] who show that given any Wold representation for $\{\phi^{-t}\Delta c_t,\phi^{-t}(c_t-\tilde{y}_t)\}$, one can always find an $\{e_t,H_t\}$ representation such that the corresponding reduced form of our model coincides with the given Wold representation. Of course, it is also true that the testable restrictions of the RWH ultimately stem from restrictions on unobservables (eg., α_t, b_t and H_t.) Thus, viewed from a broad perspective, all of the testable restrictions of our model derive from assumptions regarding the probability law of the exogenous, unobserved forcing variables. This type of result is by no means unique to our model. We conjecture that such a result holds in any dynamic stochastic general equilibrium model when the econometrician does not observe any exogenous variables.

As noted above, our parameterization of the technology shocks implies that consumption and output grow at the same geometric rate over time. While this parameterization is very restrictive, it does have an important compensating advantage: it implies that the model applies to consumption and output data which have been detrended assuming a common geometric trend. This allows us to accommodate growth in an internally consistent way while preserving the applicability of a set of econometric tools developed for nongrowing time series.

We suppose that $b_t = b\phi^t$ and $\alpha_t = \alpha\phi^{2t}$ where $\phi > 1$ and $\alpha > 0$. In order for the representative consumer's problem to be well-defined we require that $\beta\phi^2 < 1$. By allowing b_t to grow over time we are able to avoid the implication that the consumer becomes satiated. The fact that α_t grows at the geometric rate ϕ^2 implies, in conjunction with the other assumptions in our model, that neither leisure nor labor's share of net output exhibits a trend.

We model H_t as an AR(1) random variable with time varying drift:

$$(2.17) \qquad H_t = H\phi^t + \epsilon_t/(1-fL), \ |f| < 1, \ 0 < H < b(\phi-1).$$

The lower bound on H guarantees that the drift in the marginal utility of consumption, H_t, is positive. The upper bound on H ensures that the drift in consumption, $\Delta b_t - H_{t-1}$, is positive. Since $H_t = \tilde{\delta}\alpha_t/\tau_t$, condition (2.17) implies that the labor requirement per unit of capital is stochastic. In addition we assume that

$$(2.18) \qquad (1-L)e_t = e\phi^t + \eta_t/(1-aL), \ e\phi \le b(\phi-1) - H, \ |a| < 1.$$

The condition on e, together with (2.10), guarantees that the deterministic component of savings, $k_t - k_{t-1}$, is positive.

Let $\mathbf{x}_t = [\epsilon_t \eta_t]'$. The vector $\mathbf{x}_t$ is white noise, orthogonal to $I_{t-1} = \{k_{t-1-s}, H_{t-1-s}, e_{t-s}, s \ge 0\}$, and satisfies

$$(2.19) \qquad E\mathbf{x}_t\mathbf{x}_t' = \mathbf{V}\phi^{2t},$$

where $\mathbf{V}$ is a two by two positive definite symmetric matrix of constants.

According to our specification, all deterministic terms and innovation standard deviations grow at the rate ϕ.[2.10] Thus it is not surprising that we can "detrend" Δc_t and $c_t - y_t$ by ϕ^t to obtain a stationary stochastic process. Define,

$$(2.20) \qquad c_t^* = \phi^{-t}c_t,\ y_t^* \equiv y_t,\ \epsilon_t^* \equiv \phi^{-t}\epsilon_t, \text{ and } \eta_t^* \equiv \phi^{-t}\eta_t,$$

and

$$(2.21) \qquad q_t^{*\prime} = \left[c_t^*-y_t^*, c_t^*-\phi^{-1}c_{t-1}^*\right].$$

Relations (2.12), (2.14), and (2.17)-(2.19) imply that q_t^* has the VAR(2) representation:

$$(2.22) \qquad \mathbf{A}(\phi^{-1}L)q_t^* = \mathbf{T} + \mathbf{x}_t$$

where,

$$(2.23) \qquad \mathbf{A}(L) = \mathbf{I} + \mathbf{A}_1 L + \mathbf{A}_2 L^2,$$

$$\mathbf{A}_1 = \begin{bmatrix} \beta a^2 - f & -\beta a(a-f) \\ \beta^{-1} - f & -a(1-f\beta) \end{bmatrix} / (1-\beta a) \quad \mathbf{A}_2 = \begin{bmatrix} -a(a-f) & 0 \\ -a(\beta^{-1}-f) & 0 \end{bmatrix} / (1-\beta a)$$

$$\mathbf{X}_t = \begin{bmatrix} \beta a & 1 \\ 1 & 1 \end{bmatrix} \begin{pmatrix} \eta_t^*/(1-\beta a) \\ (1-\beta)\varepsilon_t^*/(1-f\beta) \end{pmatrix}$$

$$\eta_t^* = \phi^{-t}\eta_t \text{ and } \varepsilon_t^* = \phi^{-t}\varepsilon_t$$

and

$$E[\mathbf{X}_t \mathbf{X}_t'] = \mathbf{V}_d.$$

In (2.22) $\mathbf{T}$ is a two dimensional vector of positive constants and $\mathbf{V}_d$ is a two by two positive definite symmetric matrix of constants.

Relations (2.22) and (2.23) display the basic properties of the DSLR model which distinguish it from random walk versions of the model. First, the fact the (2,1) elements of $\mathbf{A}_1$ and $\mathbf{A}_2$ are not equal to zero implies that $\phi^{-t}\Delta c_t$ is

Granger caused by $\phi^{-t}(c_t - y_t)$. Second, the fact that the (2,2) element of $\mathbf{A}_1$ is not equal to zero implies that $\phi^{-(t-1)}\Delta c_{t-1}$ should be useful for predicting $\phi^{-t}\Delta c_t$.

3. Empirical Results for the Discrete Stochastic Labor Requirement Model

3.A Estimation Strategy

In section 2 we derived the implications of our model for the vector q_t^* which is defined in terms of detrended aggregate consumption and detrended net output, y_t^* (see (2.21)). We choose not to estimate the model using NNP data for two reasons. First, the data on aggregate depreciation is not particularly reliable. Second, there is little reason to believe that our model of depreciation is consistent with the model of depreciation used by the Department of Commerce. Consequently we implement our model in the following way using data on GNP. Let $\tilde{y}_t^*$ denote detrended gross output:

$$(3.1) \qquad \tilde{y}_t^* \equiv \phi^{-t}\tilde{y}_t$$

and define

$$(3.2) \qquad \tilde{q}_t' \equiv [c_t^* - \tilde{y}_t^*, c_t^* - \phi^{-1}c_{t-1}^*].$$

It follows that[3.1]

$$(3.3) \qquad q_t^* = H(\phi^{-1}L)\tilde{q}_t$$

where

$$(3.4) \qquad H(z) \equiv \frac{1}{1-(1-d)z}\begin{bmatrix} 1-z & 0 \\ 0 & 1-(1-d)z \end{bmatrix}.$$

By substituting (3.3) into (2.22), we obtain the implications of the model for $\tilde{q}_t$, which involves consumption and gross output.[3.2] The assumptions we have imposed on the structural parameters are sufficient to guarantee that $\tilde{q}_t$ is a covariance stationary stochastic process with conditionally homoscedastic disturbances. Since $|\phi^{-1}| < 1$, c_t and $\tilde{y}_t$ have unconditional growth rates equal to ϕ. Throughout our analysis we fix the value of ϕ at exp[.004568]. (See section 2.B for a description of the growth properties of our measures of consumption and output.) Given ϕ we form a time series on $\tilde{q}_t$ and denote the demeaned value of $\tilde{q}_t$ by Q_t.

Parameter estimates were obtained by maximizing the frequency domain approximation to the exact Gaussian likelihood function implied by the model. A full description of this estimator can be found in Hansen and Sargent [1981a].[3.3] In implementing this procedure we fixed the value of β at .99 which corresponds to a four percent annual rate of time preference. Consequently the free parameters of the DSLR model are a, f, d, and the three independent elements of V_d.

3.B Empirical Results

Table 3.1 summarizes the results of estimating the DSLR model. We use two methods to assess the overall performance of this model: (1) a formal statistical test of the overidentifying restrictions, and (2) an informal comparison of the constrained VAR for Q_t implied by the model with the corresponding unconstrained VAR.

Our formal statistical test is based on the fact that all of the models in this paper are nested within scalar autoregressive vector moving average (SARMA) representations for Q_t. While the DSLR model implies a constrained SARMA(3,3) for Q_t, the continuous time models of section 4 imply a constrained SARMA(3,4) for Q_t.[3.4] In order to allow all of the structural models to be nested within a common unconstrained specification we use the SARMA(3,4) as our unconstrained model.

Let J_T denote twice the difference between the maximized log likelihood for the unconstrained SARMA specification and the maximized log likelihood L_T, for the constrained SARMA specification. Then J_T is asymptotically distributed as a chi-square random variable with degrees of freedom equal to the number of restrictions imposed in the constrained specification. The test statistic J_T can be multiplied by an adjustment factor suggested by Whittle [1953], Lissitz [1972], and Sims [1980] designed to correct for small sample bias.[3.5] We denote the resulting test statistic by $\mathbf{J}_T$.

The values of J_T and $\mathbf{J}_T$ reported in Table 3.1 imply that the DSLR model is rejected at the five percent significance level, although not at the one percent significance level. However, this last result reflects in part the overparameterization of our unconstrained specification. When the model is compared to an unconstrained SARMA(3,3), the J_T statistic equals 30.36, with significance level .002, and $\mathbf{J}_T$ equals 28.75, with significance level .004. Thus, when compared to a more parsimoniously parameterized alternative, the DSLR model is rejected at the one percent significance level.[3.6]

When $d > 0$, the DSLR model implies that Q_t has an infinite ordered constrained VAR. However, when d is close to zero, this infinite ordered VAR is well approximated, in a sense to be made precise below, by a finite ordered VAR. As it turns out these VARs are more revealing for the purpose of diagnosing the empirical shortcomings of the model than the corresponding SARMA representations. We took as our benchmark an unconstrained VAR of lag length 2 since the hypothesis that the second lag is zero can be rejected at any conventional significance level, while the hypotheses that lags 3, 4, 5, and 6 are zero cannot be rejected at the five percent significance level. Table 3.1 displays the unconstrained VAR of Q_t and the truncated VAR implied by the DSLR model for Q_t. Here we use the truncation rule of not reporting matrix coefficients whose maximal elements are smaller than .02 in absolute value.

Recall that unlike random walk versions of the model, the DSLR model does not impose any zero restrictions on the law of motion for $\phi^{-t}\Delta c_t$. Consequently, it can in principle accommodate serial persistence in $\phi^{-t}\Delta c_t$ and any pattern of Granger causality between the elements of Q_t. However in practice our particular parameterization of the (e_t, H_t) process and the corresponding cross equation restrictions imposed by the DSLR model prevent it from fitting the degree of serial correlation observed in $\phi^{-t}\Delta c_t$. This can be seen by comparing the VAR implied by the DSLR model with the corresponding unconstrained VAR (see Table 3.1). Notice that the first row of coefficient matrices in these two VARs closely resemble each other. This suggests that the DSLR model fits the $\phi^{-t}(c_t-\tilde{y}_t)$ process fairly well. However the (2,2) element of the coefficient matrix on the first lagged value in the unconstrained VAR is .026, more than three standard deviations away from the unconstrained estimate of .313. Of course under the random walk hypothesis this coefficient would be zero, which is another way of seeing that the zero restrictions implied by that hypothesis are incompatible with the data.

4. The Continuous Time Permanent Income Hypothesis

4.A The Model

In this subsection we present the continuous time analogue of the discrete time model of consumption and output discussed in section 2. Our notation is the same as that used in sections 2 and 3 except that all random variables are assumed to evolve in continuous rather than discrete time. In addition, we adopt the convention of placing the time index of a continuous time random variable in parentheses.

The preferences of the representative consumer are given by:

$$(4.1) \qquad E_{0}\int_{0}^{\infty} e^{-rt}\{-\tfrac{1}{2}\,[c(t)-b(t)]^{2}-\alpha(t)h(t)\}dt,$$

where $b(t) = b\exp(\theta t)$, $\alpha(t) = \alpha\exp(2\theta t)$, r, b, α, $\theta > 0$, and $r-2\theta > 0$. In (4.1), $b(t)$ is the representative consumer's time t bliss point for consumption, $\alpha(t)$ measures the disutility of work at time t, $c(t)$ is consumption at time t and $h(t)$ work effort at time t.

As before, there is an aggregate technology that converts capital, $k(t)$, and labor effort into consumption goods:4.1/

$$(4.2) \qquad \tilde{y}(t) = \min\{\tilde{\delta}k(t),\tau(t)h(t)\} + e(t).$$

Here, $\tau(t)$ represents the (possibly) stochastic labor requirement per unit of capital, $\tilde{\delta} > 0$, and $e(t)$ is an aggregate shock to the time t production function. We impose the continuous time analogue to condition (2.4):

(4.3) $$r = \delta, \text{ where } \delta \equiv \tilde{\delta} - d,$$

where $d > 0$ is the depreciation rate on capital.

The economy-wide resource constraint is given by

(4.4) $$\tilde{y}(t) = c(t) + Dk(t) + dk(t),$$

where D denotes the time derivative operator.

The representative consumer's problem is to maximize (4.1) over contingency plans for setting $c(t)$, $Dk(t)$, $h(t)$, and $y(t)$ as a function of $I(t)$, subject to (4.2) and (4.4) and the constraint

(4.5) $$E_0 \int_0^\infty e^{-rt} k(t)^2 dt < \infty.$$

The set $I(t)$ is composed of all model variables dated t and earlier. We assume that

(4.6) $$\tilde{\delta} k(t) = \tau(t) h(t).$$

Relations (4.1), (4.2), (4.4) and (4.6) imply that the representative consumer's problem is to maximize:

(4.7) $$E_0 \int_0^\infty e^{-rt} \{- \tfrac{1}{2}[\delta k(t) - Dk(t) + e(t) - b(t)]^2 - H(t)k(t)\} dt$$

by choice of a contingency plan for $Dk(t)$, subject to (4.5). In (4.7), $H(t) \equiv \tilde{\delta}\alpha(t)/\tau(t)$. Let

(4.8) $$x_p(t) = \delta \int_0^\infty e^{-\delta\tau} E_t x(t+\tau) d\tau,$$

for any x process such that (4.8) converges. In Appendix B we show that the equilibrium laws of motion for $Dk(t)$ and $c(t)$ can be written as

$$(4.9)\qquad Dk(t) = e(t) - e_p(t) + b_p(t) - b(t) - H_p(t)/\delta$$

$$c(t) = e_p(t) + b(t) - b_p(t) + \delta k(t) + H_p(t)/\delta.$$

As before we let $y(t)$ denote net output: $y(t) = c(t) + Dk(t)$. Then

$$(4.10)\qquad c(t) - y(t) = e_p(t) - e(t) + [b(t)-b_p(t)] + H_p(t)/\delta$$

and

$$(4.11)\qquad Dc(t) = \mu_{e_p}(t) - \mu_{b_p}(t) + Db(t) + \mu_{H_p}(t)/\delta - H(t)$$

where $\mu_{x_p}(t)$ is the change in the value of $x_p(t)$ due to a disturbance in $x(t)$ that is unpredictable on the basis of $I(t-\tau)$, for all $\tau > 0$. (See Appendix B for a more careful discussion of this point.) Finally, we say that consumption satisfies the continuous time RWH if $Dc(t)$ is a continuous time white noise process with deterministic time varying drift. This does not imply that the detrended first difference of measured consumption will satisfy the discrete time random walk hypothesis. Consequently the empirical results of section 2.B cannot be used as evidence against this version of the model. In the next subsection we describe a parameterization of the unobserved exogenous forcing variables which satisfies the continuous time random walk hypothesis. We refer to this version of the model as the continuous time random walk (CRW) model.

4.B The Continuous Time Random Walk Model (CRW)

Given our assumptions on the $b(t)$ process, relation (4.11) implies that $c(t)$ is a random walk with deterministic drift if, and only if, $H(t)$ is deterministic. Accordingly, we assume

$$(4.12)\qquad H(t) = H\exp(\theta t),$$

where $H > 0$. The shock to endowment income is assumed to satisfy,

$$(4.13)\qquad e(t) = e_1(t) + e_2(t),$$

where

$$De_i(t) = e_i \exp(\theta t) + \frac{\eta_i(t)}{(a_i+D)},$$

where $a_i > 0$, $i = 1, 2$, and a_1 is not equal to a_2. Let $x(t) = [\eta_1(t)\ \eta_2(t)]'$. The vector $x(t)$ is the continuous time linear least squares innovation to the joint $[e_1(t)\ e_2(t)]$ process and satisfies,

$$(4.14)\qquad E[x(t)x(t-u)'] = \exp(2\theta t)\xi(u)\overline{V},$$

for all real values of u. Here, $\xi(u)$ is the Dirac delta generalized function and $\overline{V}$ is a two by two positive definite symmetric matrix of constants. Thus, $e(t)$ is the sum of two stochastic processes whose first derivatives are AR(1) continuous time stochastic processes. The reason for assuming that the endowment process is the sum of two stochastic processes, the realizations of which are separately observed by agents, is to guarantee that the observed bivariate consumption and income process is of full spectral rank. An alternative way of ensuring this condition is to suppose that $a_1 = a_2$ and assume that observations on the average value of consumption and income over the discrete sampling interval are contaminated by measurement error. Under these circumstances the true consumption and output process would not be of full spectral rank but the output and consumption process observed by the econometrician would not display any stochastic singularities. (See Hansen and Sargent [1980a] for a general discussion of error terms in linear rational expectations models.)

Substituting (4.12) and (4.13) into (4.11), we obtain

$$(4.15)\qquad Dc(t) = \eta_1(t)/(a_1+\delta) + \eta_2(t)/(a_2+\delta) + T_{c1}\exp(\theta t)$$

where T_{c1} is a positive scalar constant. According to (4.15), the derivative of consumption is a serially uncorrelated continuous time white noise process.4.2/

Substituting (4.12) and (4.13) into (4.10) we obtain,

$$(4.16)\qquad c(t) - y(t) = De_1(t)/(a_1+\delta) + De_2(t)/(a_2+\delta) + T_{c2}\exp(\theta t)$$

where T_{c2} is a positive scalar. We define the vector

$$q^*(t) = \left[c^*(t)-y^*(t),(D+\theta)c^*(t)\right]',$$

where $c^*(t) \equiv \exp[-\theta t]c(t)$, $y^*(t) \equiv \exp[-\theta t]y(t)$. Relations (4.13), (4.15), and (4.16) imply that $q^*(t)$ has a continuous time VAR(2) representation:

$$(4.17) \qquad A_c(D+\theta)q^*(t) = X(t) + T_c$$

where

$$(4.18) \qquad A_c(D) = I + A_{c1}D + A_{c2}D^2,$$

$$A_{c1} = \begin{bmatrix} (a_1+a_2)/(a_1a_2) & -1/(a_1a_2) \\ 0 & 0 \end{bmatrix}, \; A_{c2} = \begin{bmatrix} 1/(a_1a_2) & 0 \\ 0 & 0 \end{bmatrix},$$

$$X(t) = \begin{bmatrix} 1/a_1 & 1/a_2 \\ 1 & 1 \end{bmatrix} \begin{pmatrix} \eta_1^*(t)/(a_1+\delta) \\ \eta_2^*(t)/(a_2+\delta) \end{pmatrix},$$

$$\eta_i^*(t) = \exp[-\theta t]\eta_i(t), \; i = 1, 2,$$

and

$$E\left[X(t)X(t-u)'\right] = \xi(u)V_c.$$

In (4.17) T_c is a two dimensional vector of positive constants and in (4.20) V_c is a two by two positive definite symmetric matrix of constants. An implication of (4.17) and (4.18) is that $c(t)$ satisfies the continuous time RWH. This does <u>not</u> imply that measured consumption will satisfy the discrete time RWH.

4.C The Continuous Time Stochastic Labor Requirement Model (CSLR)

In this subsection we display the continuous time analogue to the discrete time model of section 2.C. Our specification of $b(t)$ and $\alpha(t)$ is the same as that given in section 4.B. However, we abandon the assumption that the labor requirement per unit of capital is nonstochastic. Instead, we assume

$$(4.19) \qquad H(t) = H\exp(\theta t) + \varepsilon(t)/(f+D)$$

where $E\left[\varepsilon(t)\varepsilon(t-u)'\right] = \exp(2\theta t)\xi(u)\sigma_\varepsilon^2$ and $\sigma_\varepsilon^2 > 0$. The shock to endowment income, $e(t)$, is assumed to satisfy

(4.20) $\quad De(t) = e\,\exp(\theta t) + \eta(t)/(a+D)$

where $E[\eta(t)\eta(t-u)] = \exp(2\theta t)\xi(u)\sigma_\eta^2$ and $\sigma_\eta^2 \geq 0$. Let $\mathbf{x}(t) = [\varepsilon(t)\ \eta(t)]'$. The vector $\mathbf{x}(t)$ is the continuous time linear least squares innovation to the joint $[H(t),e(t)]$ process and satisfies

(4.21) $\quad E[\mathbf{x}(t)\mathbf{x}(t-u)'] = \exp(2\theta t)\xi(u)\bar{\mathbf{V}}$

where $\bar{\mathbf{V}}$ is a two by two positive definite symmetric matrix of constants.

Substituting (4.19)-(4.20) into (4.10) and (4.11) we obtain

(4.22) $\quad c(t) - y(t) = De(t)/(a+\delta) + \delta H(t)/(f+\delta) + T_{c1}\exp(\theta t)$

$$Dc(t) = \eta(t)/(a+\delta) + [D-\delta]\delta H(t)/(f+\delta) + T_{c2}\exp(\theta t)$$

where T_{c1} and T_{c2} are positive scalar constants. Relations (4.19)-(4.22) imply that $q^*(t)$ has the continuous time VAR(2) representation:

(4.23) $\quad \mathbf{A}_c(D+\theta)q^*(t) = \mathbf{X}(t) + \mathbf{T}_c$

where

(4.24) $\quad \mathbf{A}_c(D) = I + \mathbf{A}_{c1}D + \mathbf{A}_{c2}D^2,$

$$A_{c1} = \begin{bmatrix} a+f\delta/a & f/a-1 \\ -a(f+\delta) & f+\delta \end{bmatrix}/f(a+\delta),\quad \mathbf{A}_{c2} = \begin{bmatrix} 1-f/a & 0 \\ -(f+\delta) & 0 \end{bmatrix}/f(a+\delta),$$

$$\mathbf{X}(t) = \begin{bmatrix} 1/a_1 & 1/f \\ 1 & -\delta/f \end{bmatrix} \begin{pmatrix} \eta^*(t)/(a+\delta) \\ \varepsilon^*(t)/(f+\delta) \end{pmatrix},$$

$$\eta^*(t) = \exp[-\theta t]\eta(t),\quad \varepsilon^*(t) = \exp[-\theta t]\varepsilon(t),$$

and

$$E[\mathbf{X}(t)\mathbf{X}(t-u)'] = \xi(u)\mathbf{V}_c.$$

In (4.23) $\mathbf{T}_c$ is a two dimensional vector of positive constants and in (4.24) $\mathbf{V}_c$ is a two by two dimensional positive definite symmetric matrix of constants. Relations (4.23) and (4.24) imply that c(t) does not satisfy the continuous time RWH.

5. Empirical Results for the Continuous Time Models

Before considering the empirical performance of the fully parameterized versions of our continuous time model, we first consider implications of the CRW model which do not depend on the specific parameterization of the e(t) process. From the perspective of the continuous time model, measured consumption is the time average of consumption over the discrete time sampling interval. It then follows that measured Δc_t is the average of innovations to underlying continuous time consumption from the beginning of quarter t - 1 to the end of quarter t. Since the quarter t - 1 innovations also appear in Δc_{t-1}, the continuous time RWH implies that Δc_t and Δc_{t-1} have nonzero covariance. In addition, Δc_t and y_{t-1} will also be correlated because continuous time output is correlated with previous innovations to continuous time consumption and y_{t-1} is the average value of output during quarter t - 1. For these reasons, time aggregation can in principle account for the rejections of the RWH reported in section 2.B. A simple extension of the preceding argument shows that the continuous time RWH implies that the detrended first difference of measured consumption is uncorrelated with consumption and income lagged two periods and more. In subsection 5.A we test the latter orthogonality conditions, and the continuous time RWH's implications for the autocorrelation structure of $\phi^{-t}\Delta c_t$. In subsection 5.B we discuss the way in which the fully parameterized versions of the continuous time models are estimated and tested. Finally, in subsection 5.C we report empirical results for the CRW and CSLR models.

5.A Simple Tests of the Continuous Time RWH

Abstracting from growth, Working [1960] showed that the first difference of a time averaged continuous time random walk has an autocorrelation at lag one of .25 and zero at higher lags. Christiano and Marshall [1987] showed that, after rounding to three digits, this result is also valid when $\phi = \exp[.004568]$. Thus, the continuous time RWH implies that ψ_t^0, defined in (2.15)', has mean zero and autocorrelation .25 and 0 at lags 1 and 2, respectively.

We began our investigation of the continuous time RWH by estimating the autocorrelation of $\phi^{-t}\Delta c_t$ at lag 1. Define the functions

$$(5.1) \qquad H_t \equiv H(\phi^{-t}\Delta c_t, \rho_1, C) = [\psi_t, (\psi_t\psi_{t-1} - \rho_1\psi_t^2)]',$$

and

$$(5.2) \qquad H_t^0 \equiv H(\phi^{-t}\Delta c_t, \rho_1^0, C^0),$$

where ψ_t is defined in (2.15)' and $\rho_i^0 = E\psi_t^0\psi_{t-i}^0/E(\psi_t^0)^2$, i = 1, 2, ... Under the continuous time RWH, $\rho_1^0 = .25$. Moreover, $E_{t-2}H_t^0 = 0$, which implies $EH_t^0 = 0$.5.1/ Since $\phi^{-t}\Delta c_t$ is by assumption a stationary and ergodic stochastic process, this set of unconditional moment restrictions can be used to estimate the parameters C^0 and ρ_1^0 using the GMM procedure described in Hansen [1982]. Define the function $g_T = (1/T)\sum_{t=1}^{T} H_t$. Our estimator of (C^0, ρ_1^0) is the argmax of $g_T' W_T^{-1} g_T$, where W_T is a consistent estimate of the spectral density of H_t^0 evaluated at frequency 0.5.2/

Table 5.1 reports point estimates and standard errors for ρ_1. The column marked c_{nd} denotes the consumption concept used in Flavin [1981] and the column marked $c_{nd} + c_{sd}$ denotes the consumption concept used in Hall [1978]. Notice that for all measures of consumption and all sample periods, the estimated value of ρ_1 is well within one standard deviation of .25. Proceeding as above, we also used Hansen's [1982] GMM procedure to estimate and test the null hypothesis $\rho_2^0 = 0$. In no case can we reject this null hypothesis, at the five percent significance level. Consequently, this set of tests provides virtually no evidence against the continuous time RWH.

Next, we tested the continuous time RWH's implication that $\phi^{-t}\Delta c_t$ is uncorrelated with elements of agents' time t - 2 information sets, i.e., $E_{t-2}\psi_t^0 = 0$. This conditional moment restriction implies the unconditional moment restriction $E\psi_t^0\hat{z}_{it} = 0$ for all $\hat{z}_{it}$ contained in I_{t-2}, i = 1, 2. In practice $\hat{z}_{1t}$ and $\hat{z}_{2t}$ were specified as:

$$(5.3) \qquad \hat{z}_{1t} = [1, \phi^{-(t-2)}\Delta c_{t-2}, \ldots, \phi^{-(t-4)}\Delta c_{t-4}, \phi^{-(t-2)}\tilde{y}_{t-2}, \ldots, \phi^{-(t-4)}\tilde{y}_{t-4}]'$$

and,

$$(5.3)' \quad \hat{z}_{2t} = [1, \phi^{-(t-2)}\Delta c_{t-2}, \ldots, \phi^{-(t-4)}\Delta c_{t-4}, \phi^{-(t-2)}(c_{t-2} - \tilde{y}_{t-2}), \ldots, \phi^{-(t-4)}(c_{t-4} - \tilde{y}_{t-4})]'.$$

We tested the null hypotheses $E\psi_t^0\hat{z}_{it} = 0$, i = 1, 2 using the GMM procedure described in section 2.B. In particular, our estimator of C^0 is the argmax of $J_{iT} = g_{iT}'W_T^{-1}g_{iT}$, where $g_{iT} = (1/T)\sum_{t=0}^{T}\psi_t\hat{z}_{it}$ and W_T is a consistent estimate of the spectral density matrix of $\psi_t^0\hat{z}_{it}$ evaluated at frequency zero. Under our null hypothesis, the minimized value of J_{it}, i = 1, 2 is asymptotically distributed as a chi-square random variable with six degrees of freedom.

Significance levels of the computed test statistics appear in the "Lags 2 - 4" portion of Table 5.2. Three features of these results are worth noting. First, in only one case do the significance levels in the "Lags 2 - 4" part of the table fail to exceed their counterparts in the "Lags 1 - 4" part of the table. Second, in only one case can we reject the null hypothesis, $E\psi_t^0 z_{1t} = 0$, at the one percent significance level. Also, in only one case can we reject the null hypothesis $E\psi_t^0 z_{2t} = 0$ at the five percent significance level.

We conclude this section by reporting the results of testing the joint hypotheses, $\{E\psi_t^0\hat{z}_{it}=0, \rho_1^0=.25\}$, i = 1, 2. Each of these null hypotheses imply the set of nine unconditional moment restrictions, $EH_t^0 Z_{it} = 0$, where Z_{it} is the two by nine block diagonal matrix with first and second diagonal blocks $\hat{z}_{it}'$ and 1, respectively, i = 1, 2. Our estimator of parameters ρ_1^0 and C^0 were the argmax of $J_{iT} = g_{iT}'W_{iT}^{-1}g_{iT}$, where $g_{iT} = (1/T)\sum_{t=1}^{T}H_t Z_{it}$ and W_{iT} is a consistent estimate of the spectral density matrix of $H_t^0 Z_{it}$ evaluated at frequency zero, i = 1, 2. The minimized value of J_{iT}, i = 1, 2 is asymptotically distributed as a chi-square random variable with seven degrees of freedom. Significance levels of the test statistics are reported in Table 5.3. Two important results emerge here. First, we can never reject the null hypothesis that $EH_t^0 Z_{2t} = 0$ at the one percent significance level. Second, in only one case can we reject the null hypothesis that $EH_t^0 Z_{1t} = 0$

at the one percent significance level. However, this exception occurs when our data set is used over the sample period 1951,3-1985,3 in which case the significance level of the chi-square statistic is .008.

In sum, the evidence against the continuous time RWH is sufficiently weak to warrant investigating the CRW model described in section 4.B. Since that model embodies a simple parameterization of the e(t) process, it allows us to investigate the cross dynamics between $\phi^{-t}\Delta c_t$ and $\phi^{-t}(c_{t-1}-\tilde{y}_{t-1})$. At the same time the evidence reported in this subsection with our data set contains sufficient evidence against the continuous time RWH to suggest that deviations from the CRW model are worth investigating. In subsection 5.B we discuss the ways in which the CRW and CSLR models are estimated and tested.

5.B Estimation Strategy for CRW and CSLR Models

In section 4 we derived the constrained continuous time VAR representations for $q^*(t)$ implied by the CRW and CSLR models [see (4.17)-(4.18) and (4.23)-(4.24) respectively]. In order to proceed with estimation we must deduce the implications of these VARs for the probability law of the vector of observable variables. We define $\tilde{q}(t)$ to be the 2x1 continuous time stochastic process whose first element is the difference between detrended quarterly averaged consumption and gross output, and whose second element is the detrended first difference of quarterly averaged consumption. The vectors $q^*(t)$ and $\tilde{q}(t)$ differ in two important respects. First, $q^*(t)$ involves a measure of detrended NNP, whereas $\tilde{q}(t)$ involves a measure of detrended GNP. Second, $q^*(t)$ represents point in time measured variables, whereas $\tilde{q}(t)$ represents variables which have been averaged over the discrete data sampling interval.

Our strategy for obtaining the probability law for $\tilde{q}(t)$ is to derive the linear mapping relating $q^*(t)$ and $\tilde{q}(t)$, and then to use this expression to substitute out for $q^*(t)$ in terms of $\tilde{q}(t)$ in (4.17) and (4.23). We proceed by first obtaining the linear mapping between undetrended $q^*(t)$ and undetrended $\tilde{q}(t)$. Let

$z(t)$ denote the undetrended process underlying $\tilde{q}(t)$, i.e., $z(t) \equiv [c(t)-\tilde{y}(t), Dc(t)]'$. Let $\bar{z}(t)$ denote the undetrended, averaged data underlying $\tilde{q}(t)$, ie., $\bar{z}(t) = \exp[\theta t]\tilde{q}(t)$. Formally,

$$(5.4) \qquad \bar{z}(t) = \begin{pmatrix} \int_0^1 [c(t-\tau)-\tilde{y}(t-\tau)]d\tau \\ \\ \int_0^1 [c(t-\tau)-c(t-1-\tau)]d\tau \end{pmatrix} = \begin{pmatrix} \int_0^1 [c(t-\tau)-\tilde{y}(t-\tau)]d\tau \\ \\ \int_0^1 [\int_0^1 Dc(t-\tau-\mu)d\mu]d\tau \end{pmatrix}.$$

Here we have used the fact that $\int_0^1 Dc(t-\mu)d\mu = c(t) - c(t-1)$.[5.4] In operator notation:[5.5]

$$(5.5) \qquad \bar{z}(t) = G(D)z(t),$$

where

$$(5.6) \qquad G(D) = [(1-e^{-D})/D]\begin{bmatrix} 1 & 0 \\ 0 & [(1-e^{-D})/D] \end{bmatrix}.$$

Let $q(t)$ denote the undetrended value of $q^*(t)$, i.e., $q(t) = \exp(\theta t)$ $q^*(t) = [c(t)-y(t), Dc(t)]$. In operator notation, the link between $q(t)$ and $z(t)$ is given by[5.6]

$$(5.7) \qquad q(t) = H(D)z(t)$$

where

$$(5.8) \qquad H(D) \equiv \left[\frac{1}{D+d} \begin{matrix} D & 0 \\ 0 & D+d \end{matrix}\right].$$

Substituting (5.7) into (5.5), we obtain

$$(5.9) \qquad q(t) = H(D)G(D)^{-1}\bar{z}(t)$$

which provides a mapping between the continuous time processes $q(t)$ and $\bar{z}(t)$ i.e., between undetrended $q^*(t)$ and undetrended $\tilde{q}(t)$. Finally, the link between $\tilde{q}(t)$ and $q^*(t)$ is obtained by multiplying both sides of (5.9) by $\exp(-\theta t)$:

$$(5.10) \qquad q^*(t) = H(D+\theta)G(D+\theta)^{-1}\tilde{q}(t).$$

Substituting (5.10) into (4.17) and (4.23), we obtain the time series representations for $\tilde{q}(t)$ implied by the CRW and CSLR models.

We now describe the procedure used to estimate the continuous time models. Define

$$(5.11) \qquad Q(t) \equiv \tilde{q}(t) - E\tilde{q}(t).$$

Suppose we have a sample on $Q(t)$, $t = 1, 2, 3, \ldots, T$. Our estimation criterion is the frequency domain approximation to the Gaussian density function suggested by Durbin [1961], Hannan [1970] and Hansen and Sargent [1981a].[5.7] This criterion requires that we compute the theoretical spectral density of the discrete process $\{Q(t),\ t \text{ integer}\}$ at frequency ω, $Z(\omega)$. We accomplish this in two steps. First, using results in Phillips [1958] it can be shown that the spectral density of $\{Q(t),\ t \text{ real}\}$ implied by the CRW model and equations (4.17) and (5.10) is given by:

$$(5.12) \qquad Z^c(\omega) = \psi(i\omega+\theta)A_c(i\omega+\theta)^{-1}V_c[A_c(-i\omega+\theta)']^{-1}\psi(-i\omega+\theta)'$$

for $-\infty \le \omega \le \infty$, where $\psi(s) \equiv G(s)H(s)^{-1}$. The corresponding spectral density implied by CSLR model is

$$(5.13) \qquad \mathbf{Z}^c(\omega) = \psi(i\omega+\theta)\mathbf{A}_c(i\omega+\theta)^{-1}\mathbf{V}_c[\mathbf{A}_c(-i\omega+\theta)']^{-1}\psi(-i\omega+\theta)'.$$

Second, Hannan [1970, p. 45] shows that the following "folding operator" links $Z(\omega)$ and $Z^c(\omega)$:

$$(5.14) \qquad Z(\omega) = \sum_{k=-\infty}^{\infty} Z^c(\omega+2\pi k).$$

Equations (5.12) or (5.13) and (5.14) provide a computationally feasible algorithm for obtaining $Z(\omega)$ for a given ω from $[\psi, A, V_c]$ or $[\psi, \mathbf{A}, \mathbf{V}_c]$. Because this algorithm is relatively slow, we used an alternative method based on a partial fractions decomposition of Z^c (see Durbin [1961], Hannan [1970, pp. 405-407] and Hansen and Sargent [1981]).

The preceding estimation strategy assumes that the values of θ and $E\tilde{q}(t)$ are known. We proceeded as in the discrete time case, by replacing $E\tilde{q}(t)$ by its sample mean and setting θ to .004568.[5.8] Finally we note that an implication of results in Christiano and Marshall [1987], is that both the CRW and CSLR models give rise to constrained SARMA(3,4) representations for $Q(t)$.[5.9]

5.C Empirical Results

Our estimates of the CRW and CSLR models are reported in Table 5.4. The parameter r is set equal to .0098, which implies an annual rate of time preference of four percent.[5.10] According to both the unadjusted and adjusted likelihood ratio statistics (J_T and J_T respectively) neither the CRW nor the CSLR model can be rejected, at the five percent significance level.[5.11] This is to be contrasted with our findings that the discrete RWH and the DSLR model can be rejected at close to the one percent level. Thus there is some evidence that the continuous time formulations are in greater conformity with the data than their discrete time counterparts.

The large number of parameters in the unconstrained SARMA used to construct the likelihood ratio tests raises questions regarding the power of our specification tests. Since the SARMA(3,4) is the most parsimoniously parameterized unconstrained model that nests the continuous time structural models, we cannot formally compare the performance of these models with a more tightly parameterized alternative. However, the point estimates reported in Table 5.4 suggest a way of reformulating the continuous time models so that they are nested in an unconstrained SARMA(2,3). The point estimates of a_2 and a are extremely large so that the $e_2(t)$ and $e(t)$ processes are virtually indistinguishable from continuous time random walks.[5.12] It follows that in the SARMA(3,4) representations implied by both the CRW and CSLR models, the MA matrix coefficient in the fourth lag and the AR coefficient on the third lag are approximately zero.[5.13] Hence we can compute likelihood ratio statistics by comparing the likelihood values given in

Table 5.4 with the value obtained for the unconstrained SARMA(2,3). The resulting test statistics, which are distributed asymptotically with 12 degrees of freedom, are given below. (Significance levels are in parentheses.)

	CRW Model	CSLR Model
J_T	22.74 (.030)	22.20 (.035)
J_T	21.61 (.042)	21.10 (.049)

According to these results, neither model is rejected at the three percent significance level.

In Table 5.5 we report the constrained VAR for Q_t implied by our estimates of the CRW model. In principle this VAR is infinite ordered, so we use the truncation rule of not reporting matrix coefficients whose maximal element are smaller than .02 in absolute value. Notice that the CRW model and discrete random walk versions of the model differ substantially in their implications for $\phi^{-t}\Delta c_t$. By construction, the implies that $\phi^{-t}\Delta c_t$ is uncorrelated with lagged values of both $\phi^{-t}\Delta c_t$ and $\bar{\phi}^{t}(c_t-\tilde{y}_t)$. While the CRW model embodies this restriction for the continuous time point-in-time sampled data, it does <u>not</u> imply this restriction for the actual measured, discrete time data.

Our evidence suggests that the effect discussed by Working [1960] is the major factor accounting for the improved fit of the CRW model relative to the discrete time RWH model since it accounts for the coefficient 0.27 that appears on the first own lag of $\phi^{-t}\Delta c_t$ in the constrained VAR. This is within one standard error of the point estimate (.313) of the corresponding coefficient in the unconstrained VAR. Moreover the first own lag on $\phi^{-t}\Delta c_t$ in the unconstrained VAR is more than three standard deviations away from zero. Taken together these observations suggest that the change in the value of the coefficient of once lagged $\phi^{-t}\Delta c_t$ on $\phi^{-t}\Delta c_t$ from 0 in discrete random walk versions of the model to 0.27 in the CRW model has a substantial effect on the likelihood ratio statistic.

The nonzero values of the coefficients on lagged values of $\phi^{-t}(c_t-\tilde{y}_t)$ in the second row of the VAR in Table 5.5 corresponding to the CRW model also reflect the effects of time averaging. However, when we compare these point estimates to the corresponding entries in the unconstrained VAR we see that this effect may be harmful with regards to the overall fit of the CRW model. This is because the sign on $\phi^{-(t-1)}(c_{t-1}-\tilde{y}_{t-1})$ in the $\phi^{-t}\Delta c_t$ equation of the constrained VAR is positive, in contrast to the negative sign of the corresponding term in the unconstrained VAR. Since the latter coefficient is not precisely estimated, this effect is not sufficiently important to negate the favorable impact of the effects suggested by Working [1960].

Next, we contrast the empirical performance of the DSLR and CSLR models. The constrained VAR(2) implied by the DSLR model for Q_t is reported in Table 3.1. The corresponding VAR implied by the CSLR model is reported in Table 5.5. Because the constrained VAR's are in principle infinite ordered we again use the truncation rule of not reporting matrices whose maximal element is smaller than .02 in absolute value. Comparing Tables 3.1 and 5.5 we see that the DSLR and CSLR models do not differ in any substantial way regarding the dynamics of $\phi^{-t}(c_t-\tilde{y}_t)$. However they do differ substantially in their implications for $\phi^{-t}\Delta c_t$. In the VAR corresponding to the CSLR model the coefficient on $\phi^{-(t-1)}\Delta c_{t-1}$ in the $\phi^{-t}\Delta c_t$ equation is approximately .27 while the corresponding coefficient in the DSLR model is approximately .03. Thus, the principal difference between the DSLR and CSLR models is that the latter model is able to handle substantially more serial correlation in $\phi^{-t}\Delta c_t$.

In summary we find that the CRW and CSLR models appear to be empirically more plausible than discrete random walk versions of the model and the DSLR model. In our view the evidence against the CRW model and the CSLR model is far from overwhelming. This is surprising given the simplicity and parsimonious parameterization of both these models. In both instances the impact of moving to a continuous time model is an enhanced ability to mimic the serial correlation properties of the quasi difference of consumption.

6. Conclusion

This paper develops and tests fully specified equilibrium models of consumption and output which are consistent with the fact that measured aggregate consumption does not behave as a random walk. It is not particularly challenging to develop theories which can explain this fact in principle. The random walk hypothesis is clearly a special case of the permanent income hypothesis. However, as much of the recent literature on the macroeconomics of consumption reveals, it is quite challenging to develop **empirically plausible** models of the comovements in aggregate consumption and output.

We investigated two possible reasons why the change in consumption fails to behave like a white noise. The first possibility is that exogenous shocks to the economic system generate serial persistence in the first difference of consumption. We modeled this shock as a stochastic perturbation to the amount of labor required to make capital productive. As it turns out, there is a great deal of evidence against this version of our model when it is implemented under the assumption that agents' decision intervals coincide with the data sampling interval. However, there is surprisingly little evidence against the continuous time version of this model.

The second possibility is that the RWH holds in the (unobserved) continuous consumption process, with serial persistence in measured consumption being an artifact of temporal aggregation. Our results indicate that when temporal aggregation bias is taken into account, the fit of the random walk model improves substantially. This suggests that the random walk hypothesis may yet be a useful way to conceptualize the relation between aggregate consumption and output.

While both of the continuous time models that we tested outperform their discrete time counterparts, it is very difficult, at least on the basis of aggregate consumption and output data, to distinguish between the two continuous time models. However the CRW model does have a number of implications which we did not test in this paper but which call its plausibility into question. One such implication is that the capital-labor ratio is deterministic. This implication is obviously counterfactual. While this could be remedied by allowing for measurement error, we regard the CLSR model as a more promising starting point for future research.

A different set of implications which were not explored in this paper concern the equilibrium wage rate and real interest rate. Unlike the quantity variables, our models imply that these price processes are nonlinear functions of the state variables in the system (see footnote 2.5). Consequently, deriving the laws of motion for measured wages and interest rates that are implied by our continuous time models involves technical difficulties not encountered in this paper. Nonetheless, we believe that our results for the consumption and output are sufficiently encouraging to warrant an empirical investigation of the model's implications for relative prices.

Footnotes

2.1/ In a formulation which allows for positive population growth, the expression on the left hand side of (2.3) must be replaced by $c_t + k_t - [(1-d)/n]k_{t-1}$, where n denotes the gross growth rate of the population.

2.2/ Hansen's [1986] model differs from ours in that he sets $\alpha_t \equiv 0$.

2.3/ This terminology is slightly unconventional since β^{-1} is not the gross rate interest in our model economy. (See footnote 2.5.)

2.4/ Hall [1978] and Flavin [1981] do not distinguish between the hypothesis that consumption follows a random walk (possibly with drift) and the hypothesis that consumption is a martingale (possibly with drift). While the latter hypothesis is the actual focus of attention in the literature, it is typically referred to as the random walk hypothesis (RWH).

2.5/ The derivation of the RWH in this paper and the derivation in Hall [1978] impose strong restrictions on the underlying economic model. Hall derives the RWH by directly restricting the stochastic structure of the risk free real rate of interest, r_t, which he assumes to be constant. In our model, however, the risk free real rate of interest, denominated in units of the consumption good, is stochastic even under those circumstances for which the RWH is satisfied. To see this notice that the representative consumer's intertemporal Euler equation for one period risk free consumption loans can be written as $(b_t-c_t) = \beta r_t E_t(b_{t+1}-c_{t+1})$. Relation (2.14) implies $E_t(b_{t+1}-c_{t+1}) = (b_t-c_t) + H_t$. Consequently, r_t equals $\beta^{-1}[1+H_t/(b_t-c_t)]^{-1}$ so that r_t will be stochastic even if H_t is deterministic (so that the RWH is satisfied) as long as e_t is stochastic. Thus a constant risk free real interest rate is not a necessary condition for the RWH to hold.

2.6/ Our measure of government investment is a revised and updated version of the measure discussed in Musgrave [1980].

2.7/ Our estimate of the size and trend of the government's share in total consumption may be distorted by the fact that we ignore the service flow from the stock of government capital in our measure of consumption and output.

2.8/ Hayashi [1982] also makes this assumption. (See also footnote 2.10 below.)

2.9/ In contrast to our measures of consumption and income $(c,\tilde{y})$ and in contradiction to the models of sections 2 and 4, the per capita growth rates of consumption and income in both these data sets are quite different. Over the sample period 1950,2-1985,3 the growth rates in $c_{nd} + c_s$, c_{nd}, and y_d are .004987, .003076, and .005502 respectively. (That $c_{nd}+c_s$ has a higher growth rate than c_{nd} alone is consistent with the evidence in Table 2.1).

2.10/The idea of inducing stationarity in the observable time series by geometrically scaling variables by the growth rate of the standard error in the innovations to the underlying shocks in linear-quadratic rational expectations models is discussed in Hansen and Sargent [1981]. We know of no analytically tractable alternative way to accommodate the observed heteroskedasticity in the data, which is consistent with the essential linearity of our model.

3.1/Equation (3.4) can be seen as follows:

$$q_t^* = \begin{pmatrix} c_t^*-y_t^* \\ c_t^*-\phi^{-1}c_{t-1}^* \end{pmatrix} = \begin{pmatrix} -(k_t^*-\phi^{-1}k_{t-1}^*) \\ c_t^*-\phi^{-1}c_{t-1}^* \end{pmatrix} = \begin{pmatrix} -\dfrac{(1-\phi^{-1}L)}{1-(1-d)\phi^{-1}L}\left[k_t^*-(1-d)\phi^{-1}k_{t-1}^*\right] \\ c_t^*-\phi^{-t}c_{t-1}^* \end{pmatrix}$$

$$= H(\phi^{-t}L)\tilde{q}_t$$

where the last equality follows from (3.1) and (3.2) and $k_t^* = \phi^{-t}k_t$.

3.2/According to the DSLR model $\tilde{q}_t$ has a constrained ARMA(3,1) representation.

3.3/Let $Q_t = \tilde{q}_t - E\tilde{q}_t$ and let $Z(\omega)$ and $I(\omega)$ denote the theoretical spectral density matrix and the periodogram of the Q_t process respectively at frequency $\omega_j = 2\pi j/T$, $j = 1, 2, \ldots, T$. Using results in Hannan [1970] it can be shown that the Gaussian log likelihood function can be approximated by

$$L_T = -.5T\log[2\pi] - .5\sum_{j=1}^{T-1}\log\{\det[Z(\omega_j)]\} - .5\sum_{j=1}^{T-1}\mathrm{tr}[Z(\omega_j)^{-1}I(\omega_j)].$$

The theoretical spectral density matrix of the Q(t) process is

$$Z(\omega_j) = H(\phi^{-1}e^{-i\omega_j})^{-1}A(\phi^{-1}e^{-i\omega_j})^{-1}\,V_dA(\phi^{-1}e^{-i\omega_j})^{-1'}H(\phi^{-1}e^{-i\omega_j})^{-1'}.$$

3.4/That Q_t is predicted to be SARMA(3,3) is proved in Christiano, Eichenbaum, and Marshall [1987].

3.5/Whittle's [1953] correction for small sample bias is as follows: Let N equal the total number of parameters under the alternative hypothesis (excluding the covariance matrix of the observables), M = number of equations, and T = number of observations. Then $J_T = J_T(1-N/MT)$ where J_T is the unadjusted likelihood ratio statistic and J_T is the adjusted likelihood ratio statistic. When the unconstrained alternative is a SARMA(3,4), N = 19, M = 2, and T = 141, so J_T = $0.9326241J_T$.

3.6/ We also estimated the DSLR and unconstrained SARMA(3,4) models using a one step version of the estimation method described in section 3.A in which the growth rate ϕ is estimated simultaneously with the other parameters of the model. Under the assumption that J_T is asymptotically distributed as a Chi-square random variable with degrees of freedom equal to the number of restrictions imposed in the constrained model we can summarize our results as follows. Testing the DSLR model, which has 16 degrees of freedom, yields a J_T statistic of 32.67 with associated significance level .008. Thus this model is rejected at the one percent significance level.

4.1/ Equation (4.2) differs from (2.2) in the timing of the productive inputs. (4.2) results as a limiting case of (2.2) if we rewrite the latter as $\tilde{y}_t = \min\{\tilde{\delta}_{t-\varepsilon}, \tau_{t-\varepsilon}h_{t-\varepsilon}\} + e_t$ and let $\varepsilon \to 0$ from above.

4.2/ While Dc(t) is not a physically realizable process, its average over any discrete interval of time is physically realizable.

5.1/ To see that $E_{t-2}H_t^0 = 0$, consider the Wold representation $\psi_t^0 = \epsilon_t + \theta\epsilon_{t-1}$, where $|\theta| < 1$ and ϵ_t is fundamental for ψ_t^0. This representation exists and is unique under the stationarity assumption on ψ_t^0 used to justify GMM. Then $E_{t-2}\psi_t^0\psi_{t-1}^0 = E_{t-2}(\epsilon_t+\theta\epsilon_{t-1})(\epsilon_{t-1}+\theta\epsilon_{t-2}) = \theta E\epsilon_t^2 = E\psi_t^0\psi_{t-1}^0$ which, together with the definition of ρ_1^0, establishes $E_{t-2}[\psi_t^0\psi_{t-1}^0-\rho_1^0(\psi_t^0)^2] = 0$. Trivially, $E_{t-2}\psi_t^0 = 0$, establishing the result sought.

5.2/ The condition, $E_{t-2}H_t^0 = 0$ implies that H_t^0 is autocorrelated at lag one but not higher. Consequently, inference was carried out using Hansen's [1982] correction for serial correlation, imposing the restriction that H_t^0 has an MA(1) representation. See Hansen, Heaton, and Ogaki [1987] for a discussion of the efficiency gains associated with imposing the exact MA(1) structure on error terms in GMM estimation problems. Similar MA(1) corrections were used for all the inference carried out in this subsection.

5.4/ Treating measured consumption and income as unit integrals of the underlying instantaneous quantities is a rough approximation to the methods used by the Department of Commerce to gather data.

5.5/ In deriving (5.6) we use the fact that

$$\int_0^1 x(t-\tau)d\tau = \int_0^1 e^{-\tau D}x(t)d\tau = \left[(1-e^{-D})/D\right]x(t).$$

5.6/ Equation (5.8) can be derived as follows:

$$q(t) = \begin{pmatrix} c(t)-y(t) \\ Dc(t) \end{pmatrix} = \begin{pmatrix} -Dk(t) \\ Dc(t) \end{pmatrix} = H(D)\begin{pmatrix} -(D+d)k(t) \\ Dc(t) \end{pmatrix} = H(D)\tilde{q}(t)$$

where the last equality follows from (4.4) and (5.5).

5.7/ For time domain methods of estimating continuous time models from discrete data, see Bergstrom [1983], Harvey and Stock [1986] and Zadrozny [1984].

5.8/ The relationship between θ and ϕ is $\phi = e^{\theta}$.

5.9/ The particular SARMA representation corresponding to a given continuous time model is characterized by a third order scalar polynomial, $E^d(\cdot)$, the two by two fourth order matrix polynomial, $C^d(\cdot)$, and the two by two positive semidefinite matrix V^d which satisfy:

$$Z(\omega) = C^d(e^{-i\omega})V^dC^d(e^{i\omega})'/[E^d(e^{-i\omega})E^d(e^{i\omega})].$$

Here, we impose the normalizations $C^d(0) = I$, $\det[C^d(z)] = 0$ implies $|z| \geq 1$, and $E^d(0) = 1$. The algorithm we used to calculate E^d, C^d, and V^d is the one described in Rozanov [1967, chapter I, section 10]. Thus both continuous time models are nested within the SARMA specification:

$$E^d(L)Q_t = C^d(L)X_{ct},$$

where X_{ct} is the serially uncorrelated innovation in $Q(t)$, with variance V^d, and

$$E^d(L) = 1 + E_1^dL + E_2^dL^2 + E_3^dL^3,$$

$$C^d(L) = I + C_1^dL + C_2^dL^2 + C_3^dL^3 + C_4^dL^4.$$

5.10/ The relationship between β in the discrete formulations and r in the continuous formulations is $\beta = e^{-r}$.

5.11/ We also estimated the CRW, CSLR, and unconstrained SARMA(3,4) models using a one step version of the estimation method described in section 5.B in which the growth rates of consumption and output are estimated simultaneously with the other parameters of the model. Under the assumption that J_T is asymptotically distributed as a Chi-square random variable with degrees of freedom equal to the number of restrictions imposed in the constrained model we can summarize our results as follows: (i) testing the CRW model, which has 16 degrees of freedom, yields a J_T statistic of 23.10 with associated significance level .111, (ii) testing the DSLR model, which has 16 degrees of freedom, yields a J_T statistic of 24.35 with associated level .080. Thus neither continuous time model can be rejected at the five percent significance level.

5.12/ That the reported point estimates imply that $e_2(t)$ and $e(t)$ behave essentially as continuous time random walks can be seen from the following argument. If $x(t)$ is a continuous time first order autoregression: $x(t) = \varepsilon(t)/(a+D)$, $\varepsilon(t)$ continuous time white noise, then $x(t)$ has an exponentially declining impulse response function:

$$x(t) = \int_0^\infty e^{-a\tau}\varepsilon(t-\tau)d\tau.$$

Therefore, the impulse response function for $De_2(t)$ is $e^{-28\tau}$ and for $De(t)$ is $e^{-12\tau}$. These functions decline so steeply that past impulses have negligible effect on current values of $De_2(t)$ and $De(t)$.

5.13/ In the SARMA(3,4) implied by the CRW model, the MA matrix coefficient on the fourth lag consists entirely of zeroes and the AR coefficient on the third lag equals 0.9×10^{-12}. In the case of the CSLR model, the MA matrix coefficient on the fourth lag has no element greater than 3×10^{-7} in absolute value and the AR coefficient on the third lag is -7×10^{-6}.

Table 2.1

Time Average of Relative Magnitude of Components of Our Measure of Total Consumption[1]

Sample Period	c_{nd}/c	c_s/c	c_{sd}/c	c_g/c	i_g/g
50, 1-59, 4	.384 (.014)[2]	.335 (.013)	.018 (.0013)	.264 (.024)	.227 (.018)
80, 1-85, 3	.324 (.002)	.428 (.003)	.029 (.0004)	.219 (.004)	.176 (.008)

[1]The table provides time averages and standard deviations for the indicated ratios over two sample periods. For variable definitions, see the discussion in the text.

[2]Numbers in parentheses are the sample standard deviation.

Table 3.1

The Discrete Stochastic Labor Requirement (DSLR) Model

	f	a	d	V_d^*
Point Estimate**	.915 (0.24)	.215 (.085)	.066 (.035)	$\begin{bmatrix} 656.25 & -495.49 \\ (156.99) & (135.81) \\ -495.49 & 442.58 \\ & (118.17) \end{bmatrix}$

$$L_T = -846.44 \qquad J_T^+ = \underset{(.010)}{32.14} \qquad J_T^+ = \underset{(.018)}{29.97}$$

Unconstrained VAR(2) Representation for Q_t^{++}

$$Q_t = \begin{bmatrix} \underset{(.084)}{1.070} & \underset{(.132)}{-.136} \\ \underset{(.055)}{-.074} & \underset{(.085)}{.313} \end{bmatrix} Q_{t-1} + \begin{bmatrix} \underset{(.084)}{-.204} & \underset{(.130)}{-.206} \\ \underset{(.054)}{.065} & \underset{(.085)}{.125} \end{bmatrix} Q_{t-2} + X_t^{+++}$$

$$EX_tX_t' = \begin{bmatrix} 217.02 & -.867 \\ -.867 & 92.74 \end{bmatrix}.$$

Constrained VAR (Truncated) Implied by the DSLR Model

$$Q_t = \begin{bmatrix} 1.166 & -.189 \\ -.120 & .026 \end{bmatrix} Q_{t-1} + \begin{bmatrix} -.201 & 0 \\ .034 & 0 \end{bmatrix} Q_{t-2} + X_t$$

$$EX_tX_t' = \begin{bmatrix} 261.14 & -18.64 \\ -18.64 & 107.85 \end{bmatrix}.$$

---

*V_d is the innovation in the SARMA representation implied by the DSLR model.

**Standard errors in parentheses.

+J_T is defined in footnote 3.5. Significance level of J_T and J_T in parentheses.

++Q_t is demeaned $[\phi^{-t}(c_t-\tilde{y}_t), \phi^{-t}\Delta c_t]$.

+++X_t is the disturbance term in the VAR.

Table 5.1:
First Order Autocorrelations of
Detrended Consumption First Differences[1]

Sample Period	c_{nd}	$c_{nd} + c_s$	c
51,3-85,3	.260 (.070)	.237 (.072)	.256 (.078)
52,3-85,3	.276 (.072)	.269 (.065)	.276 (.087)
51,3-79,1	.250 (.082)	.201 (.080)	.269 (.086)

[1]This table reports estimates of the first order autocorrelation of the detrended first difference of consumption under the maintained hypothesis that higher order autocorrelations are zero. The standard errors appear in parentheses. Column headings indicate the measure of consumption used. For variable definitions, see section 3.A.

Table 5.2a:

Detrended First Difference of Consumption[1] on Lagged Values of Itself and Lagged Detrended Income

Sample Period	Lags 1 - 4			Lags 2 - 4		
	$c, \tilde{y}$	$c_{nd} + c_s, y_d$	c_{nd}, y_d	$c, \tilde{y}$	$c_{nd} + c_s, y_d$	c_{nd}, y_d
51, 3-85, 3	.001	.002	.008	.029	.111	.013
52, 3-85, 3	.000	.001	.009	.029	.157	.007
51, 3-79, 1	.013	.007	.066	.159	.142	.105

[1]Significance levels of tests of null hypothesis that coefficients on lagged detrended first difference of consumption and lagged detrended income are zero. Columns labelled "Lags 1 - 4" refer to tests that allow nonzero coefficients on lags 1 - 4 of the explanatory variables under the alternative hypothesis (H_A). Columns labelled "Lags 2 - 4" refer to tests that allow only lags 2, 3, 4 to have nonzero coefficients under H_A. Column headings indicate the measure of consumption and income used in the calculations. These variable definitions are given in section 3.A.

Table 5.2b:

Detrended First Difference of Consumption[1] on Lagged Values of Itself and Lagged Detrended Consumption Minus Income

Sample Period	Lags 1 - 4			Lags 2 - 4		
	$c, \tilde{y}$	$c_{nd} + c_s y_d$	c_{nd}, y_d	$c, \tilde{y}$	$c_{nd} + c_s, y_d$	c_{nd}, y_d
51, 3-85, 3	.002	.003	.013	.072	.141	.017
52, 3-85, 3	.000	.001	.015	.086	.220	.056
51, 3-79, 1	.018	.019	.084	.240	.149	.080

[1]For an explanation, see Table 5.2a. The only difference between this table and the latter one is that here detrended income is replaced by detrended consumption minus income as an explanatory variable.

Table 5.3

GMM Test of Continuous Time Random Walk Hypothesis for Consumption[1]

Sample Period	Lagged Detrended Income			Lagged Detrended Consumption Minus Income		
	$c,\tilde{y}$	$c_{nd} + c_s, y_d$	c_{nd}, y_d	$c,\tilde{y}$	$c_{nd} + c_s, y_d$	c_{nd}, y_d
51, 3-85, 3	.008	.170	.024	.028	.210	.029
52, 3-85, 3	.041	.187	.012	.118	.270	.091
51, 3-79, 1	.040	.174	.135	.242	.180	.090

[1]Significance level of test of joint hypothesis that regression on explanatory variables lagged 2, 3, and 4 periods are zero and that the first order autocorrelation of detrended consumption first differences are zero. The first set of three columns pertains to the case when the explanatory variables are detrended first difference of consumption and detrended income. The second set of three columns pertain to the case where the explanatory variables are detrended consumption first differences and detrended consumption minus income.

Table 5.4

The Continuous Time Random Walk Model		The Continuous Time Stochastic Labor Requirement Model	
Parameter	Point Estimate*	Parameter	Point Estimate
a_1	.152 (.060)	f	.089 (.035)
a_2	27.57 (81.81)	a	11.75 (11.29)
d	.0032 (.058)	d	.058 (.042)
Estimated Covariance Matrix ** V_c		Estimated Covariance Matrix** $\mathbf{V}_c$	
$\begin{bmatrix} 488.51 & -521.63 \\ -521.63 & 719.52 \end{bmatrix}$		$\begin{bmatrix} 659.82 & 483.60 \\ 483.60 & 465.28 \end{bmatrix}$	
$\mathcal{L}_T = -843.249$		$\mathcal{L}_T = -842.982$	
J_T** = 25.76 (.058)		J_T** = 25.22 (.066)	
$\mathbf{J}_T$*** = 24.02 (.089)		$\mathbf{J}_T$*** = 23.52 (.100)	

*Standard errors in parentheses.

**V_c is the covariance matrix of the vector $[\eta_1^*(t)/(a_1+\delta), \eta_2^*(t)/(a_2+\delta)]$. $\mathbf{V}_c$ is the covariance matrix of the vector $[\eta^*(t)/(a+\delta), \delta\varepsilon^*(t)/(f+\delta)]$.

***Significance level of J_T and $\mathbf{J}_T$ in parentheses.

Table 5.5

Comparison of VAR(2) Representations for Q_t*

Truncated VAR Implied by the Continuous Time Random Walk Model

$$Q_t = \begin{bmatrix} 1.126 & -.033 \\ .011 & .270 \end{bmatrix} Q_{t-1} + \begin{bmatrix} -.300 & .018 \\ -.016 & -.073 \end{bmatrix} Q_{t-2} + \begin{bmatrix} .080 & -.007 \\ .008 & .019 \end{bmatrix} Q_{t-3}$$

$$+ \begin{bmatrix} -.021 & .003 \\ -.003 & -.005 \end{bmatrix} Q_{t-4} + X_{ct}^{**}$$

$$EX_{ct}X_{ct}' = \begin{bmatrix} 255.87 & -17.93 \\ -17.93 & 101.52 \end{bmatrix}.$$

Truncated VAR Implied by the Continuous Time Stochastic Labor Requirement Model

$$Q_t = \begin{bmatrix} 1.241 & -.072 \\ -.078 & .277 \end{bmatrix} Q_{t-1} + \begin{bmatrix} -.343 & .038 \\ -.010 & -.075 \end{bmatrix} Q_{t-2} + \begin{bmatrix} .096 & -.016 \\ .007 & .020 \end{bmatrix} Q_{t-3}$$

$$+ \begin{bmatrix} -.024 & .006 \\ -.001 & -.005 \end{bmatrix} Q_{t-4}$$

$$EX_{ct}X_{ct}' = \begin{bmatrix} 268.25 & -22.90 \\ -22.90 & 100.13 \end{bmatrix}.$$

*Q_t is demeaned $\left[\phi^{-t}(c_t-\tilde{y}_t), \phi^{-t}\Delta c_t\right]$.

**$X_{ct} = Q_t - E\left[Q_t \middle| Q_{t-s}; s = 1,2,\ldots \text{ under the null hypothesis of the CRW model}\right]$.

$X_{ct} = Q_t - E\left[Q_t \middle| Q_{t-s}; s = 1,2,\ldots \text{ under the null hypothesis of the CSLR model}\right]$.

Appendix A

Derivation of (2.10) and (2.11)

The social planner chooses a contingency plan for capital to maximize (2.7) subject to (2.5). The resulting stochastic Euler equation is:

$$E_t\{[1-\delta\beta L^{-1}][1-\delta L]k_t\} = E_t\{[1-\delta\beta L^{-1}][e_t-b_t]-H_t\}$$

or, since $\delta\beta = 1$,

$$\text{(A.1)} \qquad E_t\{[1-L^{-1}][1-\delta L]k_t\} = E_t\{[1-L^{-1}][e_t-b_t]-H_t\}.$$

Note that we can rewrite the characteristic polynomial of (B.1) as

$$[1-Z^{-1}][1-\delta Z] = [\delta-Z^{-1}][1-Z].$$

The condition $\delta\beta = 1$ implies that constraint (2.5) is binding (see Hansen [1986]), so (A.1) can be solved by applying the forward operator $[\delta-L^{-1}]^{-1}$ to both sides of the equation, yielding:

$$\begin{aligned} k_t - k_{t-1} &= E_t\{[\delta-L^{-1}]^{-1}[1-L^{-1}](e_t-b_t)-[\delta-L^{-1}]^{-1}H_t\} \\ &= E_t \frac{\beta}{1-\beta L^{-1}} \{(e_t-b_t)-(e_{t+1}-b_{t+1})-H_t\} \\ &= \beta E_t\{\sum_{j=0}^{\infty}\beta^j(e_{t+j}-b_{t+j})-\sum_{j=1}^{\infty}\beta^{j-1}(e_{t+j}-b_{t+j})-\sum_{j=0}^{\infty}\beta^j H_{t+j}\}. \end{aligned}$$

Rearranging terms,

$$\text{(A.2)} \qquad k_t - k_{t-1} = (\beta-1)E_t\sum_{j=0}^{\infty}\beta^j(e_{t+j}-b_{t+j}) + e_t - b_t - \beta E_t\sum_{j=0}^{\infty}\beta^j H_{t+j}.$$

Rewriting (A.2) using the notation defined in (2.8) results in (2.10). Equation (2.11) is obtained by using (2.10) and the fact that $c_t = \delta k_{t-1} - k_t + e_t$.

Appendix B

Derivation of Decision Rules for the Continuous Time Model

This appendix provides an informal derivation of the decision rules, (4.9) - (4.10) for the continuous time planning problem, (4.7). Proceeding as in Hansen and Sargent [1980] we can show that the Euler equation for the social planner's problem is

(B.1) $$D(D-\delta)k(t) = D\left[e(t)-b(t)\right] + H(t).$$

The unique solution to this problem which satisfies (4.5) is

(B.2) $$Dk(t) = e(t) - b(t) - \delta E_t \int_0^\infty e^{-\delta\tau}\{e(t+\tau)-b(t+\tau)\}d\tau - \delta E_t \int_0^\infty e^{-\delta\tau}H(t+\tau)d\tau.$$

This is easily shown to equal the first equation in (4.9) after the definition (4.8) is taken into account. The second equation in (4.9) is obtained by substituting the first into the relationship $c(t) = \delta k(t) - Dk(t) + e(t)$. Equation (4.10) is just the sum of the two equations in (4.9).

To derive (4.11), we first present some preliminary results regarding $x_p(t)$. Suppose the fundamental representation for $x(t)$ is $x(t) = C(D)\varepsilon(t)$. Here, $C(s) = 0$ implies $\mathrm{Real}(s) \leq 0$ and the poles of $C(s)$ lie in the closure of the left side of the complex plane (see Sargent [1982] for a discussion of the link between these conditions and $\varepsilon(t)$ being fundamental for $x(t)$). Then,

(B.3) $$x_p(t) = -\delta E_t \frac{1}{D-\delta} x(t) = -\delta E_t \frac{C(D)}{D-\delta}\varepsilon(t) = -\delta \frac{C(D) - C(\delta)}{D-\delta}\varepsilon(t),$$

by a formula due to Hansen and Sargent [1980]. Multiply both sides of (C.3) by D - δ and rearrange, to obtain

(B.4) $$(D-\delta)x_p(t) + \delta x(t) = \delta C(\delta)\varepsilon(t).$$

We identify $u_{x_p}(t)$ with $\delta C(\delta)\varepsilon(t)$. To see why, first write,

(B.5) $$x(t) = C(D)\varepsilon(t) = \int_0^\infty c(\tau)\varepsilon(t-\tau)d\tau,$$

where the function $c(\tau)$, $\tau \geq 0$ is uniquely defined by

(B.6) $$C(s) = \int_0^\infty c(\tau)e^{-s\tau}d\tau, \ \mathrm{Real}(s) > 0.$$

Equation (B.5) defines c as the impulse response function of $x(t)$ to $\varepsilon(t)$.

Consider the effect of a disturbance in $x(t)$ that is uncorrelated with $I(t-\tau)$, $\tau > 0$. This arises from a pulse in $\varepsilon(t)$, which leads to a revision in the forecast of $x(t+\tau)$ in the amount $c(\tau)\varepsilon(t)$ $\tau \geq 0$. The permanent value of this revision is $\delta C(\delta)\varepsilon(t)$. Thus the effect of the pulse in $\varepsilon(t)$ is to disturb $x_p(t)$ by $\delta C(\delta)\varepsilon(t)$, which is why we identify $\delta C(\delta)\varepsilon(t)$ with $\mu_{x_p}(t)$. We conclude that

(B.7) $$(D-\delta)x_p(t) + \delta x(t) = \mu_{x_p}(t).$$

From (4.10),

(B.8) $$Dc(t) = De_p(t) + Db(t) - Db_p(t) + \delta Dk(t) + DH_p(t)/\delta.$$

Equation (4.11) is obtained by first substituting the first equation in (4.9) into (B.8) and then making use of (B.7).

References

Bergstrom, A. R., (1983), "Gaussian Estimation of Structural Parameters in Higher Order Continuous Time Dynamic Models," Econometrica 51, no. 1, 117-52.

Brayton, F., and Mauskopf, E, (1985), "The MPS Model of the United States Economy," Board of Governors of the Federal Reserve System, Division of Research and Statistics, Washington, D.C., February.

Christiano, Lawrence J., (1984), "A Critique of Conventional Treatments of the Model Timing Interval in Applied Econometrics," manuscript.

Christiano, Lawrence J., (1985), "A Method of Estimating the Timing Interval in a Linear Econometric Model, With an Application to Taylor's Model of Staggered Contracts," Journal of Economic Dynamics and Control 9, pp. 363-404.

Christiano, L., Eichenbaum, M. and Marshall, D., (1987), "The Permanent Income Hypothesis Revisited Technical Appendix," manuscript.

Christiano, L. and Marshall, D., (1986), "A Note on the Estimation of Continuous Time Models from Temporally Averaged Data," manuscript.

Dunn, K. B. and Singleton, K. J., (1986), "Modeling the Term Structure of Interest Rates Under Nonseparable Utility Functions and Durability of Goods," Journal of Financial Economics Vol. 17; 27-55.

Durbin, J., (1961), "Efficient Fitting of Linear Models for Continuous Stationary Time Series from Discrete Data," Bulletin of the International Institute of Statistics, 38: 273-82.

Eichenbaum, M. and Hansen, L. P., (1986), "Estimating Models with Intertemporal Substitution Using Aggregate Time Series Data," manuscript.

Flavin, M. A., (1981), "The Adjustment of Consumption to Changing Expectations about Future Income," Journal of Political Economy, 89: 974-1009.

Geweke, J., (1978), "Temporal Aggregation in the Multivariate Regression Model," Econometrica, 46: 643-62.

Hall, R. E., (1978), "Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence," Journal of Political Economy, 86: 971-87.

Hall, R. E. and Mishkin, F. S., (1982), "The Sensitivity of Consumption to Transitory Income: Estimates from Panel Data on Households," Econometrica 50: 461-81.

Hannan, E. J., (1970), Multiple Time Series, New York: Wiley.

Hansen, Lars P., (1986), "Econometric Modeling of Asset Pricing Under Rational Expectations," manuscript.

Hansen, Lars P., J. Heaton, and M. Ogaki, (1987), "Efficiency Bounds Implied By Multiperiod Conditional Moment Restrictions," manuscript, University of Chicago.

Hansen, Lars P., Roberds, William and Sargent, Thomas J., (1987), "Time Series Implications of Present-Value Budget Balance and of Martingale Models of Consumption or Taxes," manuscript, September.

Hansen, L. P. and Sargent, T. J., (1980), "Methods for Estimating Continuous Time Rational Expectations Models From Discrete Time Data," Federal Reserve Bank of Minneapolis, Staff Report No. 59.

Hansen, L. P. and Sargent, T. J., (1980a), "Formulating and Estimating Dynamic Linear Rational Expectations Models," Journal of Economic Dynamics and Control, vol. 2.

Hansen, L. P. and Sargent, T. J., (1981), "Formulating and Estimating Continuous Time Rational Expectations Models from Discrete Data," manuscript.

Hansen, L. P. and Sargent, T. J., (1981a), "Exact Linear Rational Expectations Models: Specification and Estimation," Federal Reserve Bank of Minneapolis Staff Report 71, September.

Hansen, L. P. and Singleton, K. J., (1982), "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models," Econometrica, 50: 1029-54.

Harvey, A. C. and Stock, James H., (1986), "Estimating Integrated Higher Order Continuous Time Autoregressions," manuscript, June.

Hayashi, F., (1982), "The Permanent Income Hypothesis: Estimation and Testing by Instrumental Variables," Journal of Political Economy, 90: 895-916.

Kydland, F. E. and Prescott, E. C., (1982), "Time to Build and Aggregate Fluctuations," Econometrica, 51: 1821-37.

Lissitz, R. W., (1972), "Comparison of the Small Sample Power of the Chi-Square and Likelihood Ratio Tests on the Assumptions for Stochastic Models," Journal of the American Statistical Association, September, Vol. 67; 574-577.

Long, J. and Plosser, C., (1983), "Real Business Cycles," Journal of Political Economy, 91: 39-69.

Lucas, R. E., Jr., (1985), "Models of Business Cycles," manuscript.

Marcet, A., (1986), "Temporal Aggregation of Economic Time Series," manuscript.

Michener, Rou, (1984), "Permanent Income in General Equilibrium," Journal of Monetary Economics 13, 297-305.

Musgrave, J., (1980), "Government Owned Fixed Capital in the United States," Survey of Current Business, March: 33-43.

Phillips, A. W., (1958), "The Estimation of Parameters in Systems of Stochastic Differential Equations," Biometrika, 59: 657-65.

Rozanov, Y. A., (1967), Stationary Random Processes, Holden-Day.

Sargent, T. J., (1986), "Equilibrium Investment Under Uncertainty: Measurement Errors and the Investment Accelerator," University of Minnesota.

Sims, C. A., (1971), "Discrete Time Approximations to Continuous Time Lag Distributions in Econometrics," Econometrica, 39: 545-564.

Sims, C. A., (1980), "Macroeconomics and Reality," Econometrica, 48: 1-48.

Whittle, P., (1953), "The Analysis of Multiple Stationary Time Series," Journal of the Royal Statistical Society, series B, 15: 125-39.

Working, H., (1960), "Note on the Correlation of First Differences of Averages in a Random Chain," Econometrica, 28: 916-918.

Zadrozny, P., (1984), "The Exact Gaussian Likelihood of Continuous-Time Autoregressions for Stock and Flow Data at Different Frequencies," manuscript, Center for Economic Studies, U.S. Bureau of the Census, December.

Zeldes, S., (1985), "Consumption and Liquidity Constraints: An Empirical Investigation," Rodney L. White Center for Financial Research Paper No. 24-85.