Rational Expectations, Econometric Exogeneity, and Consumption

Thomas J. Sargent
University of Minnesota and Federal Reserve Bank of Minneapolis

Estimates of a rational expectations version of Friedman’s time-series consumption model are obtained by imposing the pertinent restrictions across the stochastic processes for consumption and income. A likelihood ratio test is used to test the adequacy of three joint hypotheses: namely, Friedman’s model, rational expectations, and some arbitrary conditions on the disturbance process in the consumption function. The paper treats both the cases in which income is econometrically exogenous with respect to consumption and those in which it is not. The macroeconomics of this exogeneity condition are briefly discussed.

This paper computes estimates of the parameters of Friedman’s (1957) permanent-income model of the consumption function. The estimates are obtained from aggregate time-series data for the United States by imposing the restrictions across the consumption and income processes that are implied by the hypothesis of rational expectations in the context of the permanent-income theory of consumption. That the hypothesis of rational expectations is an important element of Friedman’s application of his theory, especially to the cross section data, has recently been emphasized by Lucas (1976).

To implement the rational expectations theory in the present context requires characterizing the stochastic structure of the consumption-income process from the point of view of extracting optimal forecasts of future income from records of past income and consumption. For us the important aspects of this characterization happen to coincide with the Wiener–Granger notion of “causality.” According to Wiener and Granger’s definition (in Granger 1969), a process $C$ is said to cause a process $Y$ if, given

Research on this paper was supported by the Federal Reserve Banks of Minneapolis and Chicago, neither of which is responsible for the conclusions. Robert Litterman ably performed the calculations. Helpful comments on an earlier draft were received from Robert E. Lucas, Jr., Michael Mussa, and Christopher A. Sims.
past values of $Y$, past values of $C$ help to predict future $Y$. A theorem of Sims (1972) informs us that, if $C$ fails to Granger-cause $Y$, then there exists a model consistent with the data in which $Y$ is strictly econometrically exogenous with respect to $C$. This is of interest here because, as is well known, in standard Keynesian stochastic macroeconomic models, income $Y$ is predicted not to be exogenous with respect to consumption $C$. Despite this fact, this paper devotes some space to describing an estimator that is appropriate in the case in which $Y$ is strictly econometrically exogenous with respect to $C$. By way of indicating that the hypothesis of exogeneity of $Y$ with respect to $C$ is not uniformly rejected by macroeconomic theory, Section I describes a version of Tobin's dynamic aggregative model (1955) which predicts that $Y$ is exogenous with respect to $C$. In Tobin's model, movements in government purchases leave GNP unaltered in the short run, while movements in the money supply cause sympathetic movements in real GNP, in sharp contrast to the Keynesian model. The differing implications of Tobin's model and the Keynesian model for the Granger-causal structure of income and consumption indicate that the econometric exogeneity tests reported in Section VI are of independent macroeconomic interest, over and above their implications about appropriate procedures for estimating the consumption function under rational expectations. Section II digresses somewhat from the main theme of this paper by indicating how in Tobin's model there may emerge a spurious investment accelerator, this despite the fact that in Tobin's model firms have no investment demand schedules.

Sections III and IV describe estimation procedures for the cases in which income is and is not exogenous, respectively, while Section V takes a stand on the question of whether seasonally adjusted or unadjusted data are the appropriate ones to use. Exogeneity tests are reported in Section VI, while the rational expectations restrictions are imposed and tested in Section VII.

This paper stays with Friedman's (1957) original very simple specification of the consumption function, supplemented only by the imposition of Muth's hypothesis (1961) that expectations of future income are "rational." Proceeding in this way is subject to the valid criticism that it ignores the insights into the form of optimal consumption rules provided by Merton (1970). Further, from the viewpoint of equilibrium theories of the business cycle, which envision agents as confronting not incomes but, rather, sequences of wages and prices at which they can supply labor and consume (as in Lucas and Rapping 1969), the consumption schedule estimated in this paper is misspecified.¹ Both of these criticisms ought to be taken seriously.

¹ Similarly, Hall's (1977) proposition that consumption is an uncaused (in Granger's 1969 sense) first-order Markov process will not in general hold in Lucas and Rapping's framework. To see this in Hall's context, make instantaneous utility depend on current
Let me warn the reader in advance that the econometric findings are not encouraging from the viewpoint of the permanent-income model of consumption. But finding confirmation of that theory ought not to be the only reason for reading this paper. The paper describes a useful technology for imposing rational expectations on time-series models, a technology which is applicable to a number of dynamic econometric models of rational agents. The paper illustrates the feasibility of that technology and also the strongly overidentifying nature of the restrictions embodied in rational expectations models.

The reader not interested in the macroeconomic issues associated with the question of the econometric exogeneity of consumption can without loss advance to Section III and find a discussion of the estimation technology.

I. Econometric Exogeneity of Income in the Consumption Function

Most macroeconomists would find unacceptable the specification that income is strictly econometrically exogenous in the consumption function. Indeed, ever since Haavelmo's famous paper (1943), the failure of the consumption function to be a regression equation in the context of a Keynesian model has been a classic example of simultaneous equations bias. It perhaps bears pointing out, however, that there exist interesting macroeconomic models in which $Y$ is econometrically exogenous in the consumption function. One such model is illustrated here, though it is certainly not the only possible example of such a model.\(^2\)

Consider the following stochastic version of Tobin's dynamic aggregative model (1955):

\[
\begin{align*}
\log Y_t &= \sigma \log N_t + \varepsilon_{t1}, \sigma \approx 1 \quad \text{production function (a)} \\
\log w_t - \log p_t &= g \cdot (\log N_t - \log K_t) + \varepsilon_{2t}, \quad g < 0 \quad \text{aggregate demand schedule for employment (b)} \\
r_t &= h(\log N_t - \log K_t) + \varepsilon_{3t}, \quad h > 0 \quad \text{marginal productivity for capital (c)} \\
\log M_t - \log p_t - \log Y_t &= br_t + \varepsilon_{4t}, b < 0 \quad \text{portfolio-balance schedule (d)}
\end{align*}
\]

consumption and employment, confront the household with an exogenous real-wage process à la Lucas and Rapping, and then calculate the household's Euler equations for employment and consumption.

\(^2\) Eq. (e) could be replaced by a Phillips curve without affecting the properties of interest here. I use the simple specification (e) here in order to emphasize that the exogeneity of $Y_t$ with respect to $C_t$ in this model is not a result of imposing classical labor market assumptions.
\[
\log w_t = \sum_{i=1}^{n} v_i \log w_{t-i} + \varepsilon_{5t}
\]

exogenous money wage process \hfill (e)

\[
C_t = BY_{pt} + u_t, \quad B > 0
\]
cconsumption function \hfill (f)

\[
Y_{pt} = (1 - \alpha) \sum_{j=0}^{\infty} \alpha^j E_t Y_{t+j}, \quad 0 < \alpha < 1
\]
definition of permanent income \hfill (g)

\[
C_t + I_t + G_t = Y_t
\]
GNP identity \hfill (h)

\[
K_{t+1} - K_t = I_t.
\]

(i)

Here \( Y \) is GNP, \( N \) is employment, \( r \) the interest rate, \( G_t \) government purchases, \( p_t \) the price level, \( w_t \) the money wage, \( M_t \) the money supply, and \( I_t \) investment; \( \varepsilon_{1t}, \varepsilon_{2t}, \varepsilon_{3t}, \varepsilon_{4t}, \varepsilon_{5t}, \) and \( u_t \) are stationary random processes with finite means and variances. I assume that the \( u \) process is statistically independent of the \( \varepsilon_i \)'s, \( i = 1, 2, 3, 4, 5 \), and \( \varepsilon_4 \) processes, though these latter five processes can be statistically dependent. The notation \( E_t Y_{t+j} \) denotes the linear least-squares forecast of \( Y_{t+j} \) based on information available at time \( t \).

Tobin's model is of substantial interest to macroeconomists partly because it is a concrete example of a model in which "pure" fiscal policy (e.g., a change in government expenditures or a lump-sum transfer) does not affect output in the short run, while monetary policy does. The present version of the model is nine equations in the nine endogenous stochastic processes \( Y_t, N_t, r_t, p_t, w_t, C_t, Y_{pt}, I_t, \) and \( K_t \), with exogenous driving processes \( M_t, G_t, \) and \( \varepsilon_i \)'s. As it happens, equations (a)-(e) form a recursive block that each period determines \( Y_t, r_t, p_t, w_t, \) and \( N_t \), given values of the exogenous \( M_t \) and the predetermined \( K_t \) (predetermined in a fashion to be explained shortly). Given \( Y_t \) so determined, equations (f) and (g) determine \( C_t \), while (h) determines \( I_t \), which through equation (i) determines \( K_{t+1} \).

To indicate in more detail how (a)-(e) determine \( Y, r, p, w, \) and \( N \), substitute (a) and (b) into (d) to get the modified "LM" curve

\[
br_t = \log M_t - \log w_t + (g - \sigma) \log N_t - g \log K_t + \varepsilon_{2t} - \varepsilon_{1t} - \varepsilon_{4t},
\]

(LM)

which, given \( M_t, w_t, \varepsilon_{2t}, \) and \( \varepsilon_{4t} \), depicts the \((r_t, N_t)\) locus that assures portfolio balance, assuming that (a) and (b) are satisfied. Equation (c) is a second curve in the \((r, N)\) plane, one along which the demand for the existing stock of capital equals the supply. Both (c) and the (LM) curve are upward-sloping curves in the \((r, N)\) plane. Stability\(^3\) requires that the (LM) curve be the steeper, which is the condition that \((g - \sigma)/b > h\).

Substituting (LM) into (c) and solving for \( \log N_t \) gives the reduced form

\(^3\) See Henderson and Sargent (1973) or Sargent (1975).
\[ b[h - (g - \sigma)b]\log N_t = \log M_t - \log w_t + (bh - g)\log K_t + (\varepsilon_{2t} - \varepsilon_{1t} - \varepsilon_{4t} - b\varepsilon_{3t}) \]

The assumption of stability guarantees that the coefficient on \( \log N_t \) is positive, so that employment varies directly with \( M_t \), inversely with \( w_t \). Because the sign of \( (bh - g) \) is ambiguous, so is the effect of capital on employment and also, given (a), on output. Increases in \( \log K_t \) have two contrary effects on employment. On one hand, an increase in \( K_t \), given \( N_t \), calls for a higher real wage via (b) and, given the \( w_t \) fixed by (e), a lower \( p_t \). This lowers the portfolio-balancing interest rate and is expansionary (see eq. [LM]). On the other hand, an increase in \( K_t \) lowers the marginal product of capital given \( N_t \) and therefore lowers the capital-market clearing interest rate, which is contractionary (see eq. [c]). The two effects are opposing.

In this model, \( Y_{pt} \) will be strictly econometrically exogenous in the consumption schedule under two conditions. First, the \( u_t \) process must, as I have assumed, be statistically independent of the \( e \) processes. Second, we must have the condition \( (bh - g) = 0 \), so that capital does not appear in the reduced form for employment, the two opposing effects of the preceding paragraph just offsetting one another. With these two conditions met, (a)–(i) is an example of a model in which \( Y_{pt} \) is strictly econometrically exogenous in the consumption function.

It is worth remarking that, were we to add the side condition that \( u_t \) be serially independent, then even without our second condition the consumption function (f) is a regression equation; that is, it satisfies the population least-squares orthogonality condition \( E u_t Y_{ps} = 0 \) for all \( t \geq s \). The force of our second condition is to guarantee \( E u_t Y_{ps} = 0 \) for \( s > t \). Under the side condition that \( u \) be serially uncorrelated, estimation of the consumption function (f) by variants of least-squares methods (methods modified to extract proxies for the unobservable \( Y_{pt} \)) would yield consistent estimates even without our second condition.

Obviously, a priori setting down of models like the one above (or like

4 That \( G_t \) does not appear in the reduced form for \( N_t \) establishes the lack of a short-run effect of \( G_t \) on \( N_t \) or \( Y_t \) (for "short-run" read "given \( K_t \)"). It is possible for the \( G \) process to affect subsequent \( Y_t \)'s through the effects of \( G_t \) on \( I_t \) and therefore subsequent values of \( K \). This is the usual "classical" route by which fiscal policy affects the rate of growth of output, though not its level in the "short run."

5 The reader may question the exclusion of capital from the production function (a). However, see the time-series regressions reported by Bodkin and Klein (1967) and Lucas (1970). The exclusion of capital from (a) together with its inclusion in the aggregate demand schedule for labor (b) is rationalized theoretically by Lucas (1970).

6 The reader is invited to work things out where (c) is replaced by \( r_t = h(\log N_t - \log K_t) + E_t(\log p_{t+1} - \log p_t) + \varepsilon_{3t} \), where \( E_t(x) \) is again the linear least-squares forecast of the random variable \( x \). With this modification the two conditions in the text continue to imply that \( Y_{pt} \) is strictly exogenous in the consumption function. One of the virtues of Tobin's model is that it readily delivers procyclical movements of the interest rate even if variance in \( M_t \) is the prime force behind the cycle and even if \( E_t(p_{t+1} - p_t) \) is omitted from the right side of (c). Compare this with the predictions of standard Keynesian models with downward-sloping "IS" curves.
the standard Keynesian model) cannot settle the issue of whether \( Y_p t \) is properly regarded as econometrically exogenous in the consumption function. All that the model in this section is intended to suggest is that testing for econometric exogeneity in the consumption-income process is an interesting enterprise from the perspective of alternative macroeconomic models.

II. A Spurious Accelerator in Tobin’s Dynamic Aggregative Model

The preceding model is one in which a spurious investment accelerator emerges in the data, despite the fact that the model is one in which firms have no investment demand schedules. To illustrate things simply, I assume the special version of the model in the preceding section,

\[
C_t = \frac{B}{1 - \lambda L} Y_t + u_t, \quad 0 < \lambda < 1, \quad (k)
\]

\[
Y_t = C_t + I_t, \quad (m)
\]

where \( \mathbb{E} u_t \cdot Y_s = 0 \) for all \( t \) and \( s \), so that \( Y \) is strictly econometrically exogenous in \( (k) \). Assume that \( Y \) and \( u \) are covariance stationary stochastic processes. Now solve \( (k) \) and \( (m) \) for \( I_t \) to get

\[
I_t = \frac{(1 - B) - \lambda L}{1 - \lambda L} Y_t - u_t. \quad (n)
\]

Equation \( (n) \) expresses \( I \) as a one-sided distributed lag of \( Y \) with a disturbance \( -u_t \) that is orthogonal to the entire \( Y \) process. That is, \( Y \) is strictly econometrically exogenous in \( (k) \). If \( (1 - B) \approx \lambda \), then \( (n) \) approximately assumes the form of the distributed-lag accelerator. To see this more generally, the model \( (m) \) implies that the spectrum of \( Y \) is related to the cross-spectra of \( Y \) with \( C \) and \( I \), respectively, by \( S_Y(w) = S_{CY}(w) + S_{IY}(w) \). Dividing by the spectrum of \( Y \) gives

\[
\frac{S_{IY}(w)}{S_Y(w)} = 1 - \frac{S_{CY}(w)}{S_Y(w)}. \quad (o)
\]

If the gain \( |S_{CY}(w)/S_Y(w)| \) of consumption on income has the characteristic permanent-income-consumption-function form, with gain generally falling with increases in angular frequency \( w \), it follows from \( (o) \) that the gain of investment on income will display the characteristic (distributed-lag) accelerator form with the gain of investment on income generally rising with increases in angular frequency. Indeed, as \( (o) \) shows, the gain of investment on income is in a sense the reflection of the gain of consumption on income.

It would obviously be a mistake to interpret \( (n) \) as a structural invest-
ment demand schedule reflecting firms' behavior. Rather, it is a saving schedule reflecting choices of consumers. The present illustration is in effect an extreme example of the fact that least-squares estimates of investment schedules of the form \( (n) \) in the context of most macroeconomic models typically represent a confounding of firms' investment demand with households' saving behavior.

Of course it is possible to set up examples of macroeconomic models in which the dual of the error above occurs: that is, models in which there is a true structural investment demand schedule and in which the projection of consumption on income looks like a consumption function despite its truly being a mere reflection of the investment schedule.

### III. Estimation of the Consumption Function under Rational Expectations with Income Exogenous

Suppose the consumption function is

\[
C_t = B Y_{pt} + u_t, \tag{1}
\]

\[
Y_{pt} = (1 - \alpha)[Y_t + \alpha E_t Y_{t+1} + \alpha^2 E_t Y_{t+2} + \ldots],
\]

\[
= (1 - \alpha) \sum_{j=0}^{\infty} \alpha^j E_t Y_{t+j}, \quad 0 < \alpha < 1, \tag{2}
\]

where \( \alpha \) is a discount factor, \( E_t Y_{t+j} \) is the linear least-squares forecast of \( Y_{t+j} \) conditional on information assumed to be known at time \( t \), and \( u_t \) is a stationary random process that obeys \( E_t Y_s = 0 \) for all \( t \) and \( s \), which is equivalent to the condition that \( Y_s \) is strictly econometrically exogenous in \( (1) \). Subsequently, I will assume that information at time \( t \) consists of \( C_t, C_{t-1}, \ldots, Y_t, Y_{t-1}, \ldots \), so that \( E_t x = E x | C_t, C_{t-1}, \ldots, Y_t, Y_{t-1}, \ldots \), where \( E \) is the linear least-squares projection operator. The assumption that \( Y_t \) is econometrically exogenous in \( (1) \) implies that, given lagged values of \( Y_t \), lagged values of \( C \) do not help predict \( Y_t \). Therefore, as far as concerns the covariance structure of the \( (C, Y) \) process, it involves no further restrictions to assume that \( E_t Y_{t+j} \) is simply the regression (projection) of \( Y_{t+j} \) on current and past \( Y_t \)'s, since even if current and lagged \( C \)'s were included in the conditioning set, they would bear zero coefficients.

7 Actually the results below are valid as long as the information set that the public uses to forecast includes at least \( C_t, C_{t-1}, \ldots, Y_t, Y_{t-1}, \ldots \).

8 This is an implication of the two theorems proved by Sims (1972).

9 As will be noted in the text below, the estimator proposed in this section is statistically consistent under weaker conditions than those that have been imposed on the \( u_t \)'s in this paragraph.
By the linearity of the projection operator $E_t$, (2) can be rearranged to read\(^{10}\)

$$Y_{pt} = (1 - \alpha)E_t \sum_{j=0}^{\infty} \alpha^j Y_{t+j}. \quad (3)$$

Suppose that $Y_t$ is a linearly regular stationary stochastic process with covariance-generating function

$$R_y(z) = \sigma^2 b(z)b(z^{-1}), \quad (4)$$

where $b(z) = \sum_{j=0}^{\infty} b_j z^j$, $\sigma^2 > 0$. Any linearly regular (strictly linearly indeterministic) stochastic process has a covariance-generating function that can be represented in the form (4). By a famous formula of Wiener and Kolmogorov,\(^{11}\) the $z$-transform of the coefficients in the projection of $Y_{t+j}$ against $Y_t, Y_{t-1}, \ldots$ is given by $(1/b(z))[z^{-1}b(z)]_+$, where $[ \quad ]_+$ means ignore all negative powers of $z$, that is,

$$\left[ \sum_{j=-\infty}^{\infty} h_j z^j \right]_+ \equiv \sum_{j=0}^{\infty} h_j z^j.$$

\(^{10}\) Hall (1977) has recently suggested that the permanent-income hypothesis implies that aggregate consumption follows an uncaused (in Granger's sense) first-order Markov process. Within the context of the consumption schedule (1), sufficient conditions for Hall's proposition are (i) that theoretically permanent income follows an uncaused first-order Markov process and (ii) that consumption is an exact linear function of permanent income alone, $u_t$ being identically zero. In general, under the definition (2) of permanent income in the text, permanent income is not predicted to be an uncaused first-order Markov process. (A process $X_t$ is said to be uncaused relative to $\Omega_t$ in Granger's sense if $E X_t+1 | \Omega_t$ is only a function of $X_t, X_{t-1}, \ldots$, where $\Omega_t$ is an information set including at least $X_t, X_{t-1}, \ldots$.) To see this, notice that $Y_{pt}$ given by (2) is a solution to the stochastic difference equation $Y_{pt} = (1 - \alpha)Y_t + \alpha E_t Y_{pt+1}$, $0 < \alpha < 1$. Rearranging, we have

$$E_t Y_{pt+1} = \frac{1}{\alpha} Y_{pt} - \left( 1 - \frac{1}{\alpha} \right) Y_t. \quad (*)$$

In general, permanent income $Y_{pt}$ does not follow an uncaused first-order Markov process; i.e., $E_t Y_{pt+1}$ is not a function of $Y_{pt}$ alone; but there are two possible ways of restricting things to get that special result. The first way involves driving the discount factor $\alpha$ to unity. Heuristically, (*) indicates that with $\alpha = 1$, permanent income follows a random walk (i.e., an uncaused first-order Markov process with Markov parameter unity). For this argument to be rigorous, permanent income must be redefined as a Cesaro sum,

$$Y_{pt} = \lim_{n \to \infty} \left( \frac{1}{n} \sum_{j=0}^{n} E_t Y_{t+j} \right),$$

since $Y_{pt}$ given in (2) is not well defined with $\alpha = 1$. A second way of making $E_t Y_{pt+1} = Y_{pt}$ would be first to redefine permanent income as

$$Y_{pt} = (1 - \alpha) \sum_{j=0}^{\infty} \alpha^j E_t Y_{t+j+1}. \quad (***)$$

Then follow Muth (1960) and require that $Y_t$ follow the special process $Y_t = Y_{t-1} + a_t - \lambda a_{t-1}$, $|\lambda| < 1$, where $a_t$ is the process of one-step-ahead errors in predicting $Y_t$ from its own past. Muth showed that, with this process, definition (***) implies that permanent income follows a random walk for any $|\lambda| < 1$.

\(^{11}\) See Whittle (1963, p. 32). It is necessary to assume that $Y_t$ possesses an autoregressive representation, i.e., that $b(z)$ is invertible.
Then, using (3), \( Y_{pt} \) can be written

\[ Y_{pt} = h(L) Y_t, \]  

(5)

\[ h(z) = \frac{(1 - \alpha)}{b(z)} \left[ \frac{1}{(1 - \alpha z^{-1})} b(z) \right]^+. \]  

(6)

In practice, \( b(z) \) could be assumed to be a ratio of finite-order polynomials in \( z \) so that \( Y \) would be a moving average, autoregressive process. Under such assumptions, by using the method of partial fractions, convenient closed-form expressions for \((1 - \alpha)/b(z)\) and \(\{1/(1 - \alpha z^{-1})b(z)\}_+\) in terms of the roots of \([b(z) = 0]\) can be obtained. In this way, for fixed \( \alpha \) and given \( b(z) \), a closed-form expression for \( h(z) \) can be obtained.

Rather than pursue those calculations here, I propose to exhibit the following more compact calculations. I assume temporarily that \( \{Y_t\} \) is a covariance stationary stochastic process, though the calculations below permit limited departures from stationarity. From Wold’s decomposition theorem,\(^{12}\) it follows that \( Y_t \) can be represented uniquely as the sum of two orthogonal stochastic processes \( Y_t = Y^i_t + \epsilon_t \), where \( \text{E}Y^i_t \epsilon_s = 0 \) for all \( t \) and \( s \). Here \( \epsilon_t \) is a linearly deterministic stochastic process, which can be predicted arbitrarily well, arbitrarily far into the future from knowledge of past \( \epsilon \)'s or past \( Y^i \)'s alone; \( Y^i_t \) is a linearly indeterministic process which generally cannot be predicted perfectly from its own past. Wold’s decomposition theorem states that \( Y^i_t \) always has a moving average representation \( Y^i_t = \sum_{j=0}^{\infty} d_j a_{t-j} \) where \( \sum_{j=0}^{\infty} d_j^2 < \infty \) and \( a_t = Y_t - \text{E}Y_i | Y_{t-1}, \ldots, Y_{t-2} \). A bivariate version of Wold’s theorem asserts the existence of a similar orthogonal decomposition into deterministic and indeterministic processes for bivariate jointly covariance stationary stochastic processes.

The model formed by (1) and (2) imposes restrictions across both the deterministic and the indeterministic part of the \( C_t \) and \( Y_t \) processes. In the remainder of this section, I shall focus on characterizing the restrictions imposed on the indeterministic parts only. I shall assume that the strictly deterministic parts have been subtracted off; certain of the estimates reported in Section VII utilize as data residuals from regressions on a trend and seasonal dummies in an effort to approximate the purely indeterministic parts of the \( C_t \) and \( Y_t \) processes. I shall defer until Section IV the task of characterizing the restrictions that (1) and (2) impose across the deterministic parts of the \( (C_t, Y_t) \) process.

I now assume that \( Y_t \) is strictly linearly indeterministic (or that the deterministic parts have been subtracted). I further restrict things by supposing that \( Y_t \) is an \( n \)th-order Markov process

\[ Y_t = \delta_1 Y_{t-1} + \ldots + \delta_n Y_{t-n} + \epsilon_t, \]  

(7)

\(^{12}\) For a proof of Wold’s theorem, see Anderson (1971, chap. 7).
where \( \varepsilon_t \) is a covariance stationary stochastic process with \( \mathbb{E}\varepsilon_t = 0 \), \( \mathbb{E}\varepsilon_t^2 = \sigma_{\varepsilon}^2 \) and where \( \mathbb{E}\varepsilon_t Y_{t-s} = 0 \) for all \( s > 0 \). Write (7) in matrix form as

\[
x_t = A x_{t-1} + \eta_t
\]  
(8)

where

\[
x_t = \begin{bmatrix}
Y_t \\
Y_{t-1} \\
. \\
. \\
. \\
Y_{t-n+1}
\end{bmatrix}, \quad \eta_t = \begin{bmatrix}
\varepsilon_t \\
0 \\
0 \\
. \\
. \\
0
\end{bmatrix},
\]

and

\[
A = \begin{bmatrix}
\delta_1 & \delta_2 & \ldots & \delta_n \\
1 & 0 & \ldots & 0 \\
0 & 1 & 0 & 0 \\
. \\
. \\
0 & 0 & 1 & 0
\end{bmatrix}
\]

We can recover \( Y_t \) from \( x_t \) by \( Y_t = d x_t \), where \( d = [1, 0, \ldots, 0] \), a \((1 \times n)\) row vector. From (8) we have

\[
x_{t+1} = A x_t + \eta_{t+1}, \\
x_{t+2} = A^2 x_t + A \eta_{t+1} + \eta_{t+2}, \\
\vdots \\
x_{t+j} = A^j x_t + A^{j-1} \eta_{t+1} + \ldots + \eta_{t+j}.
\]

Since by assumption \( \mathbb{E}\varepsilon_{t+j} \cdot x_t' = 0 \) for all \( j > 0 \) (because \( \mathbb{E}\varepsilon_{t+j} Y_t = 0 \) for all \( j > 0 \)), we have

\[
E_t x_{t+j} = A^j x_t.
\]  
(9)

If the eigenvalues of \( A \) are distinct, \( A \) can be expressed as

\[
A = P \Lambda P^{-1},
\]  
(10)

where the columns of the \( n \times n \) matrix \( P \) are the eigenvectors of \( A \) and \( \Lambda \) is a diagonal matrix of eigenvalues of \( A \). Using (10), we can write

\[
A^j = P \Lambda^j P^{-1}.
\]  
(11)

Thus we have \( E_t x_{t+j} = P \Lambda^j P^{-1} x_t \) and in particular

\[
E_t Y_{t+j} = dE_t x_{t+j}, \\
= dP \Lambda^j P^{-1} x_t.
\]  
(12)
To form permanent income, we take the appropriate weighted sum of (12),

\[ Y_{pt} = (1 - \alpha) \sum_{j=0}^{\infty} \alpha^j E_t Y_{t+j}, \]

\[ = (1 - \alpha) dP \left( \sum_{j=0}^{\infty} \alpha^j \lambda^j \right) P^{-1} x_t, \tag{13} \]

where \( H_\alpha \) is the diagonal matrix

\[
H_\alpha = \begin{pmatrix}
\frac{1}{1 - \alpha \lambda_1} & 0 & \cdots & 0 \\
0 & \frac{1}{1 - \alpha \lambda_2} & & \\
& \ddots & \ddots & \\
0 & \cdots & 0 & \frac{1}{1 - \alpha \lambda_n}
\end{pmatrix},
\]

where the \( \lambda_i \)'s are the eigenvalues of \( A \). I am assuming that \( |\alpha \lambda_i| < 1 \) for all \( i \), in order to guarantee that the infinite series \( \sum_{j=0}^{\infty} \alpha^j \lambda^j \) converges to \( 1/(1 - \alpha \lambda_i) \). For \( \alpha < 1 \), this condition is weaker than the condition \( |\lambda_i| < 1 \) for all \( i \), which is the condition required if \( \{Y_t\} \) is to be a stationary stochastic process. It is possible for permanent income (2) to be a well-defined convergent series even if income is nonstationary in the sense that \( \max |\lambda_i| > 1 \). Provided \( \alpha < 1 \), permanent income is thus potentially well defined by (2) even in the face of a stochastically upward-trending rate of income. Equation (13) is a convenient closed form for \( Y_{pt} \) in terms of the eigenvalues and eigenvectors of \( A \) (i.e., in terms of the roots of \( 1 - \delta_1 z \ldots - \delta_n z^n = 0 \)), which can easily be implemented on a computer.

In summary, we have the structure

\[ C_t = B Y_{pt} + u_t, \tag{1} \]

\[ x_t = A x_{t-1} + \eta_t, \tag{8} \]

\[ Y_t = dx_t, \]

\[ Y_{pt} = (1 - \alpha) dPH_\alpha P^{-1} x_t, \tag{13} \]

We thus have the observable consumption function,

\[ C_t = B(1 - \alpha) dPH_\alpha P^{-1} x_t + u_t. \tag{14} \]

The force of the rational expectations hypothesis and the restrictions imposed on the error process \( u \) is to impose (highly nonlinear) restrictions across the exogenous \( Y \) process (8) and the observable consumption function (14). The most straightforward method of testing the model is to
estimate (8) and (14) simultaneously using the method of maximum likelihood, estimating first subject to the cross-equation constraints imposed by (8) and (14) and then with those constraints relaxed. A likelihood ratio test can be computed to evaluate the validity of the constraints.

I propose the following algorithm for estimating the vector autoregression for the \((Y, C)\) process under the restrictions implied by rational expectations, an algorithm applicable when \(Y\) is strictly econometrically exogenous with respect to \(C\). Suppose that \(u_t\) follows the second-order Markov process \(u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \xi_t\), where \(\xi_t\) is a serially uncorrelated process with mean zero. I assume that \(\xi_t\) is orthogonal to the entire \(Y\) process. (It will be clear how to proceed where \(u_t\) follows an \(n\)th-order Markov process.) Quasi differencing (1) then gives \(C_t - \rho_1 C_{t-1} - \rho_2 C_{t-2} = B[Y_{pt} - \rho_1 Y_{pt-1} - \rho_2 Y_{pt-2}] + \xi_t\). Since \(E_{t-1} \xi_t = 0\), projecting both sides of the equation above on things dated \(t - 1\) and earlier gives

\[
E_{t-1} C_t - \rho_1 C_{t-1} - \rho_2 C_{t-2} = B[E_{t-1} Y_{pt} - \rho_1 Y_{pt-1} - \rho_2 Y_{pt-2}].
\]  

But we know that

\[
E_{t-1} Y_{pt} = (1 - \alpha) dP \left( \frac{\lambda_i}{1 - \alpha \lambda_i} \right) P^{-1} x_{t-1},
\]

\[
Y_{pt-1} = (1 - \alpha) dP \left( \frac{1}{1 - \alpha \lambda_i} \right) P^{-1} x_{t-1},
\]

\[
Y_{pt-2} = (1 - \alpha) dP \left( \frac{1}{1 - \alpha \lambda_i} \right) P^{-1} x_{t-2}.
\]

Substituting these into equation (15) gives

\[
E_{t-1} C_t = \rho_1 C_{t-1} + \rho_2 C_{t-2} + B(1 - \alpha) dP \left( \frac{\lambda_i}{1 - \alpha \lambda_i} \right) P^{-1} x_{t-1}
\]

\[
- \rho_1 B(1 - \alpha) dP \left( \frac{1}{1 - \alpha \lambda_i} \right) P^{-1} x_{t-1}
\]

\[
- \rho_2 B(1 - \alpha) dP \left( \frac{1}{1 - \alpha \lambda_i} \right) P^{-1} x_{t-2},
\]

Equation (16) gives the vector autoregression for \(C\) in terms of the \(\rho\)'s that characterize the Markov process for \(u\) and the \(\delta\)'s that characterize the autoregression for \(Y\), and determine \(A = \Pi \Lambda P^{-1}\).

If we had assumed that \(u_t\) followed the \(s\)th-order Markov process
\[ u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \ldots + \rho_s u_{t-s} + \xi_t, \] where \( E_{t-1} \xi_t = 0, \) then (16) would be

\[ E_{t-1} C_t = \rho_1 C_{t-1} + \rho_2 C_{t-2} + \ldots + \rho_s C_{t-s} \]

\[ + B(1 - \alpha) \ dP \left( \frac{\lambda_i - \rho_1}{1 - \alpha \lambda_i} \right) P^{-1} x_{t-1} \]

\[ - B(1 - \alpha) \ dP \left( \frac{1}{1 - \alpha \lambda_i} \right) P^{-1} (\rho_2 x_{t-2} + \rho_3 x_{t-3} + \ldots + \rho_s x_{t-s}). \]

Proceeding with our second-order example, we can now write the vector autoregression for \((Y_t, C_t)\) as

\[ Y_t = \delta_1 Y_{t-1} + \ldots + \delta_n Y_{t-n} + a_{yt}, \]

\[ C_t = \rho_1 C_{t-1} + \rho_2 C_{t-2} + B(1 - \alpha) \ dP \left( \frac{\lambda_i - \rho_1}{1 - \alpha \lambda_i} \right) P^{-1} x_{t-1} \]

\[ - \rho_2 B(1 - \alpha) \ dP \left( \frac{1}{1 - \alpha \lambda_i} \right) P^{-1} x_{t-2} + a_{ct}, \]

where the innovations \((a_{yt}, a_{ct})\) obey \( E_{t-1} a_{yt} = E_{t-1} a_{ct} = 0\) and where \( a_{yt} \equiv \xi_t. \) Assuming that \((a_{yt}, a_{ct})\) is jointly normally distributed, given a sample over \( t = 1, \ldots, T, \) maximum likelihood estimates can be obtained by minimizing \(|V|\) where

\[ V = \begin{pmatrix}
\frac{1}{T} \sum_{t=1}^{T} \tilde{a}_{yt}^2 & \frac{1}{T} \sum_{t=1}^{T} \tilde{a}_{yt} \tilde{a}_{ct} \\
\frac{1}{T} \sum_{t=1}^{T} \tilde{a}_{ct} \tilde{a}_{ct} & \frac{1}{T} \sum_{t=1}^{T} \tilde{a}_{ct}^2
\end{pmatrix}.
\]

The minimization is carried out over the parameters \( \delta_1, \ldots, \delta_n, \rho_1, \rho_2, B, \) and \( \alpha. \) The log likelihood function at the maximum likelihood parameter estimates can be shown to be (see Wilson 1973) \( \log L = -\frac{1}{2} m T \log 2\pi - \frac{1}{2} T (\log |V| + m), \) where \( m \) is the number of variables being modeled (in our case 2, \( C, \) and \( Y)). \]

To test the model, it is appropriate to estimate the vector autoregression twice: once subject to the constraints in (16), another time unconstrained. Let \( \log L_u \) be the value of the log likelihood unconstrained, while \( \log L_r \) is its value under the constraints. Then \(-2 (\log L_r - \log L_u)\) is asymptotically distributed as \( \chi^2 \) with \( q \) degrees of freedom, where \( q \) is the number of restrictions imposed. Equivalently, we can express the likelihood ratio as \( T \{ \log |V_r| - \log |V_u| \}, \) where \( V_r \) and \( V_u \) are the values of \( V \) under the restriction and unrestricted, respectively. \(^{13}\)

\(^{13}\) Sims (personal communication) points out that in small samples using \( T \) rather than \((T - k), \) where \( k \) is the number of parameters on the right side of a typical equation under the unrestricted parameterization, biases the test statistic against the null hypothesis. As it will turn out, correcting for this bias will not reverse any of my conclusions.
Notice that the parameters $B$ and $\alpha$ are estimated only from the systematic part of the vector autoregression of the $(C, Y)$ process. In particular, information about the contemporaneous covariance matrix of the innovations $a_{ct}, a_{yt}$ is not used in estimating the parameters $B$ and $\alpha$. This estimator takes no stand on how the contemporaneous covariance between $a_{ct}$ and $a_{yt}$ is to be accounted for. In Granger's language, this estimator permits "instantaneous feedback" between $C$ and $Y$ in either direction but ignores it in estimating $B$ and $\alpha$. Notice, however, that Sims's theorems assure us that, if $C$ fails to Granger-cause $Y$, then there exists a model of the following form that is consistent with the data:

\[
C_t = \sum_{i=0}^{\infty} h_i Y_{t-i} + \xi_t, \quad (\dagger)
\]
\[
Y_t = \sum_{i=1}^{\infty} \delta_i Y_{t-i} + \psi_t,
\]

where $E\xi_t \psi_s = 0$ for all $t$ and $s$. In this model, all contemporaneous covariance between the innovations $a_{ct}$ and $a_{yt}$ is swept into the first equation, in our case the consumption function. A maximum likelihood estimator of $B$ and $\alpha$ could be based on this model provided that the instantaneous feedback between $Y$ and $C$ is partitioned a priori, as it is, for example, by the model of Section I. In that model, all instantaneous feedback flows from $Y$ to $C$ through the consumption function, thereby in effect identifying the first equation of $(\dagger)$ as the consumption function. A maximum likelihood estimator along these lines uses more information than is incorporated by the estimator described above.\footnote{The sacrifice of information occurs in the step leading to eq. (15). The estimator imposes only the restriction across the systematic part of the vector autoregression that is contained in (15).} However, at this point I prefer an estimator that does not, for example, attribute all contemporaneous covariance between $a_{yt}$ and $a_{ct}$ to the workings of the consumption function. The estimator that I have recommended above does permit instantaneous causality to flow from $C$ to $Y$, a feature that would characterize a class of macroeconomic models in which $C$ still fails to Granger-cause $Y$.\footnote{An example can easily be constructed along the lines of the model described by Sargent (1976). In particular, unexpected movements in $u_t$ are permitted to have effects on disposable income via the sort of mechanism described by Sargent (1973).} In summary, the maximum likelihood estimator that estimates $B$ and $\alpha$ by imposing the restrictions (16) requires weaker conditions on the covariance between $u_t$ and disposable income than are imposed, for example, by the macroeconomic model of Section II.\footnote{The following two-step estimator is appropriate where $Y$ is econometrically exogenous in the consumption function, so that all contemporaneous correlation between $a_{yt}$ and $a_{ct}$ is appropriately attributed to the workings of the consumption function. First estimate (8) by least squares. Then obtain the eigenvalues and eigenvectors of $A$, and estimate (14) by forming $(1 - \alpha)\delta^T \Delta^{-1} x_t$ for various values of $\alpha$ and searching over $\alpha$ for
IV. Estimation of the Consumption Function under Rational Expectations with Income Endogenous

A. Purely Indeterministic Process

Now return to the structure (1) and (2) and drop the assumption that $Y_t$ is orthogonal to $u_s$ at all $t$ and $s$. This change in assumption permits income to be econometrically endogenous and renders the preceding estimation procedures invalid; that is, the estimators for $B$ and $\alpha$ proposed above will not in general be statistically consistent when $Y$ fails to be econometrically exogenous in (1). In this section I propose an estimator that, under special assumptions, is statistically consistent even where $Y$ is not econometrically exogenous in (1). As will be shown, in obtaining consistent estimates, it is important to control for serial correlation in $u_t$ as long as our information is confined to observations on the $(C, Y)$ process alone. I begin by working things out for the case in which the $(C_t, Y_t)$ process is purely linearly indeterministic. Later in the section, it is assumed that the $(C_t, Y_t)$ process contains both deterministic and indeterministic parts.

Combining (1) with (2) gives

$$C_t = B(1 - \alpha) \left( \sum_{j=0}^{\infty} \alpha^j F \right) E_t Y_t + u_t,$$

or

$$C_t = B(1 - \alpha) \frac{1}{1 - \alpha^F} E_t Y_t + u_t,$$

(17)

where $F$ is the operator defined by $F E_t Y_{t+k} = E_t Y_{t+k+1}$, so that application of $F$ leaves the conditioning set unchanged while shifting the time index on the random variable operated upon forward by one period. Operating on both sides of (17) by the operator $(1 - \alpha F)$ gives

$$C_t = \alpha E_t C_{t+1} + B(1 - \alpha) Y_t + u_t - \alpha E_t u_{t+1}.$$  

(18)

Now suppose that $u_t$ follows the first-order Markov process

$$u_t = \rho u_{t-1} + w_t,$$

(19)

where $w_t$ is serially uncorrelated and orthogonal to lagged $Y$'s and $C$'s. Subtracting $\rho C_{t-1}$ from both sides of (18) and noting that $E_t u_{t+1} = \rho u_t$ gives

$$C_t - \rho C_{t-1} = \alpha (E_t C_{t+1} - \rho E_{t-1} C_t) + B(1 - \alpha)(Y_t - \rho Y_{t-1}) + (1 - \alpha \rho) u_t - \rho (1 - \alpha \rho) u_{t-1}.$$  

Projecting both sides of the equation above on information dated $t - 1$ and earlier gives

$$E_{t-1} (C_t - \rho C_{t-1}) = \alpha E_{t-1} (C_{t+1} - \rho C_t) + B(1 - \alpha) E_{t-1} (Y_t - \rho Y_{t-1}).$$

(20)

the value that delivers the minimum sum of squared residuals in (14). The resulting estimates of $B$ and $\alpha$ are consistent on the assumptions of the model, though they are not fully efficient. The model could be tested by adding to the right side of (14) current, past, and future values of $Y$ and testing the null hypothesis that these intruding variables bear zero coefficients. Finding nonzero coefficients on future $Y_t$'s would reflect negatively on the assumed exogeneity of $Y$ in (1).
Equation (20) is a restriction across the systematic part of the vector autoregression of the \((C_t - \rho C_{t-1}, Y_t - \rho Y_{t-1})\) process. To characterize the restriction, let the systematic part of the nth-order vector autoregression of \((C_t - \rho C_{t-1}, Y_t - \rho Y_{t-1}) \equiv (c_t, y_t)\) be\(^{17}\)

\[
E_{t-1}c_t = c_1 y_{t-1} + c_2 y_{t-2} + \ldots + c_n y_{t-n} + d_1 c_{t-1} + d_2 c_{t-2} + \ldots + d_n c_{t-n},
\]

\[
E_{t-1}y_t = a_1 y_{t-1} + a_2 y_{t-2} + \ldots + a_n y_{t-n} + b_1 c_{t-1} + b_2 c_{t-2} + \ldots + b_n c_{t-n}.
\]

Application of the "chain rule of forecasting" gives\(^{18}\)

\[
E_{t-1}c_{t+1} = c_1 E_{t-1} y_t + c_2 y_{t-1} + \ldots + c_n y_{t-n+1} + d_1 E_{t-1} c_{t} + d_2 c_{t-1} + \ldots + d_n y_{t-n+1}.
\]

Substituting from (21) into (22) and rearranging gives

\[
E_{t-1}c_{t+1} = (c_1 a_1 + c_2 + d_1 c_1) y_{t-1} + (d_1 a_2 + c_3 + d_1 c_2) y_{t-2} + \ldots + (c_1 a_{n-1} + c_n + d_1 c_{n-1}) y_{t-n+1} + (c_1 a_n + d_1 c_n) y_{t-n} + (c_1 b_1 + d_1^2 + d_2) c_{t-1} + (c_1 b_2 + d_1 d_2 + d_3) c_{t-2} + \ldots + (c_1 b_{n-1} + d_1 d_{n-1} + d_n) c_{t-n+1} + (c_1 b_n + d_1 d_n) c_{t-n}.\]

Using (21) and (23), the restrictions implied by (20) become

\[
c_1 = \beta(1 - \alpha) a_1 + \alpha(d_1 a_1 + c_2 + d_1 c_1),
\]

\[
c_2 = \beta(1 - \alpha) a_2 + \alpha(c_1 a_2 + c_3 + d_1 c_2),
\]

\[\vdots\]

\[
c_{n-1} = \beta(1 - \alpha) a_{n-1} + \alpha(c_1 a_{n-1} + c_n + d_1 c_{n-1}),
\]

\[
c_n = \beta(1 - \alpha) a_n + \alpha(c_1 a_n + d_1 c_n),
\]

\[
d_1 = \beta(1 - \alpha) b_1 + \alpha(c_1 b_1 + d_1^2 + d_2),
\]

\[
d_2 = \beta(1 - \alpha) b_2 + \alpha(c_1 b_2 + d_1 d_2 + d_3),
\]

\[\vdots\]

\[
d_{n-1} = \beta(1 - \alpha) b_{n-1} + \alpha(c_1 b_{n-1} + d_1 d_{n-1} + d_n),
\]

\[
d_n = \beta(1 - \alpha) b_n + \alpha(c_1 b_n + d_1 d_n).
\]

\(^{17}\) Notice that \((c_t, y_t)\) depends on \(\rho\).

\(^{18}\) The "chain rule of forecasting" is a simple implication of the "law of iterated projections." That law states that, where \(E\) is the linear least-squares projection operator and \(y, x, z\) are random variables,

\[
E(y \mid z) = E(E(y \mid x, z) \mid z).
\]

This law is easily proved as an implication of the fact that linear least-squares projections are uniquely determined by the condition that their forecast errors be orthogonal to the conditioning variables. To get the chain rule of forecasting, shift (21) forward one period, then project both sides of the result on information available at \((t - 1)\), applying (\(\dagger\dagger\)). An analogous "law of iterated expectations" and chain rule of forecasting are also well-known properties of conditional mathematical expectations.
These \((2n)\) equations determine \(a_1, \ldots, a_n, b_1, \ldots, b_n\) as functions of \(\beta, \alpha, c_1, \ldots, c_n, d_1, \ldots, d_n.\) For example, notice that the first equation determines \(a_1\) as a function of \(\beta, \alpha, c_1, c_2,\) and \(d_1.\) The equations in (24) embody the restrictions that equations (1) and (2) and the assumption on the error process (19) impose on the systematic part of the \(n\)th-order vector autoregression for \((c, y).\)

Write the vector autoregression for \((c_t, y_t)\) as

\[
\begin{align*}
c_t &= c_1 y_{t-1} + \ldots + c_n y_{t-n} + d_1 c_{t-1} + \ldots + d_n c_{t-n} + a_{ct}; \\
y_t &= a_1 y_{t-1} + \ldots + a_n y_{t-n} + b_1 c_{t-1} + \ldots + b_n c_{t-n} + a_{yt},
\end{align*}
\]

(25)

where \((a_{ct}, a_{yt})\) are disturbances orthogonal to \(y_{t-1}, \ldots, y_{t-n}, c_{t-1}, \ldots, c_{t-n}.\) Let \(a_t = (a_{ct}, a_{yt}).\) Given a sample extending over a period \(t = 1, \ldots, T,\) maximum likelihood estimates of (25) under the restrictions (24) are obtained by minimizing \(|V| = \left| \frac{1}{T} \sum_{t=1}^{T} a_t a_t^\top \right|\) with respect to the free parameters \(\beta, \alpha, c_1, \ldots, c_n, d_1, \ldots, d_n.\) To test the model, (25) can be estimated unconstrained and then the appropriate likelihood ratio statistic can be constructed. Maximum likelihood estimation of (25) subject to no constraints amounts to least-squares estimation of each equation separately.

### B. Deterministic Components Present

So far, my calculations have assumed that consumption and income are purely linearly indeterministic processes or else that the model (1) and (2) applies only to the indeterministic parts of consumption and income. I now relax this assumption and assume instead that the \((y_t, c_t)\) process is governed by

\[
\begin{align*}
c_t &= c_1 y_{t-1} + \ldots + c_n y_{t-n} + d_1 c_{t-1} + \ldots + d_n c_{t-n} + \eta^c_t + a_{ct}; \\
y_t &= a_1 y_{t-1} + \ldots + a_n y_{t-n} + b_1 c_{t-1} + \ldots + b_n c_{t-n} + \eta^y_t + a_{yt},
\end{align*}
\]

(26)

where \((a_{ct}, a_{yt})\) are orthogonal to \(y_{t-1}, c_{t-1}, \ldots, y_{t-n}, c_{t-n}, \eta^c_t,\) and \(\eta^y_t\) and where \(\eta^c_t\) and \(\eta^y_t\) are deterministic components that are given by

\[
\begin{align*}
\eta^c_t &= \gamma^c t + \delta^c_1 s_{1t} + \delta^c_2 s_{2t} + \delta^c_3 s_{3t} + \delta^c_4 s_{4t} , \\
\eta^y_t &= \gamma^y t + \delta^y_1 s_{1t} + \delta^y_2 s_{2t} + \delta^y_3 s_{3t} + \delta^y_4 s_{4t}.
\end{align*}
\]

(27)

Here the seasonal dummy \(s_{1t}\) equals unity every fourth quarter and zero in the intervening three quarters, and \(s_{2t+1} = s_{1t}, s_{3t+1} = s_{2t}, s_{4t+1} = s_{3t},\) and \(s_{1t+1} = s_{4t}.\) Notice that \(s_{1t} + s_{2t} + s_{3t} + s_{4t} = 1\) for all \(t.\) I assume that agents are able to forecast \(\eta^c_t\) and \(\eta^y_t\) perfectly. Then, from
the chain rule of forecasting, we have

\[ E_{t-1} c_{t+1} = c_1 \{ a_1 \gamma_{t-1} + \ldots + a_n \gamma_{t-n} + b_1 c_{t-1} + \ldots + b_n c_{t-n} + \eta_t^c \} \]

\[ + c_2 \gamma_{t-1} + \ldots + c_n \gamma_{t-n+1} \]

\[ + d_1 \{ c_1 \gamma_{t-1} + \ldots + c_n \gamma_{t-n} + d_1 c_{t-1} + \ldots + d_n c_{t-n} + \eta_t^c \} \]

\[ + d_2 c_{t-1} + \ldots + d_n c_{t-n+1} + \eta_t^{c+1} \]

\[ = \eta_{t+1}^{c+1} + d_1 \eta_t^c + c_1 \eta_t^c + \text{(terms in lagged c's and y's)}. \]

Equation (20) then places the following restriction across the \( \eta_t^c \)'s:

\[ \eta_t^c = \beta (1 - \alpha) \eta_t^y + \alpha [c_1 \eta_t^c + d_1 \eta_t^c + \eta_{t+1}^c], \]

or \[ [1 - \alpha d_1] \eta_t^c = [\beta (1 - \alpha) + \alpha c_1] \eta_t^c + \alpha \eta_t^{c+1}. \]

Substituting from (27) gives

\[ (1 - \alpha d_1) [\gamma t + \delta_1^s_{1t} + \delta_2^s_{2t} + \delta_3^s_{3t} + \delta_4^s_{4t}] = \]

\[ [\beta (1 - \alpha) + \alpha c_1] (\gamma t + \delta_1^s_{1t} + \delta_2^s_{2t} + \delta_3^s_{3t} + \delta_4^s_{4t}] \]

\[ + \alpha (\gamma t + \gamma (s_{1t} + s_{2t} + s_{3t} + s_{4t}) + \delta_5^c_{1t+1} + \delta_6^c_{2t+1} \]

\[ + \delta_7^c_{3t+1} + \delta_8^c_{4t+1}). \]

Noting that \( s_{1t+1} = s_{4t}, s_{2t+1} = s_{1t}, s_{3t+1} = s_{2t}, s_{4t+1} = s_{3t}, \) we deduce the following restrictions:

\[ (1 - \alpha d_1) \gamma = [\beta (1 - \alpha) + \alpha c_1] \gamma + \alpha \gamma^c; \]

\[ (1 - \alpha d_1) \delta_1^c = [\beta (1 - \alpha) + \alpha c_1] \delta_1^c + \alpha (\gamma^c + \delta_2^c), \]

\[ (1 - \alpha d_1) \delta_2^c = [\beta (1 - \alpha) + \alpha c_1] \delta_2^c + \alpha (\gamma^c + \delta_3^c), \]

\[ (1 - \alpha d_1) \delta_3^c = [\beta (1 - \alpha) + \alpha c_1] \delta_3^c + \alpha (\gamma^c + \delta_4^c), \]

\[ (1 - \alpha d_1) \delta_4^c = [\beta (1 - \alpha) + \alpha c_1] \delta_4^c + \alpha (\gamma^c + \delta_1^c). \]

These five equations determine \( \gamma^c, \delta_1^c, \delta_2^c, \delta_3^c, \) and \( \delta_4^c \) as functions of \( \beta, \alpha, \gamma^c, \delta_1^c, \delta_2^c, \delta_3^c, \delta_4^c, c_1, \) and \( d_1. \) It is readily verified that the restrictions of (24) continue to hold across the \( a, b, c, \) and \( d \)'s. Thus equations (24) and (28) summarize the restrictions that the theory imposes on the parameters of the vector autoregression (26). Once again, the strategy is to estimate (26) by maximum likelihood, first subject to (24) and (28), then unconstrained, and to compute the pertinent likelihood ratio statistic.

V. Appropriate Data

The preceding sections have indicated the restrictions that the rational expectations hypothesis imposes across the autoregressive representation of the \((C, Y)\) process. Evidently these restrictions are applicable to the unadulterated seasonally unadjusted data.\(^{19}\) Even in the presence of a

\(^{19}\) Here I am assuming that agents care about and see their seasonally unadjusted income flow.
linearly indeterministic seasonal in $Y$, model (1)-(2) is supposed to hold. The restrictions in general are predicted not to hold for seasonally adjusted data. Thus let seasonally adjusted consumption and income, $C_t^a$ and $Y_t^a$, respectively, be related to $C_t$ and $Y_t$ by $C_t^a = f(L)C_t$, $Y_t^a = g(L)Y_t$, where $f(L)$ and $g(L)$ are two-sided seasonal-adjustment filters. Wallis (1974) has argued that actual seasonal-adjustment procedures are well approximated by the application of such linear filters. The cross-covariance-generating matrix of the adjusted data is given by

$$
\begin{bmatrix}
S_c^a(z) & S_{cy}^a(z) \\
S_{yc}^a(z) & S_y^a(z)
\end{bmatrix}
= \begin{bmatrix}
f(z)f(z^{-1})S_c(z) & f(z)g(z^{-1})S_{cy}(z) \\
g(z)f(z^{-1})S_{yc}(z) & g(z)g(z^{-1})S_y(z)
\end{bmatrix},
$$

where $S_c(z)$, $S_y(z)$, and $S_{cy}(z)$ are the covariance-generating functions of $C$ and $Y$ and the cross-covariance-generating function between $C$ and $Y$, respectively. Restrictions that hold across elements of the cross-covariance-generating function of the unadjusted data will not in general hold across elements of the cross-covariance-generating function for the left-hand side. To take an example, with $Y$ exogenous, for the unadjusted data the generating function for the coefficients in the projection of $C_t$ on the $Y_t$ process, $S_{cy}(z)/S_y(z)$ is one-sided and has common parameters with the covariance-generating function of $Y$, $S_y(z) = \sigma^2b(z)b(z^{-1})$, as indicated in Section III. The projection of $C_t^a$ on the $Y_t^a$ process has coefficients with generating function $[f(z)/g(z)]S_{cy}(z)/S_y(z)$ which in general is two-sided, unless $f(z) = g(z)$, and in any event does not obey the Section III restrictions with the parameters of the covariance-generating function of the adjusted $Y$, $S_y^a(z) = \sigma^2g(z)g(z^{-1})b(z)b(z^{-1})$.

Below I analyze both seasonally adjusted and unadjusted data. This is done partly because other investigators have usually used seasonally adjusted data and partly because many readers may find the arguments of Sims (1976) compelling.

VI. Exogeneity Tests

For the postwar United States for both seasonally adjusted and unadjusted data I have defined the following variables: $Y_1 = \text{GNP} - \text{capital con-}

---

20 The position taken in this section is not at all unanimously accepted by serious students of rational expectations models. Sims (1976) has made an ingenious argument supporting the opposite position, that even if agents are responding to the seasonally unadjusted data it is advisable to use seasonally adjusted data to implement and test rational expectations cross-equation restrictions. Sims's argument has several elements, one of which is a desire to have procedures that are robust to departures from stationarity together with the suspicion that departures from stationarity are in some sense most likely at the seasonal frequencies. The argument advanced in the text is predicated on the assumption of stationarity in the sense that the $\delta$'s of Sec. III and the $a$, $b$, $c$, and $d$'s of Sec. IV are required to be constant over time. (As mentioned earlier, though, we can tolerate nonstationarity in the sense of maximum eigenvalues of the transition matrix $A$ exceeding unity, as long as $|\alpha\lambda| < 1$.)
sumption – federal tax receipts + transfers; $C =$ consumption of durables + consumption of nondurables + consumption of services; $I =$ residential construction + nonresidential fixed investment + change in business inventories; $Y_2 = C + I$; $\bar{I} = I +$ consumption of durables; $D_t = 0.975$ (consumption of durables) + 0.95 $D_{t-1}$; and $\bar{C}_t = C_t - 0.9875$ (consumption of durables) + 0.025 $D_t$. The data are quarterly and "real." The seasonally adjusted data are deflated by implicit price deflators and are measured in 1972 dollars. The seasonally unadjusted data were deflated by quarterly averages of the consumer price index and are measured in 1967 dollars. The sources of the data are given in the Appendix.

The disposable-income concept $Y_2$ differs from the usual disposable-income concept $Y_1$ by in effect subtracting from $Y_1$ the rate of real government deficit.\(^21\) To the extent that the government deficit is financed by issuing bonds and that agents fully discount anticipated future taxes levied to service the debt, $Y_2$ is the appropriate disposable-income concept.\(^22\)

Table 1 reports the results of testing for Granger causality over the period 1948III–72IV. The regressions computed were of the form $z_t = \sum\limits_{i=1}^{m} a_i z_{t-i} + \sum\limits_{j=1}^{n} \beta_j x_{t-j} + \text{residual}$. The null hypothesis that $x$ fails to Granger-cause $z$ is tested by testing the null hypothesis $\beta_1 = \ldots = \beta_n = 0$. The table reports the marginal significance levels associated with the $F$-statistic pertinent for testing this null hypothesis. The regressions include a constant, linear trend and, for the seasonally unadjusted data, three seasonal dummies.

For the seasonally unadjusted data, the hypothesis that $Y_1$ is not Granger-caused by $C$ or $\bar{C}$ is generally rejected even at low marginal significance levels. A similar conclusion holds for $Y_2$. These results thus point against acceptance of the version of Tobin's dynamic aggregative model described in Section I. They also indicate that we should use the Section IV estimator, which does not require exogeneity of $Y$, instead of the Section III estimator.

The results are substantially different where the seasonally adjusted data are used.\(^23\) Here $Y_1$ seems not to be Granger-caused by $C$ or $\bar{C}$;

\(^{21}\) Let the national income identity be $C + I + G + \delta K = Y$, where $Y$ is real GNP, $I$ is real investment, $C$ is real consumption, $G$ is real government purchases, and $\delta K$ is depreciation. Then we have

$$Y_1 = Y - T - \delta K,$$

$$= C + I + G - T,$$

where $T$ is the flow of real tax collections net of transfers. Hence $Y_2 = C + I = Y_1 - (G - T)$.

\(^{22}\) For discussions of discounting of taxes and appropriate definitions of wealth and disposable income, see Bailey (1972) and Sargent (1975).

\(^{23}\) The reader of Wallis (1974) and especially Sims (1974) will not be surprised that the seasonally adjusted and seasonally unadjusted data give substantially different results with respect to Granger-causality.
# TABLE 1
Granger Causality Test Results, Marginal Significant Level
A. Data Not Seasonally Adjusted

| Lags: (m, n) | 8-4 | 6-4 | 8-6 | 6-6 |
|-------------|-----|-----|-----|-----|
|             | F(4, 79) | F(4, 83) | F(6, 77) | F(6, 81) |
| Z\textsubscript{t} | X\textsubscript{t} | (1) | (2) | (3) | (4) |
| C \ Y\textsubscript{1} | .2062 | .2135 | .2841 | .3418 |
| Y\textsubscript{1} \ C | .0064 | .0200 | .0245 | .0645 |
| I \ Y\textsubscript{1} | .0943 | .0226 | .0316 | .0073 |
| Y\textsubscript{1} \ I | .0222 | .0283 | .0138 | .0516 |
| C \ Y\textsubscript{2} | .0013 | .0003 | .0003 | .0001 |
| Y\textsubscript{2} \ C | .0250 | .0082 | .0181 | .0063 |
| I \ Y\textsubscript{2} | .0291 | .0333 | .0144 | .0002 |
| Y\textsubscript{2} \ I | .0250 | .0082 | .0181 | .0063 |
| C \ I | .0013 | .0003 | .0003 | .0001 |
| \ C | .1694 | .1450 | .3105 | .2748 |
| \ Y\textsubscript{1} | .0023 | .0073 | .0106 | .0308 |
| \ I | .0909 | .0386 | .0410 | .0187 |
| \ Y\textsubscript{1} \ I | .1013 | .0980 | .0571 | .1333 |
| \ C \ Y\textsubscript{2} | .2108 | .0835 | .0181 | .0064 |
| \ Y\textsubscript{2} \ C | .0871 | .0341 | .0300 | .0009 |
| \ I \ Y\textsubscript{2} | .1945 | .0774 | .0292 | .0103 |
| \ Y\textsubscript{2} \ I | .0881 | .0347 | .0021 | .0006 |
| \ I \ C | .0013 | .0003 | .0003 | .0001 |
| \ Y\textsubscript{1} \ C | .0169 | .1450 | .3105 | .2748 |
| \ Y\textsubscript{1} \ I | .0023 | .0073 | .0106 | .0308 |
| \ I \ Y\textsubscript{1} | .0909 | .0386 | .0410 | .0187 |
| \ Y\textsubscript{1} \ Y\textsubscript{1} | .1013 | .0980 | .0571 | .1333 |
| \ Y\textsubscript{1} \ I | .0023 | .0073 | .0106 | .0308 |
| \ C \ Y\textsubscript{2} | .0131 | .0105 | .0168 | .0142 |
| \ Y\textsubscript{2} \ C | .0760 | .1199 | .0876 | .1272 |
| \ I \ Y\textsubscript{2} | .0012 | .0006 | .0004 | .0003 |
| \ Y\textsubscript{2} \ I | .0760 | .1199 | .0876 | .1272 |
| \ I \ C | .0012 | .0006 | .0004 | .0003 |
| \ \ C | .0131 | .0105 | .0168 | .0142 |
| \ Y\textsubscript{1} \ C | .0029 | .0036 | .0050 | .0059 |
| \ Y\textsubscript{1} \ I | .3145 | .5615 | .4059 | .7285 |
| \ Y\textsubscript{1} \ Y\textsubscript{1} | .0205 | .0227 | .0260 | .0410 |
| \ Y\textsubscript{1} \ I | .2848 | .3533 | .1902 | .2972 |
| \ C \ Y\textsubscript{2} | .0127 | .0117 | .0241 | .0197 |
| \ Y\textsubscript{2} \ C | .0640 | .1091 | .1647 | .2241 |
| \ I \ Y\textsubscript{2} | .0018 | .0031 | .0063 | .0085 |
| \ Y\textsubscript{2} \ I | .0666 | .1224 | .1618 | .2322 |
| \ I \ \ C | .0016 | .0025 | .0055 | .0070 |
| \ \ C | .0112 | .0099 | .0222 | .0175 |

**B. Seasonally Adjusted Data**

| Lags: (m, n) | 8-4 | 6-4 | 8-6 | 6-6 |
|-------------|-----|-----|-----|-----|
|             | F(4, 82) | F(4, 86) | F(6, 80) | F(6, 84) |
| Z\textsubscript{t} | X\textsubscript{t} | (1) | (2) | (3) | (4) |
| C \ Y\textsubscript{1} | .0144 | .0197 | .0141 | .0134 |
| Y\textsubscript{1} \ C | .3540 | .6709 | .1901 | .4771 |
| I \ Y\textsubscript{1} | .0061 | .0073 | .0190 | .0200 |
| Y\textsubscript{1} \ I | .0386 | .0843 | .1050 | .1850 |
| C \ Y\textsubscript{2} | .0131 | .0105 | .0168 | .0142 |
| Y\textsubscript{2} \ C | .0760 | .1199 | .0876 | .1272 |
| I \ Y\textsubscript{2} | .0012 | .0006 | .0004 | .0003 |
| Y\textsubscript{2} \ I | .0760 | .1199 | .0876 | .1272 |
| I \ C | .0012 | .0006 | .0004 | .0003 |
| \ \ C | .0131 | .0105 | .0168 | .0142 |
| \ Y\textsubscript{1} \ C | .0029 | .0036 | .0050 | .0059 |
| \ Y\textsubscript{1} \ I | .3145 | .5615 | .4059 | .7285 |
| \ Y\textsubscript{1} \ Y\textsubscript{1} | .0205 | .0227 | .0260 | .0410 |
| \ Y\textsubscript{1} \ I | .2848 | .3533 | .1902 | .2972 |
| \ C \ Y\textsubscript{2} | .0127 | .0117 | .0241 | .0197 |
| \ Y\textsubscript{2} \ C | .0640 | .1091 | .1647 | .2241 |
| \ I \ Y\textsubscript{2} | .0018 | .0031 | .0063 | .0085 |
| \ Y\textsubscript{2} \ I | .0666 | .1224 | .1618 | .2322 |
| \ I \ \ C | .0016 | .0025 | .0055 | .0070 |
| \ \ C | .0112 | .0099 | .0222 | .0175 |

Note.—Cols. 1 and 3: dependent variable spans 1949I–72IV; cols. 2 and 4: dependent variable spans 1948III–72IV. All regressions include constant and linear trend; those in section A also include three seasonal dummies. Regressions are of the form

\[ Z_t = \sum_{i=1}^{m} \alpha_i Z_{t-i} + \sum_{i=1}^{n} \beta_i X_{t-i} + \text{residual}. \]

Table reports marginal significance level of F-statistic pertinent for testing null hypothesis \( \beta_1 = \beta_2 = \ldots = \beta_m = 0 \), which is the null hypothesis "x fails to Granger-cause z." Where \( f \) is the calculated value of the pertinent F-statistic, the marginal significance level is defined as \( \text{prob} \{ F > f \} \) under the null hypothesis.
Y₂ seems not to be caused by C, while the F-statistics are marginal at the usual significance levels for testing whether Y₂ is Granger-caused by C. All in all, the results with the seasonally adjusted data are much more consistent with the predictions of the special version of Tobin's model described in Section I. With the seasonally adjusted data, these exogeneity results provide some justification for applying the Section III estimator.²⁴

VII. Tests of Cross-Equation Restrictions

Table 2 reports the results of implementing various versions of the Section IV estimator with the seasonally unadjusted data where n, the order of the vector autoregression, is taken as four. Observations on the dependent variables span 194811–72IV. The table reports the marginal confidence level for the likelihood ratio test, defined as follows: let X be a χ² random variable with q degrees of freedom and let x be the computed value of the test statistic. Then the marginal confidence level is \( \text{Prob} \{ X < x \} \) under the null hypothesis. High values of the marginal confidence level indicate that the null hypothesis is doing badly. Note that the marginal confidence level is one minus the marginal significance level as defined earlier. The various versions of the procedure differ in regard to whether the constraints (28) on the seasonals and trend terms were imposed and whether α and B were estimated or were imposed at the "reasonable" values of α = 0.95 and B = 0.9. The table records "yes" if the constraint on the trend (the first line of [28]) is imposed and "yes" if the constraints on the seasonals (the next four lines of [28]) are imposed. The table also reports the Kolmogorov–Smirnov statistic on the cumulated periodogram of the consumption and income residuals from the constrained autoregressions. Tables for the distribution of this statistic are reported by Durbin (1969). Finally, we set \( \rho = 1 \) in all of the Section IV calculations.

Overall, the pattern of results in table 2 calls for rejection of model (1)–(2) as supplemented by the hypothesis (19) that the disturbance \( u_t \) follows a first-order Markov process with \( \rho = 1 \). In most cases, the likelihood ratio test emphatically calls for rejection of the model. Further, very unreasonable parameter estimates are obtained, especially of B. Where they are left unconstrained, the point estimates of α do not resemble quarterly discount factors. Those estimates that correspond to likelihood ratios that do not call for rejecting the model typically involve theoretically unacceptable point estimates. Notice in particular that lines 1, 5, 11, and 15

²⁴ Hall's (1977) hypothesis that consumption is not Granger-caused by income is generally soundly rejected for the seasonally adjusted data summarized in table 1B; on the other hand, with respect to some but not all information sets, Hall's hypothesis cannot be rejected on the basis of the seasonally unadjusted data in table 1A. But the regressions underlying table 1A rather sharply contradict Hall's hypothesis that consumption is a first-order Markov process as opposed to a higher-order process, which is no surprise with seasonally unadjusted data.
of table 2, which impose the rational expectations constraints across the trends and seasonals, do not yield estimates of the discount factor $\alpha$ that would cause one to embrace the permanent-income theory over a theory in which consumption is related only to contemporaneous income.

Table 3 reports some results for the seasonally adjusted data. Qualitatively, the results are much the same as for the unadjusted data and are discouraging from the point of view of the model.

Table 4 reports results obtained by using the estimator of Section III where $Y_t$ is assumed to be an exogenous fourth-order Markov process and $u_t$ a second-order Markov process. Once again, when plausible values of $B$ and/or $\alpha$ are imposed, the likelihood ratio statistics call for rejection of the
### TABLE 3
**SECTION IV ESTIMATOR WHICH ASSUMES INCOME IS ENDOGENOUS**
(Seasonally Adjusted) *

| \( \alpha \) | \( B \) | Constraints on Trend | Constraints on Seasonal | Marginal Likelihood Ratio | \( K-S \) Statistic for C Residual | \( K-S \) Statistic for Y Residual |
|-------|------|----------------------|------------------------|--------------------------|-------------------------------|-------------------------------|
| .95   | 4.320 | Yes                  | No                     | .997                     | .0498                         | .0265                         |
| .95   | 6.06 \( \times 10^7 \) | No                  | No                     | .697                     | .0385                         | .0294                         |
| .95   | .90   | No                   | No                     | 1.000                    | .0528                         | .0381                         |

| \( \alpha \) | \( B \) | Constraints on Trend | Constraints on Seasonal | Marginal Likelihood Ratio | \( K-S \) Statistic for C Residual | \( K-S \) Statistic for Y Residual |
|-------|------|----------------------|------------------------|--------------------------|-------------------------------|-------------------------------|
| .95   | 3.471 | Yes                  | No                     | 1.000                    | .0538                         | .0233                         |
| .95   | 3.561 | No                   | No                     | 1.000                    | .0545                         | .0240                         |
| .95   | .90   | No                   | No                     | 1.000                    | .0554                         | .0380                         |

| \( \alpha \) | \( B \) | Constraints on Trend | Constraints on Seasonal | Marginal Likelihood Ratio | \( K-S \) Statistic for C Residual | \( K-S \) Statistic for Y Residual |
|-------|------|----------------------|------------------------|--------------------------|-------------------------------|-------------------------------|
| .95   | 1.22 \( \times 10^7 \) | Yes                  | No                     | .6463                    | .0278                         | .0316                         |
| .95   | 1.91 \( \times 10^7 \) | No                   | No                     | .6138                    | .0271                         | .0294                         |
| .95   | .90   | No                   | No                     | .979                     | .0335                         | .0350                         |

| \( \alpha \) | \( B \) | Constraints on Trend | Constraints on Seasonal | Marginal Likelihood Ratio | \( K-S \) Statistic for C Residual | \( K-S \) Statistic for Y Residual |
|-------|------|----------------------|------------------------|--------------------------|-------------------------------|-------------------------------|
| .95   | 2.372 | Yes                  | No                     | .980                     | .0405                         | .0240                         |
| .95   | 2.471 | No                   | No                     | .983                     | .0419                         | .0248                         |
| .95   | .90   | No                   | No                     | .997                     | .0450                         | .0311                         |

* Observations on the dependent variables span 1948II–72IV.

cross-equation restrictions imposed by the model. In the one case shown (third set) in which the likelihood ratio does not call very strongly for rejection, the estimate of \( B \) is highly implausible.\(^{25}\)

**VIII. Conclusions**

As Lucas has observed,\(^{26}\) if one's criterion of success is a high \( R^2 \), then macroeconometric theories that link flows (like consumption) to flows (like income) are bound to dominate theories that link flows to relative prices. And on that criterion, the permanent-income version of the consumption function has certainly been a success. How can we reconcile that success with the discouraging results in this paper? The answer lies partly

\(^{25}\) Once again we note that using the test statistic \( T(\log | V_r | - \log | V_r^*|) \) rather than \( (T - k)(\log | V_r | - \log | V_r^*|) \), where \( k \) is the number of parameters estimated in a typical equation under the unrestricted parameterization, biases the test against the null hypothesis. In our case, using \( (T - k) \) rather than \( T \) to define the test statistic would not materially affect our conclusions.

\(^{26}\) Personal communication.
in another observation of Lucas (1975), namely, that econometric imposition of the hypothesis of rational expectations eliminates the usual emptiness of distributed lag estimation and delivers strongly overidentified models. Another reason for my results is that, contrary to common practice, I have resisted the temptation to attribute to the workings of the consumption function the correlation between innovations in consumption and income. This, together with a serial correlation correction, was adopted by way of eliminating simultaneous equations bias.

While the results here are discouraging from the viewpoint of the simple permanent-income model of aggregate consumption, they perhaps raise
the possibility that an equilibrium model, which confronts consumers with relative prices rather than incomes, ought not to be dismissed as being so obviously econometrically dominated by the Keynesian consumption hypothesis.

Appendix

1. Construction of Data Not Seasonally Adjusted

The six series, \(Y_1, Y_2, C, I, \tilde{I},\) and \(\tilde{C},\) are generated from 14 series of quarterly data, 1947I–72IV.

The first 13 are from *The National Income and Product Accounts of the United States, 1929–74 Statistical Tables*, a supplement to the *Survey of Current Business*, published by the Department of Commerce/Bureau of Economic Analysis.

All of the series are quarterly totals, not seasonally adjusted, in current dollars, with the exception of capital consumption allowance, for which we have quarterly totals at annual rates, seasonally adjusted. The series are

| \(X_1\)  | GNP                  | Table | Line | Page |
|---------|----------------------|-------|------|------|
|         |                      | 1.22  | 1    | 62   |
| \(X_2\) | Consumption durables | 2.50  | 2    | 86   |
| \(X_3\) | Consumption nondurables | 2.50 | 6    | 86   |
| \(X_4\) | Consumption services | 2.50  | 12   | 86   |
| \(X_5\) | Investment nonresidential | 1.22 | 8    | 86   |
| \(X_6\) | Investment residential | 1.22 | 11   | 62   |
| \(X_7\) | Change in business inventories | 1.22 | 15   | 62   |
| \(X_8\) | Capital consumption allowances with capital consumption adjustment | 1.90 | 2    | 22   |
| \(X_9\) | State and local government receipts | 3.50 | 1    | 116  |
| \(X_{10}\) | Federal government receipts | 3.30 | 1    | 104  |
| \(X_{11}\) | Federal government transfers to persons | 3.30 | 11   | 104  |
| \(X_{12}\) | State and local transfers to persons | 3.50 | 9    | 116  |
| \(X_{13}\) | Federal government grants-in-aid | 3.30 | 13   | 104  |
| \(X_{14}\) | The last series was obtained from the National Bureau of Economic Research Data Bank. It consists of quarterly averages of the monthly Consumer Price Index, all items 1967 = 100. Source is the U.S. Department of Labor. |       |      |      |

The series are defined as follows: \(Y_1 = [(X_1 - X_8/4. - X_9 - X_{10} + X_{11} + X_{12} + X_{13})/X_{14}] \cdot 100,\)
\(Y_2 = C + I,\)
\(\tilde{C} = [(X_2 + X_3 + X_4)/X_{14}] \cdot 100,\)
\(\tilde{I} = I + (X_3/X_{14}) \cdot 100,\) and \(\tilde{C} = \tilde{C} - 0.9875(X_3/X_{14}) \cdot 100 + 0.025D_t,\) where \(D_t = 0.975(X_3/X_{14}) \cdot 100 + 0.95D_{t-1}\) and \(D_{47-1} = 215.\)

2. Construction of the Seasonally Adjusted Data

The series, \(Y_1, Y_2, C, I, \tilde{I},\) and \(\tilde{C},\) are generated from 14 series of quarterly data, 1947I–72IV. They are all located in the statistical tables supplement of the *National Income and Product Accounts*. All are seasonally adjusted in current or 1972 dollars. Those in current dollars are deflated by the GNP implicit price deflator.
The series are:

|   | Table | Line | Page |
|---|-------|------|------|
| 1. | GNP   | 1.5  | 1    | 14   |
| 2. | Consumption durables | 2.4  | 2    | 82   |
| 3. | Consumption nondurables | 2.4  | 6    | 82   |
| 4. | Consumption services  | 2.4  | 12   | 82   |
| 5. | Investment nonresidential | 1.1  | 8    | 2    |
| 6. | Investment residential | 1.1  | 11   | 2    |
| 7. | Change in business inventories | 1.1  | 15   | 2    |
| 8. | Capital consumption allowance | 1.9  | 2    | 22   |
| 9. | State and local government receipts | 3.4  | 1    | 110  |
| 10. | Federal government receipts | 3.2  | 1    | 98   |
| 11. | Federal government transfers to persons | 3.2  | 31   | 98   |
| 12. | State and local transfers to persons | 3.4  | 42   | 110  |
| 13. | Federal government grants-in-aid | 3.2  | 33   | 98   |
| 14. | GNP implicit price deflator | 7.1  | 1    | 264  |

References

Anderson, Theodore W. *The Statistical Analysis of Time Series*. New York: Wiley, 1971.

Bailey, Martin J. “The Optimal Full-Employment Surplus.” *J.P.E.* 80, no. 4 (July/August 1972): 649-61.

Bodkin, Ronald G., and Klein, L. R. “Nonlinear Estimation of Aggregate Production Functions.” *Rev. Econ. and Statis.* 49 (February 1967): 28-44.

Durbin, J. “Tests for Serial Correlation in Regression Analysis Based on the Periodogram of Least-Squares Residuals.” *Biometrika* 56 (March 1969): 1-15.

Friedman, Milton. *A Theory of the Consumption Function: A Study by the National Bureau of Economic Research, New York*. Princeton, N.J.: Princeton Univ. Press, 1957.

Granger, C. W. J. “Investigating Causal Relations by Econometric Models and Cross-spectral Methods.” *Economietica* 37 (July 1969): 424-38.

Haavelmo, Trygve. “The Statistical Implications of a System of Simultaneous Equations.” *Economietica* 11 (January 1943): 1-12.

Hall, Robert E. “Stochastic Implications of the Life Cycle—Permanent Income Hypothesis: Theory and Evidence.” Mimeographed. Dept. Econ., Massachusetts Inst. Tech., July 1977.

Henderson, D. W., and Sargent, Thomas J. “Monetary and Fiscal Policy in a Two-Sector Aggregative Model.” *A.E.R.* 63 (June 1973): 345-65.

Lucas, Robert E., Jr. “Capacity, Overtime, and Empirical Production Functions.” *A.E.R. Papers and Proc.* 60 (May 1970): 23-27.

Lucas, Robert E., Jr. “An Equilibrium Model of the Business Cycle.” *J.P.E.* 83, no. 6 (December 1975): 1113-44.

Lucas, Robert E., Jr., and Rapping, Leonard. “Real Wages, Employment, and Inflation.” *J.P.E.* 77, no. 5 (September/October 1969): 721-54.

Merton, Robert C. “Optimum Consumption and Portfolio Rules in a Continuous Time Model.” *J. Econ. Theory* 3 (December 1971): 373-413.
Muth, John F. “Optimal Properties of Exponentially Weighted Forecasts.” *J. American Statis. Assoc.* 55 (June 1960): 299–306.

———. “Rational Expectations and the Theory of Price Movements.” *Econometrica* 29 (July 1961): 315–35.

Sargent, Thomas J. “Rational Expectations, the Real Rate of Interest, and the Natural Rate of Unemployment.” In *Brookings Papers on Economic Activity*. Vol. 2. Edited by Arthur M. Okun and George Perry. Washington: Brookings Inst., 1973.

———. “Notes on Macroeconomic Theory.” Mimeographed. Univ. Minnesota, 1975.

———. “A Classical Macroeconometric Model for the United States.” *J.P.E.* 84, no. 2 (April 1976): 207–37.

Sims, Christopher A. “Money, Income, and Causality.” *A.E.R.* 62 (September 1972): 540–52.

———. “Seasonality in Regression.” *J. American Statis. Assoc.* 69 (September 1974): 618–26.

———. “Response to Sargent’s Comment on Wallis’s Paper.” Mimeographed. Univ. Minnesota, November 1976.

Tobin, James. “A Dynamic Aggregative Model.” *J.P.E.* 63, no. 2 (April 1955): 103–15.

Wallis, Kenneth F. “Seasonal Adjustment and Relations between Variables.” *J. American Statis. Assoc.* 69 (March 1974): 18–31.

Whittle, Peter. *Prediction and Regulation by Linear Least Square Methods*. London: English Universities Press, 1963.

Wilson, G. T. “The Estimation of Parameters in Multivariate Time Series Models.” *J. Royal Statis. Soc. B* 35, no. 1 (1973): 76–85.