Two Models of Measurements and the Investment Accelerator

Author(s): Thomas J. Sargent

Source: *Journal of Political Economy* , Apr., 1989, Vol. 97, No. 2 (Apr., 1989), pp. 251-287

Published by: The University of Chicago Press

Stable URL: http://www.jstor.com/stable/1831313

# Two Models of Measurements and the Investment Accelerator

Thomas J. Sargent

*Hoover Institution and Federal Reserve Bank of Minneapolis*

This paper describes two models of an agency that is collecting and reporting observations on a dynamical linear stochastic economy. The first is a "classical" model, with the agency reporting data that are the sum of a vector of "true" variables and a vector of measurement errors that are orthogonal to the true variables. The second is a model of an agency that uses an optimal filtering method to construct least-squares estimates of the true variables. These two models of the reporting agency imply different likelihood functions. A model of the investment accelerator is used as an example to illustrate the differing implications of the models.

## I. Introduction

"Rational expectations econometrics" aims to interpret economic time series in terms of objects that are meaningful to economists, namely, parameters describing preferences, technologies, information sets, endowments, and equilibrium concepts or coordination mechanisms. When fully worked out, rational expectations models of the equilibrium type typically deliver a well-defined mapping from these economically interpretable parameters to the moments of the time series determined by the model. If accurate observations on these time series are available, one can use that mapping to implement parameter estimation methods based either on the likelihood function or on the method of moments. However, if only error-ridden data exist for the variables of interest, then more steps are needed to extract param-

I received valuable comments on an earlier draft of this paper from Lawrence Christiano and Kenneth West. This research was supported by National Science Foundation grant SES-85-08935 to the University of Minnesota.

[*Journal of Political Economy*, 1989, vol. 97, no. 2]

eter estimates. In effect, we require a model of the data reporting agency, one that is workable enough that we can determine the mapping induced jointly by the dynamic economic model and the measurement process to the probability law for the measured data. The model chosen for the data collection agency is an aspect of an econometric specification that can make big differences in inferences about the economic structure.

This paper describes two alternative models of data generation. Although the general ideas are more widely applicable, I will adopt setups that permit the application of the powerful computational methods afforded by the theory of linear stochastic systems. Thus I assume that the dynamic economic process is described by an economic theory that can be represented in the form of a stochastic process with a linear representation. I then describe two alternative linear models linking the data reported by the agency to the "true" data turned out by the economic system. The first model is a version of a classical "errors-in-variables" model. The second model is created by assuming that, although it collects error-ridden data that satisfy a classical errors-in-variables model, the data collecting agency filters the data and reports the best estimates that it possibly can. Although the two models have the same "deep parameters," they produce quite different sets of restrictions on the data.

After describing the two models and showing how to compute a likelihood function for each, I illustrate the differing properties of the two models by studying a puzzle for which at least one of the two models seems potentially to offer an explanation. In particular, I use the models to take up an old suggestion of Sims (1972*a*) that some restrictions on a measurement error model might help to explain the empirical success of accelerator theories of factor demand, a success that is not predicted by any of a variety of theories of equilibrium investment under uncertainty. By applying the two models of measurement in this particular context, I provide a concrete illustration of how different the implications of the two models can be.

## II. Two Models of Measurements: Classical Formulation

This section sets out a classical (i.e., nonrecursive) formulation of two alternative models of measurement error.[1] The idea is that an economic theory determines a stationary stochastic process for some true

[1] Whittle (1983) and Anderson and Moore (1979) are useful references on classical and recursive approaches to the signal extraction problem. Section II describes the classical approach used throughout most of Whittle, while Sec. III uses the recursive approach described by Anderson and Moore.

variables of interest to an economist. The economist (though not an agent within the model) is dependent for data on a reporting agency that observes only error-ridden versions of the data.[2] In particular, I assume that the agency's measurements are described by a classical measurement error model, being the sum of the true variables and an orthogonal process of possibly serially correlated measurement errors. One possible model of the behavior of the measurement agency is that it simply reports the error-ridden data that it collects, making no attempt to adjust for measurement errors. Provided that the economic model is tightly enough restricted, say by cross-equation restrictions of the rational expectations variety, and provided that the measurement error process is sufficiently restricted, this model can be estimated econometrically. The free parameters of the model will include parameters of the economic model and of the measurement error processes.

An alternative model of the reporting agency is that it reports the linear projection of the true data on the record of noise-ridden data that it has. In particular, the agency optimally filters the data in light of its knowledge of the correct economic model and the measurement error process, and it reports the best estimates of the true data that it possibly can.

In other words, knowing the DGP, you can try to est. the noise and filter it out.

These two models of the reporting agency require different strategies for estimating parameters, particularly when the measured error variances are sizable. I make this point precise by studying the ramifications of the two alternative models for maximum likelihood estimation.

### A. *The Economic Model*

There is an economic model whose equilibrium can be represented as a covariance stationary stochastic process with Wold moving average

[2] The economic agents in the model are *not* dependent on the data reported by the measurement agency. This simplifies the setup by avoiding some of the computational and conceptual issues that occur in Townsend (1983). The setup in this paper differs from those in Townsend, in which both the economist and the agents in his models must solve either the same or else very similar signal extraction problems. The data available to the economist and the agents in Townsend's model are not error-corrupted, but they are partly driven by hidden state variables. The "economic models" of the present paper, which appear in the form of either eq. (1) or eq. (10), can represent an environment in which some agents are solving self-contained control and signal extraction problems that do not interact with the signal extraction problem faced by the measurement agency. The models can also accommodate any of the variety of linear rational expectations models described by Hansen and Sargent (1980, 1981) and Whiteman (1983) in which agents solve control problems with complete knowledge of current values of the relevant state variables.

representation

$$\mathbf{Z}_t = c_{\mathbf{Z}}(L)\boldsymbol{\epsilon}_{\mathbf{Z}t}, \tag{1}$$

where $\mathbf{Z}_t$ is an $n \times 1$ vector of variables,

$$c_{\mathbf{Z}}(L) = \sum_{j=0}^{\infty} c_{\mathbf{Z}j}L^j, \quad \boldsymbol{\epsilon}_{\mathbf{Z}t} = \mathbf{Z}_t - E\mathbf{Z}_t|\mathbf{Z}_{t-1}, \ldots ,$$

and $E\boldsymbol{\epsilon}_{\mathbf{Z}t}\boldsymbol{\epsilon}'_{\mathbf{Z}t} = \mathbf{I}$. The parameters of the moving average representation, $c_{\mathbf{Z}}(L)$, are functions of the deeper parameters of the model that describe the preferences, opportunities, and coordination mechanisms available to agents.

Representation (1) can be used to compute the likelihood of a sample of observations on $\{\mathbf{Z}_t\}$ as a function of the parameters of the model. If observations on $\mathbf{Z}_t$ are available, the parameters of the model can be estimated by maximizing the likelihood function. Hansen and Sargent (1980, 1981) describe methods for computing versions of representation (1) as a function of deep economic parameters and show how (1) can be used in evaluating the likelihood function.

### B. *A Classical Errors-in-Variables Model*

There is a data collecting agency that observes an error-corrupted version of $\mathbf{Z}_t$, namely,

$$\mathbf{z}_t = \mathbf{Z}_t + \mathbf{v}_t, \tag{2}$$

where the measurement error $\mathbf{v}_t$ satisfies $E\boldsymbol{\epsilon}_{\mathbf{Z}t}\mathbf{v}'_s = 0$ for all $t$ and $s$ and itself has Wold moving average representation

$$\mathbf{v}_t = c_{\mathbf{v}}(L)\boldsymbol{\epsilon}_{\mathbf{v}t}, \tag{3}$$

where $c_{\mathbf{v}}(L) = \Sigma_{j=0}^{\infty} c_{\mathbf{v}j}L^j$, $\boldsymbol{\epsilon}_{\mathbf{v}t} = \mathbf{v}_t - E\mathbf{v}_t|\mathbf{v}_{t-1}, \ldots$, and $E\boldsymbol{\epsilon}_{\mathbf{v}t}\boldsymbol{\epsilon}'_{\mathbf{v}t} = \mathbf{I}$. Equations (2) and (3) imply that the error-corrupted data have a Wold moving average representation given by

$$\mathbf{z}_t = c_{\mathbf{z}}(L)\boldsymbol{\epsilon}_{\mathbf{z}t}, \tag{4}$$

where

$$c_{\mathbf{z}}(L) = \sum_{j=0}^{\infty} c_{\mathbf{z}j}L^j, \quad \boldsymbol{\epsilon}_{\mathbf{z}t} = \mathbf{z}_t - E\mathbf{z}_t|\mathbf{z}_{t-1}, \ldots ,$$

and $E\boldsymbol{\epsilon}_{\mathbf{z}t}\boldsymbol{\epsilon}'_{\mathbf{z}t} = \mathbf{I}$. The parameters $c_{\mathbf{z}}(L)$ of (4) are related to those of (1) and (3) by the spectral factorization identity

$$c_{\mathbf{z}}(s)c_{\mathbf{z}}(s^{-1})' = c_{\mathbf{Z}}(s)c_{\mathbf{Z}}(s^{-1})' + c_{\mathbf{v}}(s)c_{\mathbf{v}}(s^{-1})', \tag{5}$$

which must satisfy the side condition that the zeros of det $c_{\mathbf{z}}(s)$ are all outside the unit circle. Equation (5) tells how the parameters of the measurement error process (3) and of the economic model (1) interact to determine the Wold representation for the error-ridden data $\mathbf{z}_t$.

Equations (1)–(5) form a model of data reported by an agency that simply reports the error-ridden data that it collects. In particular, representation (4) and equation (5) can be used to form an approximate Gaussian likelihood function for a sample of observations on $\mathbf{z}_t$. Via (5) the likelihood is a function of the parameters of the economic model (1) and the measurement error model (3).

### C. *A Reporting Agency That Filters*

Consider an agency that observes $\{\mathbf{z}_t, \mathbf{z}_{t-1}, \ldots\}$ but that reports $\tilde{\mathbf{z}}_t = E\mathbf{Z}_t|\mathbf{z}_t, \mathbf{z}_{t-1}, \ldots$. It is assumed that the agency has rational expectations, the least-squares projection $E\mathbf{Z}_t|\mathbf{z}_t, \mathbf{z}_{t-1}, \ldots$ being computed with the correct probability model for the joint $(\mathbf{Z}_t, \mathbf{z}_t)$ stochastic process. From the classical optimal filtering formula (see Whittle 1983, p. 100), we have that

$$\tilde{\mathbf{z}}_t = h(L)\mathbf{z}_t, \tag{6}$$

where

$$h(L) = [c_{\mathbf{Z}}(L)c_{\mathbf{Z}}(L^{-1})'[c_{\mathbf{z}}(L^{-1})']^{-1}]_{+}c_{\mathbf{z}}(L)^{-1}. \tag{7}$$

Here $[\ ]_+$ is the annihilation operator defined by

$$\left[\sum_{j=-\infty}^{\infty} a_j L^j\right]_+ = \left[\sum_{j=0}^{\infty} a_j L^j\right].$$

A Wold moving average representation for $\tilde{\mathbf{z}}_t$ is given by

$$\tilde{\mathbf{z}}_t = c_{\tilde{\mathbf{z}}}(L)\boldsymbol{\epsilon}_{\tilde{\mathbf{z}}t}, \tag{8}$$

where

$$c_{\tilde{\mathbf{z}}}(L) = \sum_{j=0}^{\infty} c_{\tilde{\mathbf{z}}j}L^j, \quad \boldsymbol{\epsilon}_{\tilde{\mathbf{z}}t} = \tilde{\mathbf{z}}_t - E\tilde{\mathbf{z}}_t|\tilde{\mathbf{z}}_{t-1}, \ldots,$$

and $E\boldsymbol{\epsilon}_{\tilde{\mathbf{z}}t}\boldsymbol{\epsilon}'_{\tilde{\mathbf{z}}t} = \mathbf{I}$. The operator $c_{\tilde{\mathbf{z}}}(L)$ satisfies the spectral factorization identity

$$c_{\tilde{\mathbf{z}}}(s)c_{\tilde{\mathbf{z}}}(s^{-1})' = h(s)c_{\mathbf{z}}(s)c_{\mathbf{z}}(s^{-1})'h(s^{-1})', \tag{9}$$

where $c_{\tilde{\mathbf{z}}}(s)$ must satisfy the side condition that the zeros of det $c_{\tilde{\mathbf{z}}}(s)$ are all outside the unit circle.

Equations (7), (8), and (9) determine a Gaussian likelihood function for a sample of observations on $\tilde{\mathbf{z}}_t$.

This model provides a possible way of interpreting two features of the data-reporting process: (*a*) *Seasonal adjustment:*[3] To motivate this model, assume that the components of $\mathbf{v}_t$ have strong seasonals, the diagonals of the spectral density matrix $c_{\mathbf{v}}(e^{-iw})c_{\mathbf{v}}(e^{+iw})'$ having peaks at the seasonal frequencies. Then $h(L)$ in formulas (6)–(7) will assume a shape that can be interpreted partly in terms of a seasonal adjustment filter. Notice that this filter is one-sided in current and past $\mathbf{z}_t$'s. Two-sided filters are implemented only when preliminary data are revised. (*b*) *Data revisions:* Suppose that we specify our model so that $\mathbf{Z}_t$ contains current and lagged values of some variable of interest, say a capital stock, $k_t$. For example, let $k_t$, $k_{t-1}$, and $k_{t-2}$ each be components of $\mathbf{Z}_t$. Then (6), (7), and (8) become a model that simultaneously determines a "preliminary" estimate of $k_t$ (the component $Ek_t|\mathbf{z}_t, \mathbf{z}_{t-1}, \ldots$), a "revised" estimate of last period's capital stock (the component $Ek_{t-1}|\mathbf{z}_t, \mathbf{z}_{t-1}, \ldots$), and a "final" estimate of capital two periods ago (the component $Ek_{t-2}|\mathbf{z}_t, \mathbf{z}_{t-1}, \ldots$). According to this model, efficient parameter estimates would be obtained by using all the data, including each of the preliminary, revised, and final estimates, to form the likelihood function induced by (7), (8), and (9).[4]

## D. *Computational Issues*

To implement maximum likelihood estimation via time-domain procedures, the principal challenges are to achieve rapidly the spectral factorization (5), the calculation of $h(L)$ via (7), and the spectral factorization (9).[5] These calculations are difficult with classical methods but are routine if recursive methods are used. In the next section, I describe how to apply recursive methods.

[3] This model provides the most attractive rationalization of seasonal adjustment procedures that I have encountered. Nevertheless, for me it is a weak rationalization. It is difficult to suppose that the main source of seasonality in the measured data is seasonality in measurement errors. See Sargent (1987, pp. 336–43) for an account of some of the hazards of using seasonally adjusted data to estimate econometric models in settings in which the agents being modeled make decisions on the basis of raw seasonally unadjusted data.

[4] The recursive computational method of Sec. III accommodates the simultaneous inclusion of "preliminary," "revised," and "final" estimates in the data set $\{\tilde{\mathbf{z}}_t\}$ used to construct the likelihood function. Mankiw, Runkle, and Shapiro (1984) describe and test some orthogonality implications of measurement models like the ones considered here. Their characterizations have the virtue that they are simple and require only computing linear regressions to implement tests. Their empirical results indicate that money supply revisions are more consistent with the first model of measurements than the second. However, the empirical results of Mankiw and Shapiro (1986) indicate that GNP revisions are more consistent with the second model.

[5] By using frequency domain approximations to the likelihood function, model 1 can be estimated without achieving factorization (5) (see Hansen and Sargent 1980, 1981). For model 2, the factorization of at least (5) is needed to compute $h(L)$ via (7).

## III. Two Models of Measurement: A Recursive Formulation

This section reformulates and specializes the setup of the previous section so that recursive methods can be applied. In particular, we now work with economic models and measurement error processes, which each have finite-dimensional state vectors. This permits us to formulate the models in state-space forms to which the Kalman filter is applicable.

### A. *The Economic Model*

The economic model takes the state-space form

$$\begin{aligned} \mathbf{x}_{t+1} &= \mathbf{A}\mathbf{x}_t + \boldsymbol{\epsilon}_t, \\ \mathbf{Z}_t &= \mathbf{C}\mathbf{x}_t, \\ E\boldsymbol{\epsilon}_t\boldsymbol{\epsilon}_s' &= \begin{cases} \mathbf{Q} & \text{for } s = t \\ 0 & \text{for } s \neq t, \end{cases} \\ E\boldsymbol{\epsilon}_t &= 0 \quad \text{for all } t. \end{aligned} \tag{10}$$

Here $\mathbf{x}_t$ is an $n \times 1$ state vector, $\boldsymbol{\epsilon}_t$ is an $n \times 1$ vector white noise, and $\mathbf{Z}_t$ is an $m \times 1$ vector of variables related to $\mathbf{x}_t$. Observers of the economy would like to measure some of the variables in $\mathbf{Z}_t$. The $n \times n$ transition matrix $\mathbf{A}$ and the $m \times n$ matrix $\mathbf{C}$ have elements that are typically functions of deeper parameters describing the structure of preferences, technology, endowments, and rules of the game that define the particular model at hand. Assume that the eigenvalues of $\mathbf{A}$ are less than unity in modulus. The variety of discrete-time linear rational expectations models studied by Hansen and Sargent (1980, 1981) and Whiteman (1983) have equilibria that can be represented in the form of (10).

It is also possible to use (10) to represent the equilibria of continuous-time rational expectation models of the class described by Hansen and Sargent (1981). Harvey and Stock (1985, 1987) describe how to map continuous-time models of the class studied by Hansen and Sargent into the form (10).

### B. *A Classical Model of Measurements Initially Collected by an Agency*

There is an $m \times 1$ vector of measurement errors $\mathbf{v}_t$ that obeys

$$\begin{aligned} \mathbf{v}_{t+1} &= \mathbf{D}\mathbf{v}_t + \boldsymbol{\nu}_t,\ E\boldsymbol{\nu}_t\boldsymbol{\nu}_s' = \begin{cases} \mathbf{R} & \text{for } s = t \\ 0 & \text{for } s \neq t, \end{cases} \\ E\boldsymbol{\nu}_t &= 0 \quad \text{for all } t,\ E\boldsymbol{\epsilon}_t\boldsymbol{\nu}_s' = 0 \quad \text{for all } t \text{ and } s. \end{aligned} \tag{11}$$

The eigenvalues of the $m \times m$ matrix $\mathbf{D}$ are assumed to be bounded in modulus by unity. A data collecting agency is assumed to observe a noise-corrupted version of $\mathbf{Z}_t$, namely $\mathbf{z}_t$ governed by

$$\mathbf{z}_t = \mathbf{C}\mathbf{x}_t + \mathbf{v}_t. \tag{12}$$

One possible model is that the data collecting agency simply reports $\mathbf{z}_t$.

To represent its second moments, it is convenient to obtain the (population) vector autoregression of a $\mathbf{z}_t$ process generated by (10), (11), and (12). The error vector in the vector autoregression is the innovation to $\mathbf{z}_t$ and can be taken to be the white noise in a Wold moving average representation. A Wold moving average representation can be obtained by "inverting" the autoregressive representation. The population vector autoregression, and how it depends on the parameters of (10), (11), and (12), carries insights about how to interpret estimated vector autoregressions for $\mathbf{z}_t$. Constructing the vector autoregression is also useful as an intermediate step in computing the likelihood of a sample of $\mathbf{z}_t$'s as a function of the free parameters of $\{\mathbf{A}, \mathbf{C}, \mathbf{D}, \mathbf{Q}, \mathbf{R}\}$. The particular method that will be used to construct the vector autoregressive representation also proves useful as an intermediate step in constructing a model of an optimal reporting agency.

We shall use recursive (Kalman filtering) methods to obtain the vector autoregression for $\mathbf{z}_t$. Define[6]

$$\begin{aligned} \bar{\mathbf{z}}_t &= \mathbf{z}_{t+1} - \mathbf{D}\mathbf{z}_t, \\ \bar{\boldsymbol{\nu}}_t &= \boldsymbol{\nu}_t + \mathbf{C}\boldsymbol{\epsilon}_t, \\ \bar{\mathbf{C}} &= (\mathbf{C}\mathbf{A} - \mathbf{D}\mathbf{C}). \end{aligned} \tag{13}$$

Then (10), (11), and (12) imply the state-space system[7]

$$\begin{aligned} \mathbf{x}_{t+1} &= \mathbf{A}\mathbf{x}_t + \boldsymbol{\epsilon}_t, \\ \bar{\mathbf{z}}_t &= \bar{\mathbf{C}}\mathbf{x}_t + \bar{\boldsymbol{\nu}}_t, \end{aligned} \tag{14}$$

[6] The procedure for constructing a Wold representation for the model with vector first-order serially correlated measurement errors simply applies results described in Anderson and Moore (1979, pp. 290–92). We first apply their procedure for transforming a system with first-order serially correlated measurement errors to one with white-noise measurement errors. Then we use the Kalman filter to obtain an innovations representation for the observables, thereby obtaining the desired Wold representation. The procedures for forming projections of the hidden state variables and the measurement errors on the history of observed variables are standard (see Anderson and Moore 1979). The method for constructing the likelihood function using the innovations representation is also standard (see Harvey 1981; Ljung and Söderström 1983).

[7] An alternative way of incorporating serially correlated measurement errors is available that avoids the transformation of the data leading to (14). This alternative represents the measurement errors in each series as the sum of an autoregressive process and an orthogonal white noise. The autoregressive component of the measurement error process is included in the state vector in the standard way, permitting the system to be

where $(\boldsymbol{\epsilon}_t, \bar{\boldsymbol{\nu}}_t)$ is a white-noise process with covariance matrix

$$E\begin{bmatrix}\boldsymbol{\epsilon}_t\\ \bar{\boldsymbol{\nu}}_t\end{bmatrix}\begin{bmatrix}\boldsymbol{\epsilon}_t\\ \bar{\boldsymbol{\nu}}_t\end{bmatrix}' = \begin{bmatrix}\mathbf{Q} & \mathbf{QC}'\\ \mathbf{CQ}' & \mathbf{R} + \mathbf{CQC}'\end{bmatrix} \equiv \begin{bmatrix}\mathbf{Q}_1 & \mathbf{W}_1\\ \mathbf{W}_1' & \mathbf{R}_1\end{bmatrix}. \tag{15}$$

System (14) and (15) is characterized by the five matrices $[\mathbf{A}, \bar{\mathbf{C}}, \mathbf{Q}_1, \mathbf{R}_1, \mathbf{W}_1]$. Associated with (14) and (15) is the innovations representation for $\bar{\mathbf{z}}_t$,

$$\begin{aligned}\hat{\mathbf{x}}_{t+1} &= \mathbf{A}\hat{\mathbf{x}}_t + \mathbf{K}_1\mathbf{u}_t,\\ \bar{\mathbf{z}}_t &= \bar{\mathbf{C}}\hat{\mathbf{x}}_t + \mathbf{u}_t,\end{aligned} \tag{16}$$

where

$$\begin{aligned}\hat{\mathbf{x}}_t &= E[\mathbf{x}_t|\bar{\mathbf{z}}_{t-1},\bar{\mathbf{z}}_{t-2}, \ldots, \bar{\mathbf{z}}_0, \hat{\mathbf{x}}_0]\\ &= E[\mathbf{x}_t|\mathbf{z}_t, \mathbf{z}_{t-1}, \ldots, \mathbf{z}_0, \hat{\mathbf{x}}_0],\\ \mathbf{u}_t &= \bar{\mathbf{z}}_t - E[\bar{\mathbf{z}}_t|\bar{\mathbf{z}}_{t-1}, \bar{\mathbf{z}}_{t-2}, \ldots]\\ &= \mathbf{z}_{t+1} - E[\mathbf{z}_{t+1}|\mathbf{z}_t, \mathbf{z}_{t-1}, \ldots],\\ [\mathbf{K}_1, \mathbf{S}_1] &= \text{kfilter}(\mathbf{A}, \bar{\mathbf{C}}, \mathbf{Q}_1, \mathbf{R}_1, \mathbf{W}_1),\end{aligned} \tag{17}$$

where kfilter is the matrix-valued function defined in Appendix A and

$$\mathbf{S}_1 = E(\mathbf{x}_t - \hat{\mathbf{x}}_t)(\mathbf{x}_t - \hat{\mathbf{x}}_t)'. \tag{18}$$

From (17), $\mathbf{u}_t$ is the innovation process for the $\mathbf{z}_t$ process. System (16) and definition (13) can be used to obtain a Wold vector moving average representation for the $\mathbf{z}$ process:

$$\mathbf{z}_{t+1} = (\mathbf{I} - \mathbf{DL})^{-1}[\bar{\mathbf{C}}(\mathbf{I} - \mathbf{AL})^{-1}\mathbf{K}_1\mathbf{L} + \mathbf{I}]\mathbf{u}_t. \tag{19}$$

From (14) and (16) we have that

$$\begin{aligned}E\mathbf{u}_t\mathbf{u}_t' &= \bar{\mathbf{C}}E(\mathbf{x}_t - \hat{\mathbf{x}}_t)(\mathbf{x}_t - \hat{\mathbf{x}}_t)'\bar{\mathbf{C}}' + \mathbf{R}_1\\ &= \bar{\mathbf{C}}\mathbf{S}_1\bar{\mathbf{C}}' + \mathbf{R}_1\\ &\equiv \mathbf{V}_1.\end{aligned} \tag{20}$$

The Gaussian log likelihood function for a sample $\{\mathbf{z}_t, t = 0, \ldots, T\}$, conditioned on an initial state estimate $\hat{\mathbf{x}}_0$, can be represented as

$$L^* = -T\ln 2\pi - .5T\ln|\mathbf{V}_1| - .5\sum_{t=0}^{T-1}\mathbf{u}_t'\mathbf{V}_1^{-1}\mathbf{u}_t, \tag{21}$$

represented in state-space form with a vector white-noise measurement error process. While this procedure avoids the need to perform the transformations leading to (14), it leads to system matrices that are bigger but sparser. Appendix B describes how to alter these calculations to handle higher-order serially correlated measurement errors.

where $\mathbf{u}_t$ is the function of $\{\hat{\mathbf{x}}_0; \bar{\mathbf{z}}_{t-s}, s = 0, \ldots, T - 1\}$ induced by (16). To use (16) to compute $\{\mathbf{u}_t\}$, it is useful to represent it as

$$\begin{aligned} \hat{\mathbf{x}}_{t+1} &= (\mathbf{A} - \mathbf{K}_1\bar{\mathbf{C}})\hat{\mathbf{x}}_t + \mathbf{K}_1\bar{\mathbf{z}}_t, \\ \mathbf{u}_t &= -\bar{\mathbf{C}}\hat{\mathbf{x}}_t + \bar{\mathbf{z}}_t. \end{aligned} \tag{22}$$

Given $\hat{\mathbf{x}}_0$, (22) can be used recursively to compute a $\{\mathbf{u}_t\}$ process. Equations (21) and (22) give the likelihood function of a sample of error-corrupted data $\{\mathbf{z}_t\}$.

### C. *A Model of Optimal Estimates Reported by an Agency*

Suppose that instead of reporting the error-corrupted data $\mathbf{z}_t$, the data collection agency reports linear least-squares projections of the true data on a history of the error-corrupted data. To make this operational, we have to impute to the reporting agency a model of the joint process generating the true data and the measurement errors. Here we assume that "the reporting agency has rational expectations"; that is, the reporting agency knows the economic and measurement structure leading to (14)–(15).

To prepare its estimates, the reporting agency itself computes the Kalman filter to obtain the innovations representation (16). However, rather than reporting the error-corrupted data $\mathbf{z}_t$, the agency reports $\tilde{\mathbf{z}}_t = \mathbf{G}\hat{\mathbf{x}}_t$, where $\mathbf{G}$ is a "selection matrix," possibly equal to $\mathbf{C}$, for the data reported by the agency. The data $\mathbf{G}\hat{\mathbf{x}}_t = E[\mathbf{G}\mathbf{x}_t | \mathbf{z}_t, \mathbf{z}_{t-1}, \ldots, \mathbf{z}_0, \hat{\mathbf{x}}_0]$. The state-space representation for the reported data is then

$$\begin{aligned} \hat{\mathbf{x}}_{t+1} &= \mathbf{A}\hat{\mathbf{x}}_t + \mathbf{K}_1\mathbf{u}_t, \\ \tilde{\mathbf{z}}_t &= \mathbf{G}\hat{\mathbf{x}}_t, \end{aligned} \tag{23}$$

where the first line of (23) is from representation (16). Note that $\mathbf{u}_t$ is the innovation to $\mathbf{z}_{t+1}$ and is *not* the innovation to $\tilde{\mathbf{z}}_t$.[8]

To obtain a Wold representation for $\tilde{\mathbf{z}}_t$ and the likelihood function for a sample of $\tilde{\mathbf{z}}_t$ requires that we obtain an innovations representation for (23). To add a little generality to (23) we amend it to the system

$$\begin{aligned} \hat{\mathbf{x}}_{t+1} &= \mathbf{A}\hat{\mathbf{x}}_t + \mathbf{K}_1\mathbf{u}_t, \\ \tilde{\mathbf{z}}_t &= \mathbf{G}\hat{\mathbf{x}}_t + \boldsymbol{\eta}_t, \end{aligned} \tag{24}$$

[8] One can readily specify the $\mathbf{A}$ and $\mathbf{G}$ matrices and the vectors $\mathbf{x}_t$, $\hat{\mathbf{x}}_t$, and $\tilde{\mathbf{z}}_t$ to accommodate the model of data revision in Sec. II. For example, in the model of Sec. IV, one could augment the state vector to be $[k_t, \theta_t, k_{t-1}, k_{t-2}]$ and augment $\mathbf{G}$ to pick off estimates of $k_{t-1}$ and $k_{t-2}$ based on error-ridden observations through time $t$.

where $\boldsymbol{\eta}_t$ is a type 2 white-noise measurement error process ("typos") with presumably a very small covariance matrix. We have the covariance matrix

$$E\begin{bmatrix}\mathbf{K}_1\mathbf{u}_t\\ \boldsymbol{\eta}_t\end{bmatrix}\begin{bmatrix}\mathbf{K}_1\mathbf{u}_t\\ \boldsymbol{\eta}_t\end{bmatrix}' = \begin{bmatrix}\mathbf{K}_1\mathbf{V}_1\mathbf{K}_1' & \mathbf{0}\\ \mathbf{0} & \mathbf{R}_2\end{bmatrix} = \begin{bmatrix}\mathbf{Q}_2 & \mathbf{0}\\ \mathbf{0} & \mathbf{R}_2\end{bmatrix}, \tag{25}$$

where $\mathbf{R}_2$ is possibly close or equal to zero.[9] For system (24) and (25), an innovations representation is

$$\begin{aligned}\hat{\hat{\mathbf{x}}}_{t+1} &= \mathbf{A}\hat{\hat{\mathbf{x}}}_t + \mathbf{K}_2\mathbf{a}_t,\\ \tilde{\mathbf{z}}_t &= \mathbf{G}\hat{\hat{\mathbf{x}}}_t + \mathbf{a}_t,\end{aligned} \tag{26}$$

where

$$\begin{aligned}\mathbf{a}_t &= \tilde{\mathbf{z}}_t - E[\tilde{\mathbf{z}}_t|\tilde{\mathbf{z}}_{t-1}, \tilde{\mathbf{z}}_{t-2}, \ldots],\\ \hat{\hat{\mathbf{x}}}_t &= E[\hat{\mathbf{x}}_t|\tilde{\mathbf{z}}_{t-1}, \tilde{\mathbf{z}}_{t-2}, \ldots, \tilde{\mathbf{z}}_0, \hat{\hat{\mathbf{x}}}_0],\\ \mathbf{S}_2 &= E(\hat{\mathbf{x}}_t - \hat{\hat{\mathbf{x}}}_t)(\hat{\mathbf{x}}_t - \hat{\hat{\mathbf{x}}}_t)',\\ [\mathbf{K}_2, \mathbf{S}_2] &= \text{kfilter}(\mathbf{A}, \mathbf{G}, \mathbf{Q}_2, \mathbf{R}_2, \mathbf{0}),\end{aligned} \tag{27}$$

where kfilter is again the matrix-valued function defined in Appendix A. Thus $\{\mathbf{a}_t\}$ is the innovation process for the reported data $\tilde{\mathbf{z}}_t$, with contemporaneous covariance matrix

$$E\mathbf{a}_t\mathbf{a}_t' = \mathbf{G}\mathbf{S}_2\mathbf{G}' + \mathbf{R}_2 \equiv \mathbf{V}_2. \tag{28}$$

A Wold moving average representation for $\tilde{\mathbf{z}}_t$ is found from (26) to be

$$\tilde{\mathbf{z}}_t = [\mathbf{G}(\mathbf{I} - \mathbf{A}L)^{-1}\mathbf{K}_2L + \mathbf{I}]\mathbf{a}_t. \tag{29}$$

When a method analogous to the one in the preceding section is used, a Gaussian log likelihood for $\tilde{\mathbf{z}}_t$ can be computed by first computing an $\{\mathbf{a}_t\}$ sequence from observations on $\tilde{\mathbf{z}}_t$ by using

$$\begin{aligned}\hat{\hat{\mathbf{x}}}_{t+1} &= (\mathbf{A} - \mathbf{K}_2\mathbf{G})\hat{\hat{\mathbf{x}}}_t + \mathbf{K}_2\tilde{\mathbf{z}}_t,\\ \mathbf{a}_t &= -\mathbf{G}\hat{\hat{\mathbf{x}}}_t + \tilde{\mathbf{z}}_t.\end{aligned} \tag{30}$$

Then the likelihood function for a sample of $T$ observations $\{\tilde{\mathbf{z}}_t\}$ is

$$L^{**} = -T\ln 2\pi - .5T\ln|\mathbf{V}_2| - .5\sum_{t=0}^{T-1}\mathbf{a}_t'\mathbf{V}_2^{-1}\mathbf{a}_t.$$

[9] If $\mathbf{R}_2$ is singular, it is necessary to adjust the Kalman filtering formulas by using transformations that induce a "reduced order observer." In practice, when we want to model a situation in which we think $\mathbf{R}_2$ is actually singular (say a matrix of zeros), we can continue to use the Kalman filter as described in the text if we simply use a positive definite $\mathbf{R}_2$ to approximate the singular one. This procedure is used in the calculations in Sec. IV below. In particular, we approximate a zero $\mathbf{R}_2$ matrix with the matrix $(.000001)\mathbf{I}$, where $\mathbf{I}$ is the identity matrix.

Note that relative to computing the likelihood function (21) for the error-corrupted data, computing the likelihood function for the optimally filtered data requires more calculations. Both likelihood functions require that the Kalman filter (17) be computed, while the likelihood function for the filtered data requires that the Kalman filter (27) also be computed. In effect, in order to interpret and use the filtered data reported by the agency, it is necessary to retrace the steps that the agency used to synthesize those data.[10]

## IV. Equilibrium Investment under Uncertainty, Measurement Errors, and the Investment Accelerator

This section uses the two models of measurement in an attempt to understand a puzzle about the behavior of aggregate investment. In particular, we shall try to reconcile the empirical success of the investment accelerator with equilibrium theories of investment under uncertainty. By an accelerator, I mean a one-sided projection of a rate of investment in a productive factor on current and lagged values of a measure of output. Existing equilibrium theories of investment under uncertainty do not generally lead a researcher to fit accelerator regressions. These theories also predict that accelerator relationships will not survive specification tests. Nevertheless, a desire for good statistical fits in explaining investment or employment has often driven empirical workers toward fitting investment or employment accelerators (see, e.g., Abel and Blanchard 1983). Further, accelerators have been subjected to specification tests, in the form of tests for Granger causality, by Sims (1972*a*) and Abel and Blanchard (1983), and have survived them. In particular, those studies found that a measure of shipments or output Granger-caused stocks of capital or employment but found little evidence for reverse Granger causality. Such findings contradict existing theories of equilibrium investment under uncertainty, of both the partial equilibrium variety (e.g., Lucas and Prescott 1971; Eichenbaum 1983, 1984; Sargent 1987) and the general equilibrium variety (e.g., Brock and Mirman 1972; Sargent 1980). Equilibrium theories of investment under uncertainty imply that stocks of factors of production, output, prices, and other variables are jointly determined stochastic processes and that each depends on the history of all the shocks in the system. This implies that projections of investment on output will in general be two-sided in past and present values.

[10] The Kalman filter (17) is supposed to be formed by the agency. The agency need not use Kalman filter (27) because it does not need the Wold representation for the filtered data.

This section uses a hint given by Sims (1972*a*, 1974) and the two models of measurement to explore this puzzle. The idea is to restrict the equilibrium model to have only one source of random disturbance, say the demand shock in a Lucas-Prescott model (1971) or the technology shock in a Brock-Mirman model (1972). For a linear model, the one-shock hypothesis makes the equilibrium for output, capital, and other variables (such as output in the Lucas-Prescott model or consumption in the Brock-Mirman model) a stochastic process whose spectral density is of rank one at all frequencies, each distinct variable being expressible as a distinct one-sided distributed lag of the *same* white noise. If each of those distributed lags was invertible and if all the variables were to be measured accurately, then no variable would Granger-cause any other, and the projection of one process on any other would be one-sided in present and past values. However, we posit that capital, output, and other variables are not measured accurately and that observations of each are polluted by serially correlated measurement errors. To begin with, we use our first model of the reporting agency, so that the agency simply reports the error-corrupted data. The measurement errors are posited to be ordered in variance, with output being relatively well measured and capital being relatively poorly measured. The error-ridden measured process then has the property that every variable Granger-causes every other variable but that the best measured variable (output) exhibits the least feedback from other variables and the greatest predictive power for other variables. This section illustrates how the parameters of an equilibrium model and the measurement error processes can be chosen to deliver a stochastic process for which the accelerator fits the error-corrupted data well and for which it is difficult to detect Granger causality from other variables to output in small samples. For the example under study, the empirical success of the accelerator reflects the error corruption of the data and the propensity of least-squares regressions to solve signal extraction problems, and it does not necessarily reflect poorly on equilibrium theories or help to confirm "accelerator theories" that make investment a distributed lag of income or output that is structural (invariant under interventions).[11]

[11] When time-series data are corrupted by measurement errors, least-squares procedures for fitting vector autoregressions to those data automatically solve signal extraction (or, equivalently, optimal prediction) problems. Vector autoregressions, ipso facto, estimate versions of Wold representations, their innovations being estimates of (and converging to) the fundamental vector white noise that appears in the Wold representation. That is, vector autoregressions recover versions of the Wold representation (4) or (19) for the error-ridden data. Hansen and Sargent (1982) describe two problems associated with interpreting Wold representations and the "innovation accountings" of Sims (1980) that are related to the issues raised in this section. Hansen and Sargent

After illustrating the implications of the first model of measurement error, I shall then illustrate how interpretations would change under the second model of measurement.

I shall illustrate these things by using a linear-quadratic version of a stochastic optimal growth model, a version of which is described in Sargent (1987, pp. 376–77). The model fits into the framework of Section III. With an eye to attaining an accelerator relationship, I restrict the model to have only one source of random shocks, an endowment shock.[12] If we had more sources of shocks, say a preference shock or a government expenditure shock, we would not be able to attain the accelerator as we do below. In order to obtain a sharp version of the accelerator, I also restrict the model in other ways that will be discussed below. My intent in imposing these restrictions is to explore what assumptions are needed to get the accelerator and to understand the senses in which the accelerator is very special in terms of equilibrium theories of investment.

A social planner maximizes

$$E \sum_{t=0}^{\infty} \beta^t \left( u_0 + u_1 c_t - \frac{u_2 c_t^2}{2} \right) \tag{31}$$

subject to the technology

$$c_t + k_{t+1} = f k_t + \theta_t, \quad \beta f^2 > 1, \tag{32}$$

$$a(L)\theta_t = \epsilon_t, \tag{33}$$

where $a(L) = 1 - a_1 L - a_2 L^2 - \ldots - a_r L^r$, and where $a(L)$ has the property that $a(z) = 0$ implies that $|z| > 1$. Here $c_t$ is consumption at $t$, $k_t$ is the capital stock at $t$, $E$ is the mathematical expectations operator, $\theta_t$ is an endowment or technology shock that follows an $r$th-order autoregressive process given by (33), and $\epsilon_t = \theta_t - E[\theta_t | \theta_{t-1}, \theta_{t-2} \ldots]$. The social planner's problem is to maximize (31) subject to (32) and (33) by choosing a "realizable" contingency plan for $\{c_t, k_{t+1}\}$ as a function of the information available at $t$ and subject to the boundedness condition $E_0 \Sigma_{t=0}^{\infty} \beta^t k_t^2 < \infty$. The initial capital stock $k_0$ is given, and at $t$ the information set available to the planner is $k_t, \theta_t, \theta_{t-1}, \ldots, \theta_{t-r}$.

---

describe two settings in which there is no measurement error in the usual sense but in which vector autoregressions on the data available to the economist and the innovations recovered by vector autoregressions contain less information than that available to the agents in the model. Those settings involve, respectively, an "invertibility problem" and "time aggregation." Hansen and Sargent describe methods for overcoming those problems that are closely related to ones described in the present paper.

[12] This model is an "unobservable index model" in the sense of Sargent and Sims (1977), with the modification that sufficient a priori structure is imposed on it to achieve identification.

The solution to this problem can be represented by the following optimal decision rule for $c_t$:

$$c_t = \frac{-\alpha}{f-1} + \left(1 - \frac{1}{\beta f^2}\right)\left[\frac{L - f^{-1}a(f^{-1})^{-1}a(L)}{L - f^{-1}}\,\theta_t + fk_t\right],$$

$$k_{t+1} = fk_t + \theta_t - c_t, \tag{34}$$

where $\alpha = u_1[1 - (f\beta)^{-1}]/u_2$. Equations (33) and (34) exhibit the cross-equation restrictions characteristic of rational expectations models.

In the interests of attaining an accelerator theory, we define net output or national income as

$$y_{nt} = (f-1)k_t + \theta_t. \tag{35}$$

Note that (32) and (35) imply that $(k_{t+1} - k_t) + c_t = y_{nt}$. We now make two special assumptions that will simultaneously deliver a version of Friedman's (1957) geometric distributed lag formulation of the consumption function and a distributed lag accelerator. We assume that (1) $a(L) = \mathbf{I}$ and (2) $\beta f = 1$. Assumption 1 means that the technology shock is a white noise, while assumption 2 states that the rate of return on capital equals the rate of time preference. Assumption 1 is crucial for some of our results, while assumption 2 is less so, affecting only the values of various constants.[13] Assumption 1 is used to deliver both the strict version of the distributed lag accelerator and Friedman's geometric distributed lag model. Relaxing assumption 1 would still leave consumption and investment expressible as one-sided distributed lags of income $y_{nt}$.

Under assumptions 1 and 2, (34) becomes

$$c_t = (1 - f^{-1})\theta_t + (f-1)k_t. \tag{36}$$

When (36), (35), and (32) are used, the optimal consumption, capital accumulation plan satisfies the following restrictions (see Sargent 1987, pp. 376–77):

$$c_t = \left(\frac{1-\beta}{1-\beta L}\right)y_{nt}, \tag{37}$$

$$k_{t+1} - k_t = f^{-1}\left(\frac{1-L}{1-\beta L}\right)y_{nt}, \tag{38}$$

$$y_{nt} = \theta_t + (1-\beta)(\theta_{t-1} + \theta_{t-2} + \theta_{t-3} + \ldots). \tag{39}$$

[13] The assumption that $\beta f = 1$ makes the solution (34) one for which $c_t$ follows a martingale, regardless of the value assumed by the operator $a(L)$. Hall (1978) described how such a martingale characterization for (the marginal utility of) consumption emerged from a version of problem (31)–(33).

Equation (37) is a version of Friedman's consumption model for time-series data. It expresses consumption as a geometric distributed lag of income. Further, equation (37) conforms to Friedman's (1963) conjecture that the geometric decay coefficient used in discounting the past equals the factor associated with discounting the future.[14] Equation (38) is a version of the distributed lag accelerator, expressing investment as a geometric distributed lag of the first difference of income. Equation (39) states that the first difference of disposable income is a first-order moving average process with innovation equal to the innovation of the endowment shock $\theta_t$. Thus $y_{nt}$ is a stochastic process that, as Muth (1960) showed, is optimally forecast via a geometric distributed lag or "adaptive expectations" scheme.

This setup produces an accelerator in the strict sense that investment is a distributed lag of the *change* in income. Before I move on to describe modifications of the model that will deliver weaker senses of an accelerator, note how the model above highlights the very special assumptions needed to attain (38). Assumption 1 is crucial in producing (38) and, in particular, in delivering the $(1 - L)$ term in the numerator of the right-hand side that characterizes (38) as an accelerator. As the cross-equation restrictions exhibited in (34) reveal, the consumption function and in turn the counterpart of the projection of $c_t$ on $\{y_{nt}, y_{nt-2}, \ldots\}$ given in (38) are functions of the parameters of the autoregressive process $a(L)$ for $\theta_t$. This dependence of the special form (38) on the special assumption that $a(L) = \mathbf{I}$ is important, in view of the widespread practice of positing serially correlated technology shocks (e.g., Kydland and Prescott 1982) and also because the model possesses a ready reinterpretation in which $\theta_t$ is actually the difference between a technology shock $\tau_t$ and a stochastic process for government purchases $g_t$, with $\theta_t = \tau_t - g_t$. The mathematics described above then requires that both $\tau_t$ and $g_t$ must be white noises in order for (38) to hold. Further, any projection of the form (38) will fail to be invariant with respect to hypothetical interventions that alter the predictable part of the stochastic process for $\{g_t\}$ (e.g., any intervention that changes the polynomial corresponding to $a(L)$ for $g_t$).

When assumption 1 is dropped so that $\theta_t$ is permitted to be serially correlated, an accelerator relationship obtains in a broader sense. As long as the equilibrium can be represented as expressing each variable as an invertible distributed lag of a single source of randomness, such as $\theta_t$, an accelerator relationship will characterize the model in the broad sense that investment can be represented as a one-sided

[14] Sargent (1987) uses a version of the model of the present section as a vehicle for studying Friedman's conjecture in more detail than that used in the present paper.

distributed lag of current and past income.[15] The projection of investment on income will be one-sided on past and present income and will equal investment (in the mean squared error sense). However, within this class of structures, the $z$-transform of the coefficients of the projection of investment on current and past income will in general fail to have a zero at $z = 1$, a special feature that (38) possesses because of the $(1 - L)$ term in the numerator. Further, within this broad class of models, there is no presumption that the projection of investment on current and past income can be well approximated by a *smooth* distributed lag on the change in income.[16]

Another delicacy associated with the derivation of (38) affects the attainment of both the strict and the broad form of the accelerator. In deriving (38), we have heavily exploited the fact that there is only a one-dimensional random shock, $\{\theta_t\}$, impinging on the planner's problem. If there are two or more sources of random shocks, the derivation above dissolves. There are two plausible ways that one might want to add additional sources of random shocks to model (31)–(33). First, in the notation of the previous paragraph, if $\theta_t$ in (32) is reinterpreted as the difference $\tau_t - g_t$, if both $\tau_t$ and $g_t$ are observed by the social planner, and if $\tau_t$ and $g_t$ follow stochastic processes with *different* univariate moving average representations, then a "one-shock" model fails to hold. Second, if there are random shocks to preferences (e.g., a term like $c_t w_t$, $w_t$ a random process, is added to the current-period return function in [31]), then a one-shock model will again fail.

The fact that the setup above leads to a one-dimensional-shock (or

[15] To obtain invertibility of the lag distribution in terms of $\theta_t$, it is sufficient to impose the side condition on $a(L)$ and $\beta$ that the polynomial $\{[L - \beta a(\beta)^{-1}a(L)]/(L - \beta)\}$ that operates on the right side of (34) has no zeros inside the unit circle. The invertibility condition assures that current and lagged values of the random process $\{[L - \beta a(\beta)^{-1}a(L)]/(L - \beta)\}\theta_t$ generate the same linear space as current and lagged values of $\theta_t$. If the condition were violated, then even though the spectral density of $(1 - L)c_t$, $(1 - L)k_{t+1}$, $(1 - L)y_{nt}$ is of rank one at all frequencies, the Wold representation for these variables has more than one white noise. This means that nontrivial Granger causality can extend from one variable to another. The invertibility of the operator $\{[L - \beta a(\beta)^{-1}a(L)]/(L - \beta)\}$ is discussed in Hansen and Sargent (1980, 1982) and Futia (1981). They show that invertibility of this operator imposes restrictions on $\beta$ and $a(L)$ beyond those implied by the stationarity of $\theta_t$ in (33).

[16] It always *can* be approximated arbitrarily well in the mean square metric by a sufficiently erratic distributed lag of the change in income. The apparatus of Sims (1972*c*) can be used to prove this claim. Let $y_{nt}$ be a covariance stationary stochastic process. Let $b$ and $a$ each be square summable, one-sided infinite sequences. Consider the process $b * y_{nt}$, where the asterisk denotes convolution. Then there exists a polynomial $a$ that is one-sided and square summable and possesses an inverse under convolution such that $a^{-1} * b * (y_{nt} - y_{nt-1})$ approximates $b * y_{nt}$ arbitrarily well. Here $a^{-1}$ denotes inverse under convolution. By working in terms of variables multiplied by $\beta^{-.5t}$ or by using a limiting argument, the argument above can be extended to handle the "borderline nonstationarity" exhibited by model (31), (32), and (33).

"singular") stochastic process for $\{c_t, k_{t+1} - k_t, y_{nt}\}$ is important when it comes to explaining the empirical success of the accelerator, in particular, the apparent Granger causal priority of output measures with respect to capital accumulation measures. This property of the system will hold for *any* specification of $a(L)$ that delivers an equilibrium with the property that $(1 - L)\{c_t, k_{t+1} - k_t, y_{nt}\}$ are each invertible distributed lags of the single shock $\theta_t$. Such a one-innovation representation for the equilibrium *is* a Wold representation. From such a "singular" Wold representation it can be calculated directly that no variable Granger-causes any other.[17]

Granger causality extending from one variable to another can be accommodated if measurement errors are added to the model. To illustrate the possibilities, suppose that (37), (38), and (39) are true. Suppose that output is measured without error but that consumption and capital are polluted by covariance stationary, serially correlated measurement errors, $v_{ct}$ and $v_{kt}$, respectively. We assume that $\{v_{ct}\}$ and $\{v_{kt}\}$ are mutually orthogonal stochastic processes and that each is orthogonal to the technology shock $\theta_t$. Let measured consumption and capital be denoted $\bar{c}_t$ and $\bar{k}_t$, respectively. Then adding $v_{ct}$ to the right side of (37) and $v_{kt+1} - v_{kt}$ to the right side of (38) gives the following model for the observables:

$$\bar{c}_t = \left(\frac{1 - \beta}{1 - \beta L}\right) y_{nt} + v_{ct}, \tag{40}$$

$$\bar{k}_{t+1} - \bar{k}_t = \beta\left(\frac{1 - L}{1 - \beta L}\right) y_{nt} + (v_{kt+1} - v_{kt}), \tag{41}$$

$$y_{nt} = \theta_t + (1 - \beta)(\theta_{t-1} + \theta_{t-2} + \ldots). \tag{42}$$

The model formed by (40), (41), and (42) is one in which the distributed lag accelerator survives Granger causality tests. Income Granger-causes, but is not Granger-caused by, both measured consumption and investment.[18] These statements about Granger causal-

[17] The proof of this assertion is a simple exercise in the application of Sims's (1972*b*) theorems. See Sargent (1987, chap. 11, exercise 8) for a version of this exercise.

[18] The claim that income Granger-causes investment is not true in the special case that $v_{kt+1} - v_{kt}$ is a white noise. Notice that (37), (38), and (39) imply the following moving average representation for the true variables:

$$(1 - L)c_t = (1 - \beta)\theta_t,$$

$$(1 - L)y_{nt} = (1 - \beta L)\theta_t,$$

$$k_{t+1} - k_t = \beta\theta_t.$$

Let $v_{it} = v_{kt+1} - v_{kt}$. We then have that $k_{t+1} - k_t = \beta\theta_t + v_{it}$. If $v_{it}$ is a white noise, then $\beta\theta_t + v_{it}$ is evidently a fundamental white noise for $k_{t+1} - k_t$. A triangular Wold representation with $\beta\theta_t + v_{it}$ as one of the white noises can readily be constructed, which via Sims's (1972*b*) two theorems verifies that no other observable Granger-causes ity can be verified by using (40), (41), and (42) to derive a Wold moving average representation for the vector $(\bar{c}_t, \bar{k}_{t+1} - \bar{k}_t, y_{nt})$. This can be done readily on noting that (42) is a univariate Wold representation for $y_{nt}$, that $v_{kt+1}$ is the part of the innovation in $\bar{k}_{t+1}$ that is orthogonal to the $\theta_t$ process, and that $v_{ct}$ is the part of the innovation in $\bar{c}_t$ that is orthogonal to the $\theta_t$ process. One can then immediately move from (40), (41), and (42) to a Wold moving average representation for $(\bar{c}_t, \bar{k}_{t+1} - \bar{k}_{t}, y_{nt})$ that is triangular in a way that confirms the preceding claims about Granger causality (see Sims 1972*b*).[19]

When $y_{nt}$ is also measured with error, so that each of $c_t$, $k_{t+1} - k_t$, and $y_{nt}$ is corrupted by measurement errors, each measured series will in general Granger-cause each of the others. The strength of Granger causality from one series to another, as measured by decompositions of $j$-step-ahead prediction error variances,[20] will depend on the relative variances of the measurement errors. In a one-common-index model like ours ($\theta_t$ is the common index), the best measured variables will extend more Granger causality to the less well measured series than vice versa.

To substantiate these claims in a context in which each of $c_t$, $k_{t+1} - k_t$, and $y_{nt}$ is measured with error requires that we deduce a Wold moving average representation for the measured series. We can use the recursive methods described in Section III to compute this representation. I now describe the results of these calculations for a specific numerical example of the model (31), (32), (33), and (34).

Parameters of the economic model and the measurement model are described in table 1. We assume that the reporting agency collects error-corrupted measurements on $y_{nt}$, $c_t$, and $(k_{t+1} - k_t)$. We assume that $\theta_t$ is a white noise. It is elementary to map the economic model and the measurements into the state-space representation of Section

---

$k_{t+1} - k_t$. More directly, note that from the representation $k_{t+1} - k_t = \beta\theta_t + v_{it}$, it follows that $E_{t-1}(k_{t+1} - k_t) = 0$, which establishes absence of Granger causality from other variables to investment.

[19] This construction covers the case in which the measurement error in consumption equals minus the measurement error in investment, so that the measured data on consumption and investment add up to measured income. In this special case, the covariance matrix of the innovation $\epsilon_{zt}$ in (4) has rank two, being equal to the sum of the number of "indexes" or signals (one, with the signal being $\theta_t$) and the number of linearly independent measurement errors (being one in this special case). Even though the measured process is singular in this special case, it remains true that measured income is not Granger-caused by any other variables in the system. The model (40), (41), and (42) can reconcile with Hall's (1978) model the "excess sensitivity of consumption to income" relative to that predicted by Hall's model that was observed by Flavin (1981, 1985). More generally, as the calculations below illustrate, the first model of the reported data, though not the second, is capable of reconciling Flavin's observations and Hall's theory under my assumptions about the ordering of variances of the measurement errors.

[20] See Sims's (1980) description of "innovation accountings."

TABLE 1

PARAMETERS OF MEASUREMENT ERROR MODELS

Economic model:

$$\beta = (1.05)^{-1},\ E\theta_t^2 = 1,\ \theta_t \text{ a white noise}$$

Measurement error process:

Measurement error to income ($y_n$):

$$v_{yt} = .6v_{yt-1} + \nu_{yt},\ \text{standard deviation } (v_{yt}) = .050$$

Measurement error to consumption ($c$):

$$v_{ct} = .7v_{ct-1} + \nu_{ct},\ \text{standard deviation } (v_{ct}) = .035$$

Measured error to investment ($k_{t+1} - k_t$):

$$v_{it} = .3v_{it-1} + \nu_{it},\ \text{standard deviation } (v_{it}) = .65$$

($\nu_{yt}$, $\nu_{ct}$, $\nu_{it}$) are mutually orthogonal white noises with variances chosen to yield standard deviations of $v_{yt}$, $v_{ct}$, and $v_{it}$ described above

($\nu_{yt}$, $\nu_{ct}$, $\nu_{it}$) are orthogonal to $\theta_t$

III. The *state* of the system is $\mathbf{x}_t = (k_t, \theta_t)$. The true data $y_{nt}$, $c_t$, and $k_{t+1} - k_t$ are linear functions of the state.

Table 2 gives the impulse response function of the true data to a shock in the one noise in the economic model, $\theta_t$. The table faithfully reflects the consumption-smoothing property that is displayed analytically in note 15.

Table 3 reports a population decomposition of variance of the $j$-step-ahead prediction error variance associated with the Wold representation for the measured data. Using procedures that correspond to those described by Sims (1980), I have orthogonalized the innovations, with $y_{nt}$ going first, $c_t$ going second, and $k_{t+1} - k_t$ going third in the orthogonalization procedure induced by a Cholesky decomposition of the covariance matrix of the innovations. The covariance ma-

TABLE 2

IMPULSE RESPONSE OF $y_{nt}$, $c_t$, AND $k_{t+1} - k_t$ TO AN INNOVATION IN $\theta_t$ ($\beta = 1.05^{-1}$, $\beta f = 1$)

| Lag | $y_n$ | $c$ | $k_{+1} - k$ |
|---|---|---|---|
| 0 | 1.00 | .0476 | .9524 |
| 1 | .0476 | .0476 | 0 |
| 2 | .0476 | .0476 | 0 |
| 3 | .0476 | .0476 | 0 |
| 4 | .0476 | .0476 | 0 |
| 5 | .0476 | .0476 | 0 |
| . | | | |
| . | | | |
| . | | | |

## TABLE 3

DECOMPOSITION OF VARIANCE OF $j$-STEP-AHEAD PREDICTION ERROR VARIANCE IN MEASURED DATA

| $j$ | $y_n$ | $c$ | $k_{+1} - k$ |
|---|---|---|---|
| | A. Variance Due to Innovation in $y_n$ | | |
| 1 | 1.0057 | .0023 | .9036 |
| 2 | 1.0083 | .0045 | .9036 |
| 3 | 1.0108 | .0068 | .9036 |
| 4 | 1.0131 | .0090 | .9036 |
| 5 | 1.0155 | .0113 | .9036 |
| 6 | 1.0178 | .0135 | .9036 |
| 7 | 1.0201 | .0158 | .9036 |
| 8 | 1.0223 | .0181 | .9036 |
| 9 | 1.0246 | .0203 | .9036 |
| 10 | 1.0269 | .0226 | .9036 |
| 11 | 1.0291 | .0248 | .9036 |
| 12 | 1.0314 | .0271 | .9036 |
| 13 | 1.0337 | .0294 | .9036 |
| 14 | 1.0359 | .0316 | .9036 |
| 15 | 1.0382 | .0339 | .9036 |
| 16 | 1.0404 | .0361 | .9036 |
| 17 | 1.0427 | .0384 | .9036 |
| 18 | 1.0450 | .0407 | .9036 |
| 19 | 1.0472 | .0429 | .9036 |
| 20 | 1.0495 | .0452 | .9036 |
| | B. Variance Due to Innovation in $c$ | | |
| 1 | .0000 | .0025 | .0000 |
| 2 | .0000 | .0037 | .0000 |
| 3 | .0000 | .0044 | .0000 |
| 4 | .0000 | .0047 | .0000 |
| 5 | .0000 | .0049 | .0000 |
| 6 | .0000 | .0050 | .0000 |
| 7 | .0000 | .0051 | .0000 |
| 8 | .0000 | .0051 | .0000 |
| 9 | .0000 | .0052 | .0000 |
| 10 | .0000 | .0052 | .0000 |
| 11 | .0000 | .0052 | .0000 |
| 12 | .0001 | .0052 | .0000 |
| 13 | .0001 | .0052 | .0000 |
| 14 | .0001 | .0052 | .0000 |
| 15 | .0001 | .0053 | .0000 |
| 16 | .0001 | .0053 | .0000 |
| 17 | .0001 | .0053 | .0000 |
| 18 | .0001 | .0053 | .0000 |
| 19 | .0001 | .0053 | .0000 |
| 20 | .0001 | .0053 | .0000 |
| | C. Variance Due to Innovation in $\Delta k$ | | |
| 1 | .0000 | .0000 | .4682 |
| 2 | .0000 | .0000 | .5096 |
| 3 | .0000 | .0000 | .5133 |
| 4 | .0000 | .0000 | .5136 |
| 5 | .0000 | .0000 | .5136 |

TABLE 3 (*Continued*)

| $j$ | $y_n$ | $c$ | $k_{+1} - k$ |
|---|---|---|---|
| 6 | .0000 | .0000 | .5137 |
| 7 | .0000 | .0000 | .5137 |
| 8 | .0000 | .0000 | .5137 |
| 9 | .0000 | .0000 | .5137 |
| 10 | .0000 | .0000 | .5137 |
| 11 | .0000 | .0000 | .5137 |
| 12 | .0000 | .0000 | .5137 |
| 13 | .0000 | .0000 | .5137 |
| 14 | .0000 | .0000 | .5137 |
| 15 | .0000 | .0000 | .5137 |
| 16 | .0000 | .0000 | .5137 |
| 17 | .0000 | .0000 | .5137 |
| 18 | .0000 | .0000 | .5137 |
| 19 | .0000 | .0000 | .5137 |
| 20 | .0000 | .0000 | .5137 |

D. Covariance Matrix of Innovations in Measured Data on $y_n$, $c$, and $k_{+1} - k$

| | | |
|---|---|---|
| 1.0057 | .0476 | .9533 |
| .0476 | .0047 | .0453 |
| .9533 | .0453 | 1.3718 |

E. Eigenvalues of Covariance Matrix of Innovations in Measured Data on $y_n$, $c$, and $k_{+1} - k$

| |
|---|
| .0024 |
| .2183 |
| 2.1614 |

trix of the innovations in the measured data and its eigenvalues are reported in parts D and E of table 3.

Table 3 indicates a pattern in which innovations in $y_{nt}$ account for substantial proportions of the $j$-step-ahead prediction errors in $c_t$ and $k_{t+1} - k_t$, but not vice versa. The innovations in each of $c_t$ and $k_{t+1} - k_t$ contribute mainly to prediction error variances of themselves. Those conclusions are modified somewhat, but not substantially, when we change the orthogonalization order by letting $c_t$ go first (the results are not reported here).

Table 3 affirms that the Granger causality pattern portrayed by the extreme system (40), (41), and (42), in which income is measured perfectly, continues approximately to hold in a system in which, although income is measured with error, it is the best measured variable. Table 4 further affirms this point. The table reports a Wold representation for $(y_{nt}, c_t, k_{t+1} - k_t)$ in which the innovations are not orthogonalized. The pattern of coefficients closely approximates the pattern of zeros required to assure a situation in which Granger cau-

TABLE 4

IMPULSE RESPONSE FUNCTION OF MEASURED DATA (Wold Representation)

| $j$ | $y_n$ | $c$ | $k_{+1} - k$ |
|---|---|---|---|
| | A. Response to One Standard Deviation Innovation in $y_n$ | | |
| 0 | 1.0029 | .0000 | .0000 |
| 1 | .0578 | .0134 | −.2809 |
| 2 | .0526 | .0228 | −.0843 |
| 3 | .0494 | .0294 | −.0253 |
| 4 | .0475 | .0340 | −.0076 |
| 5 | .0464 | .0372 | −.0023 |
| 6 | .0457 | .0394 | −.0007 |
| 7 | .0453 | .0410 | −.0002 |
| 8 | .0451 | .0421 | −.0001 |
| 9 | .0449 | .0429 | −.0000 |
| 10 | .0448 | .0434 | −.0000 |
| 11 | .0448 | .0438 | −.0000 |
| 12 | .0447 | .0441 | −.0000 |
| 13 | .0447 | .0443 | −.0000 |
| | B. Response to One Standard Deviation Innovation in $c$ | | |
| 0 | .0000 | .0686 | .0000 |
| 1 | −.0025 | .0492 | −.0020 |
| 2 | .0000 | .0356 | −.0006 |
| 3 | .0015 | .0261 | −.0002 |
| 4 | .0025 | .0194 | −.0001 |
| 5 | .0030 | .0147 | −.0000 |
| 6 | .0033 | .0115 | −.0000 |
| 7 | .0035 | .0092 | −.0000 |
| 8 | .0037 | .0076 | −.0000 |
| 9 | .0037 | .0064 | −.0000 |
| 10 | .0038 | .0057 | −.0000 |
| 11 | .0038 | .0051 | −.0000 |
| 12 | .0038 | .0047 | −.0000 |
| 13 | .0038 | .0045 | −.0000 |
| | C. Response to One Standard Deviation Innovation in $\Delta k$ | | |
| 0 | .0000 | .0000 | 1.1712 |
| 1 | −.0063 | .0001 | .3481 |
| 2 | −.0037 | .0001 | .1044 |
| 3 | −.0021 | .0002 | .0313 |
| 4 | −.0012 | .0002 | .0094 |
| 5 | −.0006 | .0002 | .0028 |
| 6 | −.0003 | .0002 | .0008 |
| 7 | −.0001 | .0002 | .0003 |
| 8 | .0001 | .0002 | .0001 |
| 9 | .0001 | .0002 | .0000 |
| 10 | .0002 | .0002 | .0000 |
| 11 | .0002 | .0002 | .0000 |
| 12 | .0002 | .0002 | .0000 |
| 13 | .0002 | .0002 | .0000 |

TABLE 5

DECOMPOSITION OF VARIANCE OF $j$-STEP-AHEAD PREDICTION ERROR OF FILTERED DATA ON $y_n$, $c$, AND $k_{+1} - k$

| $j$ | $y_n$ | $c$ | $k_{+1} - k$ |
|---|---|---|---|
| | A. Variance Due to Innovation in $y_n$ | | |
| 1 | .9945 | .0023 | .9019 |
| 2 | .9968 | .0045 | .9019 |
| 3 | .9991 | .0068 | .9019 |
| 4 | 1.0014 | .0090 | .9019 |
| 5 | 1.0036 | .0113 | .9019 |
| 6 | 1.0058 | .0136 | .9019 |
| 7 | 1.0081 | .0158 | .9019 |
| 8 | 1.0104 | .0181 | .9019 |
| 9 | 1.0126 | .0204 | .9019 |
| 10 | 1.0149 | .0226 | .9019 |
| 11 | 1.0171 | .0249 | .9019 |
| 12 | 1.0194 | .0271 | .9019 |
| 13 | 1.0217 | .0294 | .9019 |
| 14 | 1.0239 | .0317 | .9019 |
| 15 | 1.0262 | .0339 | .9019 |
| 16 | 1.0285 | .0362 | .9019 |
| 17 | 1.0307 | .0385 | .9019 |
| 18 | 1.0330 | .0407 | .9019 |
| 19 | 1.0352 | .0430 | .9019 |
| 20 | 1.0375 | .0452 | .9019 |
| | B. Variance Due to Innovation in $c^*$ | | |
| 1 | .0000 | .0058 | .0058 |
| 2 | .0058 | .0116 | .0058 |
| 3 | .0116 | .0174 | .0058 |
| 4 | .0174 | .0231 | .0058 |
| 5 | .0231 | .0289 | .0058 |
| 6 | .0289 | .0347 | .0058 |
| 7 | .0347 | .0405 | .0058 |
| 8 | .0405 | .0463 | .0058 |
| 9 | .0463 | .0521 | .0058 |
| 10 | .0521 | .0578 | .0058 |
| 11 | .0578 | .0636 | .0058 |
| 12 | .0636 | .0694 | .0058 |
| 13 | .0694 | .0752 | .0058 |
| 14 | .0752 | .0810 | .0058 |
| 15 | .0810 | .0868 | .0058 |
| 16 | .0868 | .0925 | .0058 |
| 17 | .0925 | .0983 | .0058 |
| 18 | .0983 | .1041 | .0058 |
| 19 | .1041 | .1099 | .0058 |
| 20 | .1099 | .1157 | .0058 |
| | C. Variance Due to Innovation in $\Delta k^*$ | | |
| 1 | .0000 | .0000 | .3000 |
| 2 | .0333 | .0333 | .3000 |
| 3 | .0667 | .0667 | .3000 |
| 4 | .1000 | .1000 | .3000 |
| 5 | .1333 | .1333 | .3000 |

TABLE 5 (*Continued*)

| $j$ | $y_n$ | $c$ | $k_{+1} - k$ |
|---|---|---|---|
| 6 | .1667 | .1667 | .3000 |
| 7 | .2000 | .2000 | .3000 |
| 8 | .2333 | .2333 | .3000 |
| 9 | .2667 | .2667 | .3000 |
| 10 | .3000 | .3000 | .3000 |
| 11 | .3333 | .3333 | .3000 |
| 12 | .3667 | .3667 | .3000 |
| 13 | .4000 | .4000 | .3000 |
| 14 | .4333 | .4333 | .3000 |
| 15 | .4667 | .4667 | .3000 |
| 16 | .5000 | .5000 | .3000 |
| 17 | .5333 | .5333 | .3000 |
| 18 | .5667 | .5667 | .3000 |
| 19 | .6000 | .6000 | .3000 |
| 20 | .6333 | .6333 | .3000 |

D. Covariance Matrix of Innovations to Filtered Data on $y_n$, $c$, and $k_{+1} - k$

| | | |
|---|---|---|
| .9945 | .0474 | .9471 |
| .0474 | .0023 | .0452 |
| .9471 | .0452 | .9019 |

E. Eigenvalues of Covariance Matrix of Innovations to Filtered Data on $y_n$, $c$, and $\Delta k$

| |
|---|
| 1.898723 |
| .000000 |
| .000009 |

* All entries in pt. B are multiplied by $10^{-3}$. All entries in pt. C are multiplied by $10^{-10}$.

sality extends from $y_n$ to $c$ and $(k_{t+1} - k_t)$, but not vice versa (see Sims 1972*b*).

These figures illustrate how the first model of the reporting agency in conjunction with the one-index economic model offers an explanation for the tendency of accelerator equations to survive Granger causality tests. I now indicate how much things can change when we use the second model of the measurement agency.

Table 5 reports a decomposition of variance for the $j$-step-ahead prediction errors for the optimally filtered data for $y_{nt}$, $c_t$, and $k_{t+1} - k_t$. Notice that the covariance matrix of the innovations is nearly singular, there being one dominant eigenvalue. This means that we are close to a situation in which there is only one shock impinging on the system for the filtered data (just as there is actually only one shock in the system governing the *true* data). The decomposition of variance of $j$-step-ahead prediction errors in table 5 confirms this. The first innovation accounts for virtually all the variances of prediction errors in *all* the variables.

Table 6 reports a Wold representation for the filtered data, with the innovations orthogonalized via a Cholesky decomposition that amounts to ordering the variables in the order $y_{nt}$, $c_t$, and $k_{t+1} - k_t$ (see Sims 1980). The Wold representation has one dominant noise, the second and third noises bearing close to zero coefficients for all variables. Notice that the moving average associated with the first noise for the filtered data closely approximates the impulse response function for the true data in response to an innovation in the endowment shock $\theta_t$ (see table 2). That the Wold representation in table 6 closely approximates the impulse response function in table 2 indicates that the filtered data mimic the true data very closely. Among other things, this means that in the present context the second model of measurements is not able to explain what we have interpreted partly as the empirical success of the accelerator: the tendency for income or output to Granger-cause, but not to be Granger-caused by, investment.

These impressions about the very good fit of the filtered to the true data are affirmed by a simulation of length 80 periods for the system. Figure 1–3 graph true and measured consumption, investment, and income, respectively. The assumed ordering of measurement error variances shows up in these figures. Figures 4, 5, 6, and 7 display true and filtered consumption, investment, income, and capital stock, respectively. Notice how closely the filtered and true series match up in each case. Table 7 reports the covariance and correlation matrix for the true, measured, and filtered data from the simulation.

We have assumed that the measurement errors to consumption and investment are orthogonal. This means that measured investment and consumption fail to add up to measured income. Figure 8 illustrates this. Figure 9 shows that filtered consumption and investment do add up to filtered income. It is easy for the optimal filter to build in the GNP identity.

## V. Concluding Remarks

It is possible to imagine many more models of measurements. Even if we retain the basic structure of the two models described here and continue to assume that the measurements collected by the agency satisfy classical errors-in-variables orthogonality assumptions, we can generate additional models of the reporting agency's "filtering" activities by imputing to the agency different models of the economy. I have adopted a very special (and workable) case by assuming that the reporting agency has the correct model (i.e., the model of the econo-

TABLE 6

WOLD REPRESENTATION OF FILTERED DATA ON $y_n$, $c$, AND $k_{+1} - k$

| $j$ | $y_n$ | $c$ | $k_{+1} - k$ |
|---|---|---|---|
| | A. Response to an Innovation in $y_n$ | | |
| 0 | .9973 | .0476 | .9497 |
| 1 | .0476 | .0476 | .0000 |
| 2 | .0476 | .0476 | .0000 |
| 3 | .0476 | .0476 | .0000 |
| 4 | .0476 | .0476 | .0000 |
| 5 | .0476 | .0476 | .0000 |
| 6 | .0476 | .0476 | .0000 |
| 7 | .0476 | .0476 | .0000 |
| 8 | .0476 | .0476 | .0000 |
| 9 | .0476 | .0476 | .0000 |
| 10 | .0476 | .0476 | .0000 |
| 11 | .0476 | .0476 | .0000 |
| 12 | .0476 | .0476 | .0000 |
| 13 | .0476 | .0476 | .0000 |
| | B. Response to an Innovation in $c$ | | |
| 0 | .0000 | .0024 | −.0024 |
| 1 | .0024 | .0024 | .0000 |
| 2 | .0024 | .0024 | .0000 |
| 3 | .0024 | .0024 | .0000 |
| 4 | .0024 | .0024 | .0000 |
| 5 | .0024 | .0024 | .0000 |
| 6 | .0024 | .0024 | .0000 |
| 7 | .0024 | .0024 | .0000 |
| 8 | .0024 | .0024 | .0000 |
| 9 | .0024 | .0024 | .0000 |
| 10 | .0024 | .0024 | .0000 |
| 11 | .0024 | .0024 | .0000 |
| 12 | .0024 | .0024 | .0000 |
| 13 | .0024 | .0024 | .0000 |
| | C. Response to an Innovation in $\Delta k$* | | |
| 0 | .0000 | .0000 | .5477 |
| 1 | −.1826 | −.1826 | .0000 |
| 2 | −.1826 | −.1826 | .0000 |
| 3 | −.1826 | −.1826 | .0000 |
| 4 | −.1826 | −.1826 | .0000 |
| 5 | −.1826 | −.1826 | .0000 |
| 6 | −.1826 | −.1826 | .0000 |
| 7 | −.1826 | −.1826 | .0000 |
| 8 | −.1826 | −.1826 | .0000 |
| 9 | −.1826 | −.1826 | .0000 |
| 10 | −.1826 | −.1826 | .0000 |
| 11 | −.1826 | −.1826 | .0000 |
| 12 | −.1826 | −.1826 | .0000 |
| 13 | −.1826 | −.1826 | .0000 |

* All entries in pt. C are to be multiplied by $10^{-5}$.

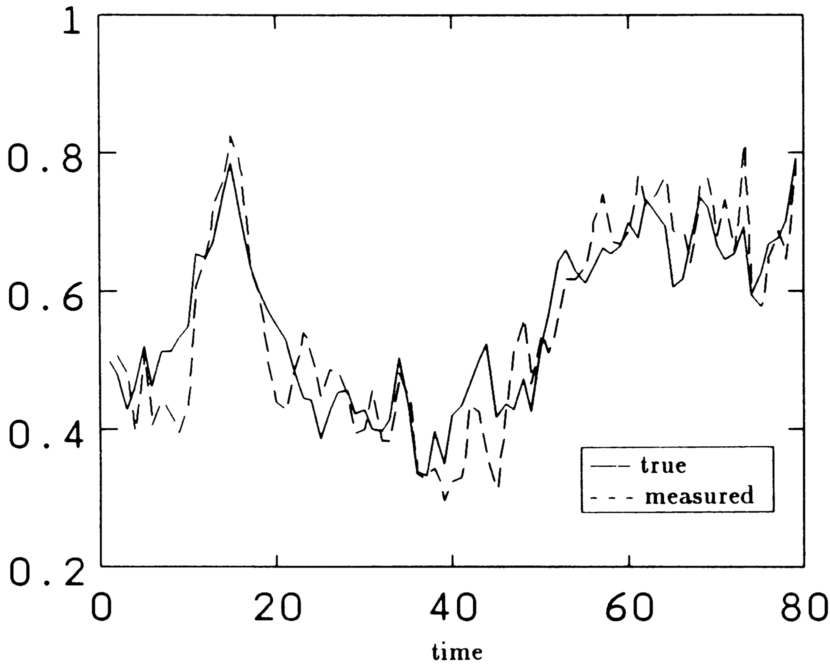


FIG. 1.—True and measured consumption

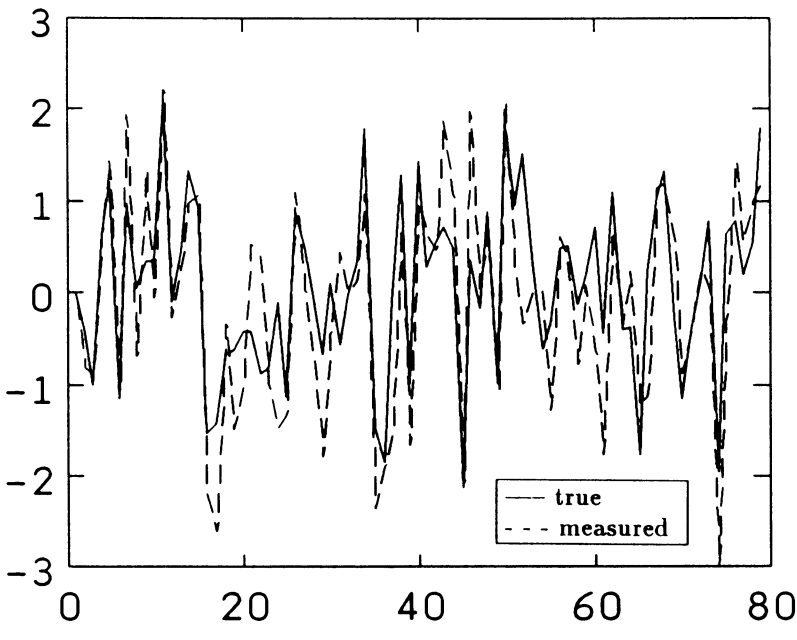


FIG. 2.—True and measured investment

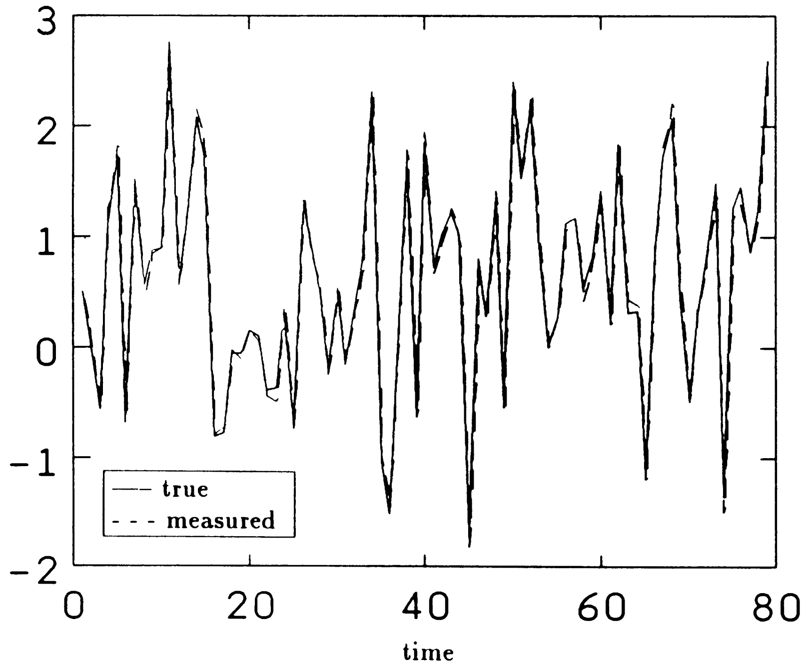


FIG. 3.—True and measured income

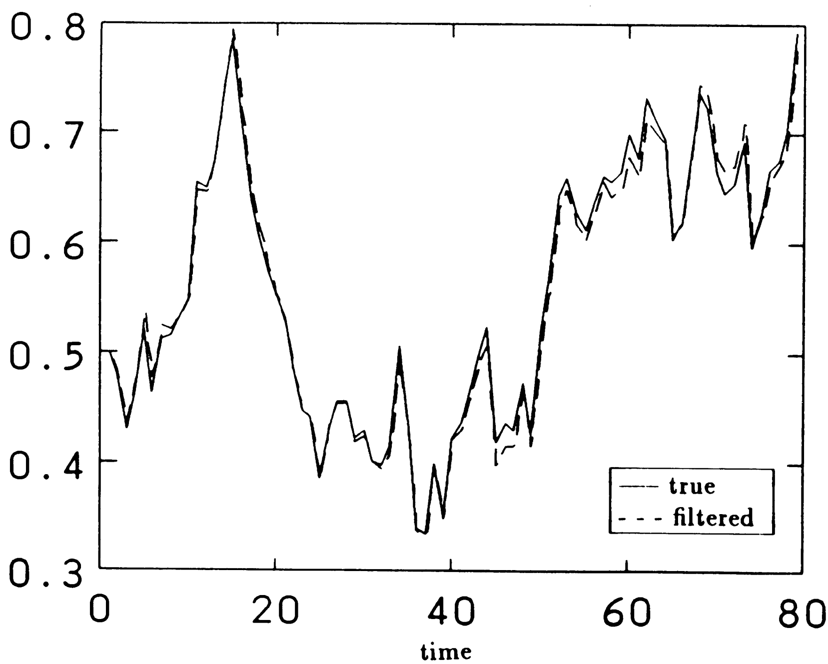


FIG. 4.—True and filtered consumption

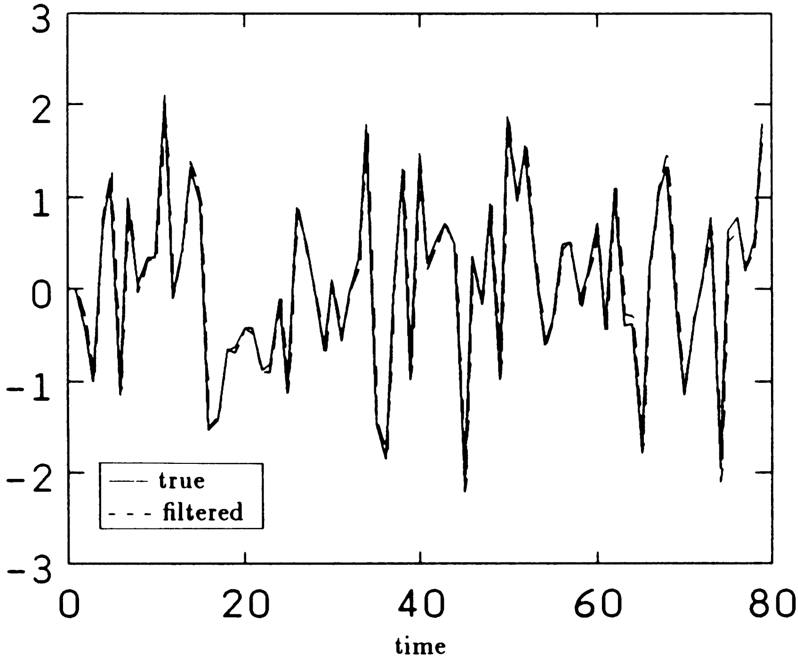


FIG. 5.—True and filtered investment

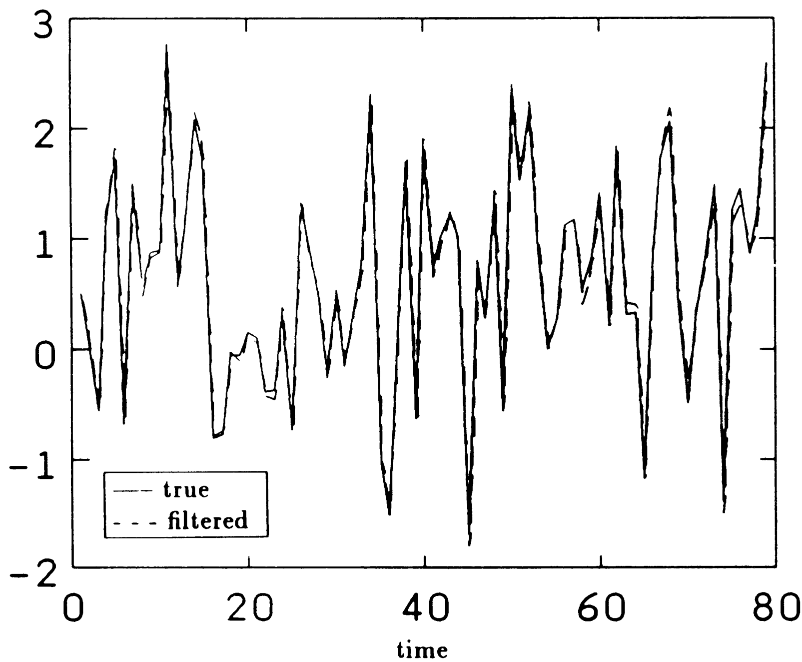


FIG. 6.—True and filtered income

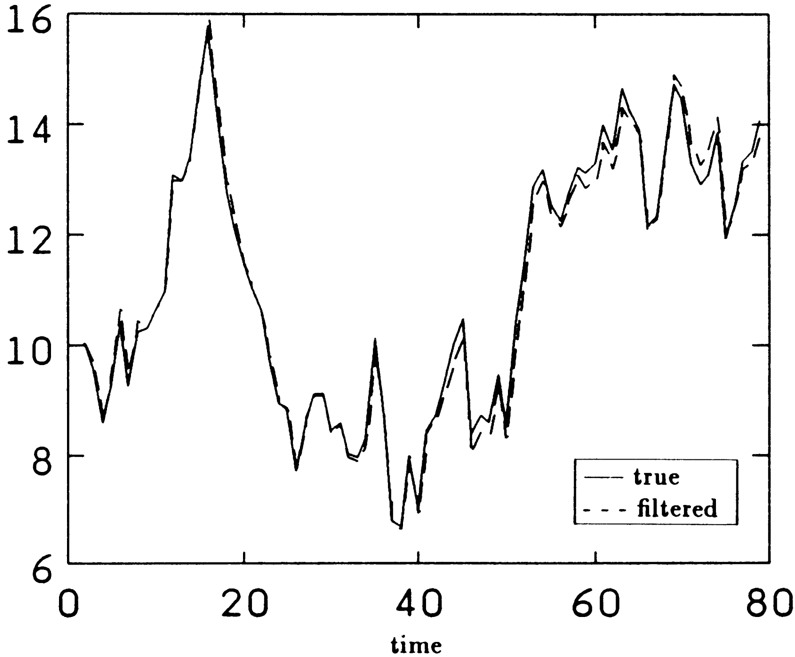


FIG. 7.—True and filtered capital stock

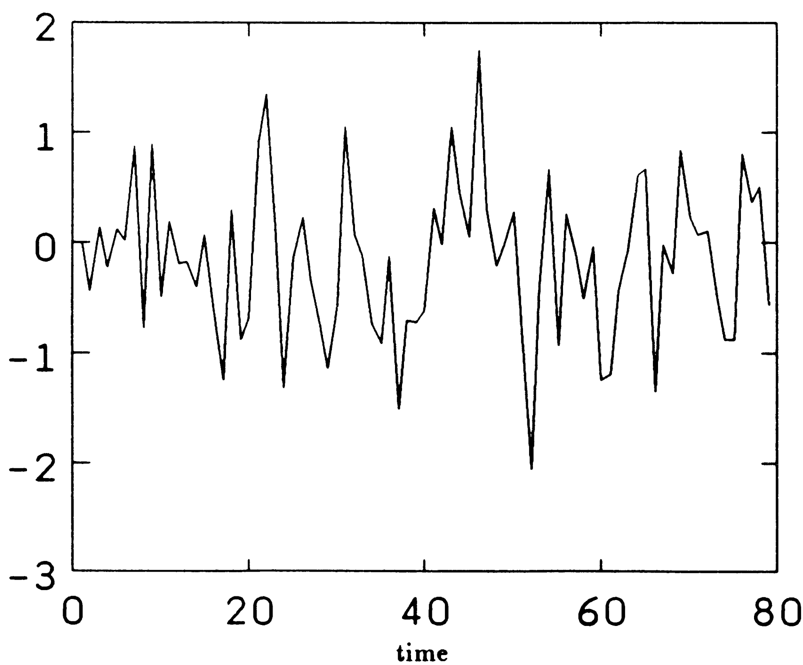


FIG. 8.—Measured consumption plus investment minus income

TABLE 7

| | | |
|---|---|---|
| A. Covariance Matrix of True, Measured, and Filtered Consumption | | |
| .0141 | .0155 | .0141 |
| .0155 | .0205 | .0156 |
| .0141 | .0156 | .0142 |
| B. Correlation Matrix of True, Measured, and Filtered Consumption | | |
| 1.0000 | .9092 | .9966 |
| .9092 | 1.0000 | .9129 |
| .9966 | .9129 | 1.0000 |
| C. Covariance Matrix of True, Measured, and Filtered Investment | | |
| .8942 | .8890 | .8939 |
| .8890 | 1.3609 | .8882 |
| .8939 | .8882 | .8981 |
| D. Correlation Matrix of True, Measured, and Filtered Investment | | |
| 1.0000 | .8058 | .9974 |
| .8958 | 1.0000 | .8034 |
| .9974 | .8034 | 1.0000 |
| E. Covariance Matrix of True, Measured, and Filtered Income | | |
| .9667 | .9713 | .9646 |
| .9713 | .9813 | .9744 |
| .9646 | .9744 | .9677 |
| F. Correlation Matrix of True, Measured, and Filtered Income | | |
| 1.0000 | .9972 | .9973 |
| .9972 | 1.0000 | .9999 |
| .9973 | .9999 | 1.0000 |

| | |
|---|---|
| G. Covariance Matrix of True and Filtered Capital | |
| 5.3672 | 5.4112 |
| 5.4112 | 5.4894 |
| H. Correlation Matrix of True and Filtered Capital | |
| 1.0000 | .9969 |
| .9969 | 1.0000 |

mist who wants to compute a likelihood function). It would evidently be possible to work out procedures analogous to the ones described above that would cover cases in which the data reporting agency filters the data using a particular erroneous model. That would lead to econometric specifications that proliferate parameters relative to the specifications described here because the likelihood function would contain parameters describing the reporting agency's erroneous model, in addition to the parameters describing the economic model and the measurement error processes that appear in the likelihood functions of the two models.

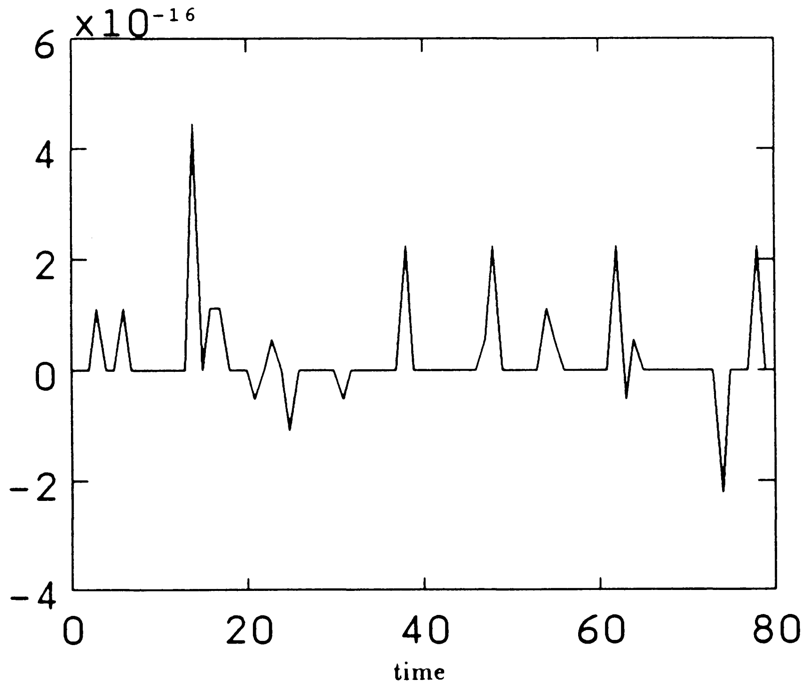


FIG. 9.—Filtered consumption plus investment minus income

## Appendix A

### The Kalman Filter

Consider the state-space system

$$\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \boldsymbol{\epsilon}_t,$$
$$\mathbf{y}_t = \mathbf{C}\mathbf{x}_t + \mathbf{w}_t,$$

where $\mathbf{x}_t$ is $n \times 1$, $\boldsymbol{\epsilon}_t$ is $n \times 1$, $\mathbf{y}_t$ is $k \times 1$, $\mathbf{w}_t$ is $k \times 1$, $\mathbf{A}$ is $n \times n$, $\mathbf{C}$ is $k \times n$, $\boldsymbol{\epsilon}_t$ and $\mathbf{w}_t$ are each vector white noises with

$$E\boldsymbol{\epsilon}_t\boldsymbol{\epsilon}_t' = \mathbf{Q},$$
$$E\mathbf{w}_t\mathbf{w}_t' = \mathbf{R},$$
$$E\boldsymbol{\epsilon}_t\mathbf{w}_t' = \mathbf{W}.$$

The matrices $\mathbf{Q}$ and $\mathbf{R}$ are each positive semidefinite. Define the matrix $\mathbf{S}$ to be the unique positive semidefinite matrix that satisfies the algebraic Ricatti equation

$$\mathbf{S} = \mathbf{Q} + \mathbf{ASA}' - (\mathbf{ASC}' + \mathbf{W})(\mathbf{R} + \mathbf{CSC}')^{-1}(\mathbf{CSA}' + \mathbf{W}'). \quad \text{(A1)}$$

Define $\mathbf{K}$ in terms of $\mathbf{S}$ by

$$\mathbf{K} = (\mathbf{ASC}' + \mathbf{W})(\mathbf{CSC}' + \mathbf{R})^{-1}. \quad \text{(A2)}$$

Methods for solving (A1) and (A2) for $\mathbf{S}$ and $\mathbf{K}$ are described in Anderson and Moore (1979). In the text, we denote the solutions of (A1) and (A2) by the function

$$[\mathbf{K}, \mathbf{S}] = \text{kfilter}(\mathbf{A}, \mathbf{C}, \mathbf{Q}, \mathbf{R}, \mathbf{W}).$$

## Appendix B

### Serially Correlated Measurement Errors

The calculations described in Section III handle only the case in which the measurement error processes are first-order autoregressive processes. This Appendix shows how to modify the calculations to permit the measurement error to be an autoregressive process of arbitrary finite order. This can be accomplished by augmenting the state vector of the measurement error $\mathbf{v}_t$ to include lagged values of the measurement errors themselves and properly defining $\mathbf{D}$ to have submatrices that are companion matrices. We require that the observer equation be modified to

$$\mathbf{z}_t = \mathbf{C}\mathbf{x}_t + \mathbf{H}\mathbf{v}_t,$$

where $\mathbf{H}$ is a selector matrix that picks off the current values of the measurement errors and ignores the others. As an example, suppose that we have univariate measurement error $v_{1t}$ that we want to model as a second-order autoregressive process

$$v_{1t} = \delta_1 v_{1t-1} + \delta_2 v_{1t-2} + \nu_{1t-1}.$$

Suppose that $v_{1t}$ is the only measurement error in the system. We would define

$$\mathbf{v}_t = \begin{pmatrix} v_{1t} \\ v_{1t-1} \end{pmatrix}$$

and write

$$\mathbf{v}_{t+1} = \mathbf{D}\mathbf{v}_t + \boldsymbol{\nu}_t,$$

where

$$\mathbf{D} = \begin{bmatrix} \delta_1 & \delta_2 \\ 1 & 0 \end{bmatrix}, \boldsymbol{\nu}_t = \begin{bmatrix} \nu_{1t} \\ 0 \end{bmatrix}.$$

Suppose that there is a single measured variable in the system, $\mathbf{z}_t = \mathbf{C}\mathbf{x}_t + v_{1t}$. We would express this as $\mathbf{z}_t = \mathbf{C}\mathbf{x}_t + \mathbf{H}\mathbf{v}_t$, where $\mathbf{H} = [1 \quad 0]$.

Consider the system

$$\begin{aligned} \mathbf{x}_{t+1} &= \mathbf{A}\mathbf{x}_t + \boldsymbol{\epsilon}_t, \\ \mathbf{z}_t &= \mathbf{C}\mathbf{x}_t + \mathbf{H}\mathbf{v}_t, \\ \mathbf{v}_{t+1} &= \mathbf{D}\mathbf{v}_t + \boldsymbol{\nu}_t, \end{aligned} \tag{B1}$$

where $\mathbf{x}_t$ is $n \times 1$, $\mathbf{z}_t$ is $k \times 1$, $\mathbf{v}_t$ is $m \times 1$, and the matrices $\mathbf{A}$, $\mathbf{C}$, $\mathbf{H}$, and $\mathbf{D}$ are conformable with the objects that they appear with. We assume that $\boldsymbol{\epsilon}_t$ and $\boldsymbol{\nu}_t$ are white noises with $E\boldsymbol{\epsilon}\boldsymbol{\epsilon}_t' = \mathbf{Q}$, $E\boldsymbol{\nu}_t\boldsymbol{\nu}_t' = \mathbf{R}$, and $E\boldsymbol{\epsilon}_t\boldsymbol{\nu}_s' = 0$ for all $t$ and $s$. Notice that

$$\bar{\mathbf{z}}_t \equiv \mathbf{z}_{t+1} - \mathbf{D}\mathbf{z}_t = (\mathbf{CA} - \mathbf{DC})\mathbf{x}_t + (\mathbf{C}\boldsymbol{\epsilon}_t + \mathbf{H}\boldsymbol{\nu}_t) + (\mathbf{HD} - \mathbf{DH})\mathbf{v}_t. \tag{B2}$$

We can combine (B1) and (B2) to obtain the state-space system

$$\begin{pmatrix} \mathbf{x}_{t+1} \\ \mathbf{v}_{t+1} \end{pmatrix} = \begin{pmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{D} \end{pmatrix} \begin{pmatrix} \mathbf{x}_t \\ \mathbf{v}_t \end{pmatrix} + \begin{pmatrix} \boldsymbol{\epsilon}_t \\ \boldsymbol{\nu}_t \end{pmatrix},$$

$$\bar{\mathbf{z}}_t = [\mathbf{CA} - \mathbf{DC}, \mathbf{HD} - \mathbf{DH}] \begin{pmatrix} \mathbf{x}_t \\ \mathbf{v}_t \end{pmatrix} + (\mathbf{C}\boldsymbol{\epsilon}_t + \mathbf{H}\boldsymbol{\nu}_t). \tag{B3}$$

The state of the system is $(\mathbf{x}_t, \mathbf{v}_t)$. The covariance matrix for state noise is seen to be

$$\begin{pmatrix} \mathbf{Q} & \mathbf{0} \\ \mathbf{0} & \mathbf{R} \end{pmatrix}.$$

The covariance matrix of the observation noise $\mathbf{C}\boldsymbol{\epsilon}_t + \mathbf{H}\boldsymbol{\nu}_t$ is computed to be $\mathbf{CQC}' + \mathbf{HRH}'$. The covariance between the state noise and the observation noise is

$$\begin{bmatrix} \mathbf{C}\,\mathbf{Q} \\ \mathbf{H}\,\mathbf{R} \end{bmatrix}.$$

To obtain an innovations representation for system (B3), we compute

$$[\mathbf{K}_3, \mathbf{S}_3] = \text{kfilter}\left(\begin{pmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{D} \end{pmatrix}, [\mathbf{CA} - \mathbf{DC}, \mathbf{HD} - \mathbf{DH}], \begin{pmatrix} \mathbf{Q} & \mathbf{0} \\ \mathbf{0} & \mathbf{R} \end{pmatrix}, \mathbf{CQC}' + \mathbf{HRH}', \begin{bmatrix} \mathbf{C}\,\mathbf{Q} \\ \mathbf{H}\,\mathbf{R} \end{bmatrix}\right).$$

The innovations representation for (B3) is then

$$\begin{bmatrix} \hat{\mathbf{x}}_{t+1} \\ \hat{\mathbf{v}}_{t+1} \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{D} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{x}}_t \\ \hat{\mathbf{v}}_t \end{bmatrix} + \mathbf{K}_3\mathbf{u}_t,$$

where

$$\bar{\mathbf{z}}_t = [\mathbf{CA} - \mathbf{DC}, \mathbf{HD} - \mathbf{DH}] \begin{pmatrix} \hat{\mathbf{x}}_t \\ \hat{\mathbf{v}}_t \end{pmatrix} + \mathbf{u}_t,$$

$$\begin{bmatrix} \hat{\mathbf{x}}_t \\ \hat{\mathbf{v}}_t \end{bmatrix} = E\left[\begin{bmatrix} \mathbf{x}_t \\ \mathbf{v}_t \end{bmatrix} \middle| \mathbf{z}_t, \mathbf{z}_{t-1}, \ldots, \mathbf{z}_0, \begin{bmatrix} \hat{\mathbf{x}}_0 \\ \hat{\mathbf{v}}_0 \end{bmatrix}\right],$$

$$\mathbf{S}_3 = E\begin{pmatrix} \mathbf{x}_t - \hat{\mathbf{x}}_t \\ \mathbf{v}_t - \hat{\mathbf{v}}_t \end{pmatrix}\begin{pmatrix} \mathbf{x}_t - \hat{\mathbf{x}}_t \\ \mathbf{v}_t - \hat{\mathbf{v}}_t \end{pmatrix}',$$

$$\mathbf{u}_t = \bar{\mathbf{z}}_t - E[\bar{\mathbf{z}}_t | \bar{\mathbf{z}}_{t-1}, \ldots, \bar{\mathbf{z}}_0, \hat{\mathbf{x}}_0, \hat{\boldsymbol{\nu}}_0].$$

## References

Abel, Andrew, and Blanchard, Olivier J. "Investment and Sales: Some Empirical Evidence." Paper presented at NBER meetings, Cambridge, Mass., July 1983.

Anderson, Brian D. O., and Moore, John B. *Optimal Filtering.* Englewood Cliffs, N.J.: Prentice-Hall, 1979.

Brock, William A., and Mirman, Leonard J. "Optimal Economic Growth and Uncertainty: The Discounted Case." *J. Econ. Theory* 4 (June 1972): 479–513.

Eichenbaum, Martin S. "A Rational Expectations Equilibrium Model of Inventories of Finished Goods and Employment." *J. Monetary Econ.* 12 (August 1983): 259–77.

———. "Rational Expectations and the Smoothing Properties of Inventories of Finished Goods." *J. Monetary Econ.* 14 (July 1984): 71–96.

Flavin, Marjorie A. "The Adjustment of Consumption to Changing Expectations about Future Income." *J.P.E.* 89 (October 1981): 974–1009.

———. "Excess Sensitivity of Consumption to Current Income: Liquidity Constraints or Myopia?" *Canadian J. Econ.* 18 (February 1985): 117–36.

Friedman, Milton. *A Theory of the Consumption Function.* Princeton, N.J.: Princeton Univ. Press (for NBER), 1957.

———. "Windfalls, the 'Horizon,' and Related Concepts in the Permanent-Income Hypothesis." In *Measurement in Economics: Studies in Mathematical Economics and Econometrics in Memory of Yehuda Grunfeld,* by Carl Christ et al. Stanford, Calif.: Stanford Univ. Press, 1963.

Futia, Carl A. "Rational Expectations in Stationary Linear Models." *Econometrica* 49 (January 1981): 171–92.

Hall, Robert E. "Stochastic Implications of the Life Cycle–Permanent Income Hypothesis: Theory and Evidence." *J.P.E.* 86 (December 1978): 971–87.

Hansen, Lars Peter, and Sargent, Thomas J. "Formulating and Estimating Dynamic Linear Rational Expectations Models." *J. Econ. Dynamics and Control* 2 (February 1980): 7–46.

———. "Formulating and Estimating Continuous Time Rational Expectations Models." Staff Report no. 75. Minneapolis: Fed. Reserve Bank, October 1981.

———. "Two Difficulties in Interpreting Vector Autoregressions." Working Paper no. 227. Minneapolis: Fed. Reserve Bank, November 1982.

Harvey, A. C. *Time Series Models.* New York: Halsted, 1981.

Harvey, A. C., and Stock, James H. "The Estimation of Higher Order Continuous Time Autoregressive Models." *Econometric Theory* 1 (April 1985): 97–112.

———. "Estimating Integrated Higher Order Continuous Time Autoregressions with an Application to Money-Income Causality." Working Paper no. E87-28. Stanford, Calif.: Hoover Inst., July 1987.

Kydland, Finn E., and Prescott, Edward C. "Time to Build and Aggregate Fluctuations." *Econometrica* 50 (November 1982): 1345–70.

Ljung, Lennart, and Söderström, Torsten. *Theory and Practice of Recursive Identification.* Cambridge, Mass.: M.I.T. Press, 1983.

Lucas, Robert E., Jr., and Prescott, Edward C. "Investment under Uncertainty." *Econometrica* 39 (September 1971): 659–81.

Mankiw, N. Gregory; Runkle, David E.; and Shapiro, Matthew D. "Are Preliminary Announcements of the Money Stock Rational Forecasts?" *J. Monetary Econ.* 14 (July 1984): 15–27.

Mankiw, N. Gregory, and Shapiro, Matthew D. "News or Noise: An Analysis of GNP Revisions." *Survey Current Bus.* 66 (May 1986): 20–25.

Muth, John F. "Optimal Properties of Exponentially Weighted Forecasts." *J. American Statis. Assoc.* 55 (June 1960): 299–306.

Sargent, Thomas J. " 'Tobin's *q*' and the Rate of Investment in General Equilibrium." In *On the State of Macroeconomics,* edited by Karl Brunner and

Allan H. Meltzer. Carnegie-Rochester Conference Series on Public Policy, vol. 13. Suppl., *J. Monetary Econ.* Amsterdam: North-Holland, 1980.

———. *Macroeconomic Theory.* 2d ed. New York: Academic Press, 1987.

Sargent, Thomas J., and Sims, Christopher A. "Business Cycle Modeling without Pretending to Have Too Much *a Priori* Economic Theory." In *New Methods in Business Cycle Research,* edited by Christopher A. Sims. Minneapolis: Fed. Reserve Bank, 1977.

Sims, Christopher A. "Are There Exogenous Variables in Short-Run Production Relations?" *Ann. Econ. and Soc. Measurement* 1 (January 1972): 17–36. (*a*)

———. "Money, Income, and Causality." *A.E.R.* 62 (September 1972): 540–52. (*b*)

———. "The Role of Approximate Prior Restrictions in Distributed Lag Estimation." *J. American Statis. Assoc.* 67 (March 1972): 169–75. (*c*)

———. "Output and Labor Input in Manufacturing." *Brookings Papers Econ. Activity,* no. 3 (1974), pp. 695–728.

———. "Macroeconomics and Reality." *Econometrica* 48 (January 1980): 1–48.

Townsend, Robert M. "Forecasting the Forecasts of Others." *J.P.E.* 91 (August 1983): 546–88.

Whiteman, Charles H. *Linear Rational Expectations Models: A User's Guide.* Minneapolis: Univ. Minnesota Press, 1983.

Whittle, Peter. *Prediction and Regulation by Linear Least-Squares Methods.* 2d ed. Minneapolis: Univ. Minnesota Press, 1983.