Forecasting the Forecasts of Others

Robert M. Townsend

Carnegie-Mellon University

This paper explores the formulation and analysis of linear equilibrium models of investment in which learning is perpetual and informationally decentralized firms need never share the same beliefs concerning time series relevant to their decisions. Recursive, Kalman filtering techniques are shown to be applicable in an illustrative, hierarchical information structure, and a nonlinear technique of undetermined coefficients is shown to be applicable in an illustrative, symmetric information structure in which there is a confounding of laws of motion with forecasting problems. The equilibrium time series of these models can display interesting movement in response to shocks and measurement errors, including persistence, certain cross-correlation properties, and damped oscillations. That is, forecast errors are serially correlated over decision makers and serially correlated over time in a certain crucial sense. More generally, these models do place restrictions on observed time series and can be fitted to data.

I. Introduction

That economic decision makers do not share the same beliefs concerning the time series of economic variables relevant to their decisions is a key part of the writings of such diverse authors as Pigou

I am extraordinarily indebted to Lars Peter Hansen, Thomas J. Sargent, Edward C. Prescott, and Ian Bain for instruction and insights; would like to thank Robert E. Lucas, Jr., Edmund S. Phelps, Carl Futia, and participants of seminars at Carnegie-Mellon University, the University of Chicago, Columbia University, the Federal Reserve Bank of Minneapolis, and the NBER Conference on Business Fluctuations in Chicago, October 1981, for helpful comments; and accept full responsibility for errors and the views expressed herein. Support for this research from the National Science Foundation and the Alfred P. Sloan Foundation and programming assistance from Riccardo Revelli and Lakshmi Subramanian are also gratefully acknowledged.

[*Journal of Political Economy*, 1983, vol. 91, no. 4]

(1929), Keynes (1936), and Lucas (1972, 1975). In his explanation of industrial fluctuations, Pigou (1929) emphasizes that in modern industry businessmen are partially but not completely interconnected, possessing only limited information on what others are doing. Pigou then argues that these interconnections tend to promote the mutual generation of forecast errors, both within and across industries. Further, such forecast errors can be mutually generated over time, leading to serial correlation and waves of optimism and pessimism. Similarly, it was Keynes's (1936) view that with the opening of stock exchanges investors shift from a concern with long-term determinants of profitability or yield to a concern with short-term movements in market prices themselves. Thus, conventional valuation would be liable to change violently as the result of sudden fluctuations in opinion due to factors which do not really make much difference to prospective yield, and even relatively informed investors would be concerned with average opinion. More recently, Lucas (1972) has emphasized the role of confusion between relative and absolute price movements in his explanation of the positive correlation between prices and economic activity. And Lucas (1975) uses economy-wide average beliefs as a state variable, along with money and capital, to explain certain stylized facts of business cycles, especially persistence and cumulation in response to serially uncorrelated economic shocks.

This paper represents an attempt to develop further an equilibrium account of some of the qualitative observations which motivate Pigou (1929) and Keynes (1936) as well as Lucas (1972, 1975). That is, it follows Muth (1961), Lucas and Prescott (1971), and Lucas (1972, 1975) in adopting the "rational expectations" hypothesis, supposing that (1) decision makers act as if they were solving explicit inference and dynamic decision problems, with well-specified information sets and objective functions, taking as given the aggregate laws of motion of observed and unobserved state variables, including perhaps the laws describing the inferences of other decision makers, and (2) these laws of motion are in turn those actually generated in the model by the inferences and decisions of the decision makers.[1] The important aspect of this formulation is that it can accommodate learning and disparate expectations across decision makers and can handle information structures in which decision makers forecast the forecasts of others.[2]

[1] The equilibrium definition employed here builds on Townsend (1978), Sargent (1979), and Prescott and Townsend (1980), among others, and is related to ongoing work in the theory of general economic equilibrium by Green (1977), Jordan (1977), Kreps (1977), Radner (1979), and Allen (1981), among others. Unlike the general equilibrium literature, however, here prices do not simultaneously clear markets and convey information.

[2] Again, the learning takes place in a rational expectations equilibrium way, as in

This paper emphasizes four qualitative characteristics of economic time series which equilibrium models of this kind can generate. The first is the "learning mechanism" for the propagation of economic shocks. That is, learning alone can convert serially uncorrelated economic shocks into serially correlated movements in economic decision variables.[3] This result is not particularly novel—it is clear from the earlier work of Lucas (1975) and Crawford (1979), for example. But perhaps this propagation mechanism has not yet received the attention it deserves.[4] Second, because decision makers have limited a priori information on underlying economic shocks, they may well respond to variables generated by the decisions of others and to the noise in such decisions, even if there are no real economic links among such decision makers. Thus, time series can display certain cross-correlation properties and may appear more noisy than if decision makers acted in isolation. Third, models with disparate but rational expectations, in which decision makers forecast the forecasts of others, can lead to some rather new and exciting dynamics. In particular, the paper shows how serially uncorrelated economic shocks can induce a certain volatility—relatively rapid oscillations in forecasts and decision variables in response to economic shocks, not the usual exponential decay. Fourth, and more generally, the paper argues that forecasts and forecast errors can be serially correlated in a certain crucial sense. That is, economic shocks which are innovations relative to the underlying driving stochastic processes of the model, and hence are serially uncorrelated, are not necessarily the innovations of the information sets of decision makers. Thus, forecasts of the economic innovations and forecast errors of the driving stochastic processes can display certain cross-correlation and serial correlation properties when expressed in terms of the economic innovations themselves. This is one way to understand the first three properties.

One might well ask at this point whether these results are easily obtained. After all, models in which decision makers forecast the

---

Townsend (1978) and Bray and Kreps (1981). In Arrow and Green (1973), Cyert and DeGroot (1974), DeCanio (1979), Blume and Easley (1982), and Bray (1982), learning can produce convergence to rational expectations equilibrium, but the learning process itself need not be rational. See also Friedman (1979).

[3] As Lucas and Sargent (1978) note, the distinction between sources of impulses and propagation mechanisms is stressed by Frisch (1933) in a classic paper. Lucas and Sargent also identify three propagation mechanisms in the literature. One stems from costs of adjustment, a second from optimal asset accumulation, and a third from the frictions of search theory.

[4] See also Jovanovic (1979), Brunner, Cukierman, and Meltzer (1980), and Kydland and Prescott (1982). A more comprehensive discussion of the relationship between economic models with uncertainty and statistical decision theory is contained in Prescott and Townsend (1980).

forecasts of others have a reputation for being difficult to formulate and analyze.[5] Indeed, the paper was originally motivated by a search for solution techniques which would prove useful in a variety of applications, under a variety of information structures. To that end, the paper adopts a highly stylized model and focuses attention on what problems arise and what formulation-solution techniques are needed as the information structure is varied. As it turns out, all the information variants examined here *are* easy to analyze. That is, techniques do vary somewhat across information variants, but, in one form or another, all techniques make use of standard formulas for conditional means and variances of normal random variables.

With interesting economic time series and tractable solution techniques, one can proceed to ask whether the models studied here have content in a more formal, econometric sense. That is, suppose an econometrician were confronted with data generated from one of the information variants of the basic model, and suppose that market beliefs are unobserved, latent variables. Would he be able to make sense of such time series, that is, interpret the time series relative to theory?[6] More specifically, would the econometrician be able to fit the data to the theoretical model and identify key underlying parameters?[7] As it turns out, in the linear-quadratic, normal random variable model under consideration, many of these questions can be answered in the affirmative.[8]

To summarize, the paper proceeds as follows. Section II describes the basic model, a partial equilibrium model of investment, following Lucas and Prescott (1971), with unobserved permanent and transitory shocks to demand, so there is a nontrivial signal extraction prob-

[5] In Lucas (1975) agents with different histories have different forecasts, and, in effect, these forecasts themselves become an object of speculation. But to simplify this complex picture, Lucas pools the forecasts and makes the average common knowledge. Futia has taken up models with disparate but rational expectations in two papers and emphasizes that Lucas's problem is one of making inferences from *endogenous* time series. Futia (1981) has solved such a problem in a hierarchical structure with informed and uninformed traders. But in general the problem is nonlinear and seems difficult to solve. It is avoided in Futia's (1980) work on capturing what are believed to be the key elements of Lucas's model; agents are endowed with exogenous, distinct information sets but with more information than they can capture from endogenous time series. Chari (1979) studies models with dispersed information and turns up more difficulties. If agents do not see economy-wide output, then there cannot be a finite list of state variables; agents must be concerned with the infinite past in an explicit way. Of course, this paper does not propose a way out of *all* the difficulties associated with the above-mentioned models. It does seek to identify a broad class of models under which such difficulties can be either handled or avoided.

[6] This paper delivers on Sargent's (1981) call for econometric analysis of models in which there are information discrepancies across agents.

[7] The need for this type of analysis, to overcome Lucas's (1976) critique of conventional econometric practice, is eloquently described in Sargent (1981).

[8] There are some caveats on identification, however.

lem; with the Lucas (1967), Gould (1968), Treadway (1969), Mortensen (1973), and Sargent (1979) costly adjustment of capital, following Haavelmo (1960), Eisner and Strotz (1963), and Jorgenson (1963), so there is already present a natural propagation mechanism; and with dispersed markets or islands restricting information flows, following Phelps (1970) and the business cycle literature (see, e.g., King 1978; Barro 1980; Grossman and Weiss 1980), so the "forecasting-the-forecasts" problem can be analyzed. Section III describes the dynamic, stochastic maximization problem confronting the individual (representative) firm in each market and derives the best decision rule. Section IV describes precisely the definition of equilibrium employed throughout the paper, a dynamic linear equilibrium with rational but disparate expectations. Section V generates a law of motion for the aggregate capital stock in each market, allowing one to focus attention entirely on inference problems. Section VI describes a "hierarchical" information variant, in which one market receives price but not quantity information from an informationally self-contained market lower in the hierarchy. This hierarchical system is easily solved with a recursive use of Kalman filtering techniques. The important point is that any firm sees a filtered version of (endogenous) time series generated by linear laws of motion of a finite number of state variables.

Section VII presents the time-series dynamics for the hierarchical structure with some numerical examples. Section VIII describes a symmetric but disparate information structure in which firms in each market see economy-wide average price contaminated with measurement error. Here there is a confounding of laws of motion of observed and unobserved state variables with the inference problems of firms. But a method of undetermined coefficients is developed, a nonlinear but tractable procedure (see also King 1978; Chari 1979; Sargent 1979; Barro 1980; Futia 1981). Moreover, two procedures for forecasting in this setting are described, at least on the assumption that there is full information after a finite number of periods. One procedure makes use of the entire relevant history, the other utilizes the distinction between moving average representations in the space of economic disturbances and representations in the space of disturbances which are innovations relative to agents' information sets (motivated by Hansen and Sargent [1981]). Section IX describes the time-series dynamics for the symmetric structure, again with numerical examples. Finally, Section X discusses how the model might be fitted to data, in principle. This section extends the techniques of Sargent (1978, 1981) and Hansen and Sargent (1980*a*, 1980*b*) and discusses the possibilities and limitations of more conventional econometric procedures, as well as the procedures of Sims (1980*a*, 1980*b*) and Kydland and Prescott (1982).

II. The Basic Model

Imagine an economy in which a single commodity can be produced in each of a set of local "islands." At the level of abstraction adopted here, these islands can be interpreted as local markets within an industry, separate markets within a country, and so on. The essential aspect of the specification is that the commodity itself cannot be moved across islands or markets. Thus one may even suppose the commodity itself varies across islands as if there were separate industries, with all variables expressed in common units. But information can flow among islands, at least to a limited extent, consistent with Phelps's (1970) island paradigm.

More formally, then, suppose there is a set of markets indexed by i, $i = 1, 2, \ldots, I$. (Subsequently I may be either finite or infinite.) There is a continuum of firms (on the unit interval) in each market i, and each firm has a production function of the form

$$y_t^i = f_0 k_t^i, \quad f_0 > 0, \tag{1}$$

where y_t^i is the output of the produced good of the (representative) firm in market i at time t and k_t^i is the capital stock of the firm, chosen at time $t - 1$. Thus, output is linear in the capital stock. Summing over all firms in market i, let Y_t^i and K_t^i denote the corresponding aggregates of market i.

Each market i is confronted with an exogenous, linear demand schedule for the produced good, a schedule which is buffeted by relatively persistent shocks and completely transitory shocks. The transitory shocks are independent across markets and the persistent component is common. Thus, to the extent that the persistent component is unobserved, markets are informationally linked. That is, economic time series in one market may contain information on the persistent component, information which is of interest to firms in another market.

More formally, then, under market clearing,

$$P_t^i = -b_1 Y_t^i + z_t^i, \quad b_1 > 0, \tag{2}$$

where P_t^i is the price of the produced good in market i at time t and z_t^i is a shock to market i demand at time t. Here z_t^i is the sum of a persistent economy-wide component θ_t and a completely transitory market-specific component ϵ_t^i, that is,

$$z_t^i = \theta_t + \epsilon_t^i. \tag{3}$$

The common shock θ_t follows a first-order autoregressive process

$$\theta_t = \rho\theta_{t-1} + v_t, \quad 0 < |\rho| < 1, \tag{4}$$

where v_t, ϵ_t^i are jointly normally distributed, independent among themselves and over time, with mean zero and covariances σ_v^2 and σ_ϵ^2, respectively.

Note that in the formulation above a trend in economic variables is precluded. In fact, the system as specified has a steady state with mean zero. In what follows, however, all variables can be interpreted more generously as deviations from (unspecified) mean values.

III. Individual Firm Maximization

The decision problem confronting each firm in market i at time 0 is to choose a sequence of contingent plans for capital stocks k_t^i so as to maximize discounted expected profits, that is,

$$\max_{\{k_t^i\}_{t=1}^{\infty}} E_0^i \sum_{t=0}^{\infty} \beta^t \left[P_t^i f_0 k_t^i - \frac{f_1}{2}(k_t^i)^2 - \frac{f_2}{2}(k_{t+1}^i - k_t^i)^2 \right], \tag{5}$$
$$f_0, f_2 > 0, f_1 \geqslant 0.$$

Here E_0^i denotes the expectation of (future) prices conditioned on information available to the firm in market i at time $t = 0$, Ω_0^i; parameter β is the discount rate, $0 < \beta < 1$; the term $P_t^i f_0 k_t^i$ represents revenue from sales at time t; the price sequence $\{P_t^i\}_{t=0}^{\infty}$ is taken parametrically and is bounded in mean; the term $(f_1/2)(k_t^i)^2$ induces a kind of "long-run" decreasing returns to scale (constant returns to scale if $f_1 = 0$); and the term $(f_2/2)(k_{t+1}^i - k_t^i)^2$ represents a cost of adjustment, thereby linking the period-by-period decisions of the firm.[9] Such adjustment cost formulations capture the idea that firms do not respond immediately to perceived movements in market demand, giving explicit dynamics to the "static" long-run adjustment stories of standard price theory (see Lucas 1967). Again, a typical firm in market i is to choose a sequence of contingent plans or decision rules,

$$k_{t+1}^i = k_{t+1}^i(\Omega_t^i), \tag{6}$$

where Ω_t^i is the information available to the representative firm at time t in market i and includes, among other things, current and past

[9] Suppose there is a linear technology which maps the amount of capital adjustment $(k_{t+1}^i - k_t^i)^2$ into a labor requirement $(f_2/2)(k_{t+1}^i - k_t^i)^2$. Suppose also the labor can be hired at a fixed wage $\overline{w}$ in any period t. Then $(\overline{w}f_2/2)(k_{t+1}^i - k_t^i)^2$ is the appropriate cost term in the objective function. There is a similar interpretation for the capital maintenance term $(f_2/2)(k_t^i)^2$. Then, dividing through by $\overline{w}$, P_t^i can be interpreted as the relative price of the produced good in terms of labor.

market i aggregate capital stocks and current and past market i prices. Various versions of the model will be considered below by altering these information sets, Ω_t^i. Of crucial importance will be the inclusion of *some* series in other markets.

To solve this decision problem, one can make heavy use of certainty equivalence, as in Holt et al. (1960), following Hansen and Sargent (1980*a*) and described in more detail in Townsend (1981). The end result is a decision rule of the form

$$k_{t+1}^i = \lambda_1 k_t^i + \frac{f_0 \beta \lambda_1}{f_2} \sum_{j=0}^{\infty} (\beta\lambda_1)^j E(P_{t+1+j}^i | \Omega_t^i), \tag{7}$$

where $E(\cdot | \Omega_t^i)$ denotes the expectation conditioned on the information set Ω_t^i. (Note that k_t^i is included in Ω_t^i.) Thus, as one might expect, the capital stock decision of the representative firm in market i is a linear function of the beginning-of-period capital stock and the expectation of all future market-specific prices. As we shall see, these latter expectations are often expressed in terms of a finite number of state variables.

IV. A Definition of Equilibrium

The sequence of expected future prices, which each firm in market i takes as given, is of course determined in an equilibrium, with market-clearing condition (2). Moreover, this sequence will be required to have the rational expectations property that the firms' expectations be consistent with the statistical distribution of prices which the model generates. Note that these expectations across markets need not be identical, as information sets across markets need not be identical. But each firm's expectations are required to be statistically correct in the sense of minimizing the mean square forecast error given the model. Thus, to accommodate rational but disparate expectations, the equilibrium is specified at the level of decision rules. That is, firms act as if they know that prices are determined at each date by market-clearing considerations, that the demand at each date is determined by the schedule (2), and that the supply at each date is determined by a law of motion describing the evolution of the market-wide average capital stock. More formally, consider then (with some repetition of earlier equations) the following definition.

DEFINITION: A dynamic linear equilibrium with rational but possibly disparate expectations is a law of motion for the aggregate capital stock in each market i,

$$K_{t+1}^i = h_1 K_t^i + h_2 M_t^i, \quad i = 1, 2, \ldots, I, \tag{8}$$

where $M_t^i = E(\theta_t | \Omega_t^i)$; a law of motion for the parameter θ,

$$\theta_{t+1} = \rho\theta_t + v_{t+1}; \tag{9}$$

(linear) forecasting formulas for the θ_{t+s} in each market i,

$$E(\theta_{t+s} | \Omega_t^i) = \rho^s M_t^i, \quad s = 0, 1, \ldots, \infty, \quad i = 1, 2, \ldots, I; \tag{10}$$

firm-specific laws of motion for the capital stock in each market i,

$$k_{t+1}^i = g_1 k_t^i + g_2 K_t^i + g_3 M_t^i, \quad i = 1, 2, \ldots, I; \tag{11}$$

and a price equation in each market i,

$$P_t^i = \theta_t + \epsilon_t^i - b_1 f_0 K_t^i, \quad i = 1, 2, \ldots, I; \tag{12}$$

such that (i) for each firm in market i, the individual law of motion (11) is maximizing, that is, is derived from (7) given the aggregate law of motion (8) and market-clearing condition (12) for i and parameter motion (9), under the forecasting formulas (10) for i;[10] (ii) for all firms in market i, the individual law of motion (11) generates the aggregate law (8), that is, $g_3 = h_2$, $g_1 + g_2 = h_1$; (iii) for each firm in market i, the forecasting formulas (10) are statistically correct, that is, based on all available information in market i at time t given the aggregate capital laws and market-clearing conditions (8) and (12) *over all* i and parameter law (9) and given the forecasting formulas (10) *over all* i.

It must be remarked here that this definition remains somewhat imprecise until the information sets Ω_t^i are specified. It is assumed that Ω_t^i contains *at least* $(M_t^i, M_{t-1}^i, \ldots, K_t^i, K_{t-1}^i, \ldots, P_t^i, P_{t-1}^i, \ldots)$. (Note here M_t^i denotes common *market* i forecasts, a state variable from the point of view of the individual firm in market i since it describes the beliefs of other firms.) Equilibrium condition iii allows the possibility that firms in market i see prices and/or capital stocks in other markets, for example, so statistically correct forecasts must use the laws of motion of these variables. Note that these laws may involve the forecasts in other markets, which, under the definition, must have the same statistically correct property. We shall concentrate on this simultaneity in what follows and examine the extent to which it can cause difficulties.

[10] On the right-hand side of (7) is the sequence of expected future prices, $E(P_{t+1}^i | \Omega_t^i)$, $E(P_{t+2}^i | \Omega_t^i)$, and so on. Substituting (8) into (12) at $t + 1$, we get $P_{t+1}^i = \theta_{t+1} + \epsilon_{t+1}^i - b_1 f_0 (h_1 K_t^i + h_2 M_t^i)$. Thus $E(P_{t+1}^i | \Omega_t^i) = \rho M_t^i + 0 - b_1 f_0 h_1 K_t^i - b_0 f_0 h_2 M_t^i$. Similarly, $P_{t+2}^i = \theta_{t+2} + \epsilon_{t+2}^i - b_1 f_0 h_1 h_1 K_t^i - b_1 f_0 h_1 h_2 M_t^i - b_1 f_0 h_2 M_{t+1}^i$. Noting $M_{t+1}^i = E(\theta_{t+1} | \Omega_{t+1}^i)$, $E(M_{t+1}^i | \Omega_t^i) = E(\theta_{t+1} | \Omega_t^i)$ by the law of iterated expectations. Thus $E(P_{t+2}^i | \Omega_t^i) = \rho^2 M_t^i + 0 - b_1 f_0 h_1^2 K_t^i - b_1 f_0 h_1 h_2 M_t^i - b_1 f_0 h_2 \rho M_t^i$. Substituting expressions like these into (7) delivers (11).

V. A Law of Motion for Market Aggregates

The definition of an equilibrium given above may appear somewhat forbidding in a computational sense. But one can make use of an insight of Sargent (1979), that the law of motion of the aggregate capital stock in each market i can be derived without directly calculating firm-specific laws of motion. Moreover, the aggregate law can be computed without being at all specific about the information sets and forecasting. The details of this procedure are described in Townsend (1981). The important result is

$$K_{t+1}^i = \gamma_1 K_t^i + \frac{f_0 \gamma_1 \beta \rho}{f_2(1 - \gamma_1 \beta \rho)} M_t^i, \tag{13}$$

which has the form of (8) in the definition of equilibrium.

Given statistically correct forecasts determining M_t^i and the law of motion (13), one can work backward to compute the firm-specific decision rule (11). But this is never done in what follows. Instead, advantage is taken of the insight that statistically correct forecasts along with (13) completely determine an equilibrium. Having already computed (13), finding an equilibrium here is equivalent to finding statistically correct forecasts, if the resulting price sequences are bounded in mean. These forecasts will be functions of the information structure, as the following sections illustrate.

VI. Equilibrium in a Hierarchical Information Structure[11]

This section focuses on a highly stylized structure in order to illustrate the interplay of forecasts across markets. It is supposed in particular that there are two markets, one which is informationally self-contained and one which receives price but not quantity information from the other. As we shall see in Section VII, this structure induces a special kind of volatility in the second market. The present section is devoted to formulation-analysis issues and makes the point that Kalman filtering techniques easily handle the inference problems of decision makers when the information structure is hierarchical, even if observed time series are endogenous. It will also set the stage for understanding the difficulties which emerge later on when the information structure is symmetric.

More formally, then, suppose there are just two markets, so that $I = 2$. Suppose also that the first market is informationally self-

[11] That hierarchical structures should be tractable with Kalman filtering was suggested to me by Thomas Sargent. See also Prescott and Townsend (1980).

contained, so that $\Omega_t^1 = (K_t^1, K_{t-1}^1, \ldots, P_t^1, P_{t-1}^1, \ldots, M_t^1, M_{t-1}^1, \ldots)$. Finally, suppose for now that $t = 0$ is a starting date, so that histories extend back to $t = 0$, and that firms in market 1 have an initial prior on the parameter θ_0, at the beginning of date $t = 0$, in particular that θ_0 is regarded as normally distributed with means ρM_{-1}^1 and variance $\sigma^2(\theta_0)$. Thus, with prices *and* quantities known at the end of each period t, the entire history of local shocks $(z_0^1, z_1^1, \ldots, z_t^1)$ is completely deducible at the end of each period t. And thus each firm in market 1 can make direct use of that history to form beliefs on θ_t at the end of period t, that is, to form the distribution of θ_t conditioned on that history. As we shall now indicate, there are two ways to derive these conditional distributions.

One way is to make direct, explicit use of the entire history at each date t. Note, in particular, that at the end of period t the representative firm in market 1 has seen

$$\begin{aligned} z_0^1 &= \theta_0 + \epsilon_0^1 \\ z_1^1 &= \rho\theta_0 + v_1 + \epsilon_1^1 \\ &\vdots \\ z_t^1 &= \rho^t\theta_0 + \rho^{t-1}v_1 + \ldots + \rho v_{t-1} + v_t + \epsilon_t^1. \end{aligned}$$

Thus, the representative firm can form a posterior distribution on the parameter vector $\mathbf{x}' = [\theta_0 \; v_1 \ldots v_t]$, conditional on the vector $\mathbf{y}' = [z_0^1 \; z_1^1 \ldots z_t^1]$, given noise vector $\mathbf{w}' = [\epsilon_0^1 \; \epsilon_1^1 \ldots \epsilon_t^1]$, via standard formulas for (conditional) means and variances of normal random variables. That is, suppose we are given that

$$\mathbf{y} = C\mathbf{x} + \mathbf{w}, \quad \mathbf{y}, \mathbf{w} \in R^l, \quad \mathbf{x} \in R^n, \tag{14}$$

and that the prior normal distribution on $\mathbf{x}$ and $\mathbf{w}$ is characterized by means and variances

$$E(\mathbf{x}) = \bar{\mathbf{x}}, \quad E(\mathbf{w}) = \bar{\mathbf{w}}, \tag{15}$$

and

$$E\{(\mathbf{x} - \bar{\mathbf{x}})(\mathbf{x} - \bar{\mathbf{x}})'\} = \Sigma_{xx}, \quad E\{(\mathbf{w} - \bar{\mathbf{w}})(\mathbf{w} - \bar{\mathbf{w}})'\} = \Sigma_{ww}, \tag{16}$$

respectively. Then the posterior distribution of $\mathbf{x}$, conditional on $\mathbf{y}$, has means and variances

$$\hat{\mathbf{x}}(\mathbf{y}) = \bar{\mathbf{x}} + \Sigma_{xx}C'(C\Sigma_{xx}C' + \Sigma_{ww})^{-1}(\mathbf{y} - C\bar{\mathbf{x}} - \bar{\mathbf{w}}) \tag{17}$$

and

$$E_{\mathbf{y}}\{[\mathbf{x} - \hat{\mathbf{x}}(\mathbf{y})][\mathbf{x} - \hat{\mathbf{x}}(\mathbf{y})]'\} = \Sigma_{xx} - \Sigma_{xx}C'(C\Sigma_{xx}C' + \Sigma_{ww})^{-1}C\Sigma_{xx}, \tag{18}$$

respectively (see Bertsekas 1976). Here, then,

$$\bar{\mathbf{x}}' = [\rho M^{1}_{-1}\ 0 \ldots 0] \qquad \bar{\mathbf{w}}' = [0\ 0 \ldots 0]$$

$$\Sigma_{xx} = \begin{bmatrix} \sigma^2(\theta_0) & 0 & \ldots & 0 \\ 0 & \sigma_v^2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \sigma_v^2 \end{bmatrix} \qquad \Sigma_{ww} = \begin{bmatrix} \sigma_\epsilon^2 & 0 & \ldots & 0 \\ 0 & \sigma_\epsilon^2 & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & & \ldots & \sigma_\epsilon^2 \end{bmatrix}.$$

We might note in passing that in this formulation the firm is learning about the parameter θ_0 at each date $t \geqslant 0$, and thus with more and more observations the posterior distribution on θ_0 should tend to become degenerate with mean θ_0 and variance zero. (In this regard the updating may be viewed as sequential, with only the most recent observations employed explicitly.) Similarly, the firm is learning about the parameter v_1 at each date $t \geqslant 1$, and so on. Thus, in general, the firm should learn parameters in the arbitrarily distant past arbitrarily well. But the variable of interest, $\theta_t = \Sigma_{j=0}^{t-1}\rho^j v_{t-j} + \rho^t\theta_0$, is constantly buffeted by new shocks v_t which are not well understood. Indeed, the posterior distribution for θ_t may now be computed directly, at the end of period t, since it is just a linear combination of a finite number of unknown parameters $(v_t, v_{t-1}, \ldots, \theta_0)$. One might expect that posterior to have a variance which reflects the above-mentioned degree of uncertainty and a mean which moves around with the realized history of the shocks v_t.

Finally, we might note that in substituting for $\mathbf{y}$ in (17) from (14) it becomes clear that the mean forecast of $\mathbf{x}$, $\hat{\mathbf{x}}(\mathbf{y})$, is a linear function of the random variables, $\mathbf{x}$ and $\mathbf{w}$. In the application here, then, the mean forecast of the parameter v_t is a linear function of the random variables $(v_t, v_{t-1}, \ldots, v_1)$ and $(\epsilon_t^1, \epsilon_{t-1}^1, \ldots, \epsilon_0^1)$. It is already apparent, then, that this learning mechanism is the cause of some persistence. It converts serially uncorrelated economic shocks, the v_t and the ϵ_t^1, into serially correlated forecasts of the shocks v_t. That is, $E(v_t|\Omega_t^1)$ is correlated with $E(v_{t-1}|\Omega_{t-1}^1)$, as both contain v_{t-1}, for example, except in the special case of full information, $\sigma_\epsilon^2 = 0$. By the same token, then, one anticipates that the mean forecasts of the variables, θ_t—say, $E(\theta_t|\Omega_t^1)$, $E(\theta_{t-1}|\Omega_{t-1}^1)$, and so on—will be serially correlated with a persistence over and above that present in θ_t itself. That is, the forecast error $[E(\theta_t|\Omega_t^1) - \theta_t]$ will be serially correlated, a moving average of past random variables v_t and ϵ_t^1. This will be made clear in a moment.

A distinct disadvantage of the method described above for computing the conditional distribution of θ_t is that the state vector $\mathbf{x}_t$ increases in dimension with the length of history (back to $t = 0$)—there are more and more innovations. An alternative way to compute the conditional distribution of θ_t exploits a recursive algorithm developed by Kalman (1960). That is, suppose we have under consideration a finite dimensional state vector $\mathbf{x}_t$, following linear law of motion

$$\mathbf{x}_{t+1} = A\mathbf{x}_t + \mathbf{v}_{t+1}, \quad \mathbf{v}_t, \mathbf{x}_t \in R^n, \tag{19}$$

with period-by-period observations

$$\mathbf{y}_t = C\mathbf{x}_t + \mathbf{w}_t, \quad \mathbf{w}_t, \mathbf{y}_t \in R^l. \tag{20}$$

Thus, the information set Ω_t is $(\mathbf{y}_0, \mathbf{y}_1, \ldots, \mathbf{y}_t)$. Suppose also that

$$E(\mathbf{v}_t) = 0, \quad E(\mathbf{w}_t) = 0, \quad E[\mathbf{v}_t\mathbf{v}_t'] = M, \quad E[\mathbf{w}_t\mathbf{w}_t'] = N, \tag{21}$$

with $\mathbf{w}_t$ and $\mathbf{v}_t$ independent normal random variables. Then, by using the fact that

$$E(\mathbf{x}_t|\Omega_t) = E(\mathbf{x}_t|\Omega_{t-1}) + E\{\mathbf{x}_t|[\mathbf{y}_t - E(\mathbf{y}_t|\Omega_{t-1})]\}, \tag{22}$$

standard formulas for conditional means and variances of normal random variables, like (17)–(18), deliver a law of motion for conditional forecasts

$$E(\mathbf{x}_t|\Omega_t) = AE(\mathbf{x}_{t-1}|\Omega_{t-1}) + \dot{\Sigma}_{t|t}C'N^{-1}[\mathbf{y}_t - CAE(\mathbf{x}_{t-1}|\Omega_{t-1})] \tag{23}$$

and law of motion for variance-covariance matrix

$$\Sigma_{t|t-1} \equiv E\{[\mathbf{x}_t - E(\mathbf{x}_t|\Omega_{t-1})][\mathbf{x}_t - E(\mathbf{x}_t|\Omega_{t-1})]'|\Omega_{t-1}\},$$

namely,

$$\Sigma_{t+1|t} = A[\Sigma_{t|t-1} - \Sigma_{t|t-1}C'(C\Sigma_{t|t-1}C' + N)^{-1}C\Sigma_{t|t-1}]A' + M. \tag{24}$$

The matrix $\Sigma_{t|t}$ in (23) is defined by

$$\begin{aligned}\Sigma_{t|t} &\equiv E\{[\mathbf{x}_t - E(\mathbf{x}_t|\Omega_t)][\mathbf{x}_t - E(\mathbf{x}_t|\Omega_t)]'\,|\,\Omega_t\} \\ &= \Sigma_{t|t-1} - \Sigma_{t|t-1}C'(C\Sigma_{t|t-1}C' + N)^{-1}C\Sigma_{t|t-1}.\end{aligned} \tag{25}$$

Moreover, under specified regularity conditions, the variance-covariance matrix $\Sigma_{t|t-1}$ in (24) converges to some constant Σ independent of the variance of the initial prior. In this case we can define a constant matrix $\bar{\Sigma}$ from (25) by substituting Σ for $\Sigma_{t|t-1}$, and thus coefficients in the linear law of motion (23) are no longer time dependent. Moreover, one can interpret this solution as having supposed an *infinite* history so that the steady state has been obtained at $t = 0$. This supposition is useful for analytic and econometric purposes. To reiterate, then, under the Kalman filtering algorithm, contemporary mean beliefs may be updated from past mean beliefs and the contemporary observation $\mathbf{y}_t$ alone—mean beliefs capture all that is necessary for forecasting from the infinite history.

Returning to the present application, the inference problem faced by firms in the first market, note again that

$$z_t^1 = \theta_t + \epsilon_t^1 \tag{26}$$

and

$$\theta_{t+1} = \rho\theta_t + v_{t+1}. \tag{27}$$

Thus let $\mathbf{x}_t = \theta_t$, $\mathbf{y}_t = z_t^1$, $A = \rho$, $C = 1$, $M = \sigma_v^2$, and $N = \sigma_\epsilon^2$. So, with appropriate substitutions and manipulations, (23) yields

$$M_t^1 = \alpha_0 M_{t-1}^1 + \alpha_1 z_t^1, \tag{28}$$

where $\alpha_0 = \rho[1 - (\bar{\Sigma}/\sigma_\epsilon^2)]$ and $\alpha_1 = (\bar{\Sigma}/\sigma_\epsilon^2)$. (Note here that the coefficients α_0 and α_1 can be expressed entirely in terms of ρ and variances σ_ϵ^2, σ_v^2.)

Now we may note also that, with an infinite past,

$$M_t^1 = E(\theta_t|\Omega_t^1) = \frac{\alpha_1 v_t}{(1 - \alpha_0 L)(1 - \rho L)} + \frac{\alpha_1 \epsilon_t^1}{(1 - \alpha_0 L)} \tag{29}$$

and

$$[E(\theta_t|\Omega_t^1) - \theta_t] = \frac{\alpha_1 v_t}{(1 - \alpha_0 L)(1 - \rho L)} + \frac{\alpha_1 \epsilon_t^1}{(1 - \alpha_0 L)} - \frac{v_t}{(1 - \rho L)}. \tag{30}$$

Thus the forecast error of θ_t is a linear combination of current and past v_t and ϵ_t^1 and is serially correlated, as was anticipated above.

With this artillery we now turn to the inference problem of firms in a second market. Suppose in particular that firms in some market 2 see price in market 1, P_t^1, in addition to the shock to their own market, z_t^2. That is, suppose $\Omega_t^2 = (K_t^2, K_{t-1}^2, \ldots, P_t^2, P_{t-1}^2, \ldots, M_t^2, M_{t-1}^2, \ldots, P_t^1, P_{t-1}^1, \ldots)$. Thus, firms in market 2 do *not* see the aggregate capital stock (output) in the first market and thus do not get a second direct observation on θ_t. But firms in market 2 do see a filtered version of the vector of state variables in market 1: the shock θ_t; the mean forecast in market 1, M_t^1; and the capital stock in market 1, K_t^1. That is, letting $\mathbf{x}_t' = [K_t^1 \; M_t^1 \; \theta_t]$, note that

$$\begin{bmatrix} K_{t+1}^1 \\ M_{t+1}^1 \\ \theta_{t+1} \end{bmatrix} = \begin{bmatrix} h_1 & h_2 & 0 \\ 0 & \alpha_0 & \alpha_1\rho \\ 0 & 0 & \rho \end{bmatrix} \begin{bmatrix} K_t^1 \\ M_t^1 \\ \theta_t \end{bmatrix} + \begin{bmatrix} 0 \\ \alpha_1 v_{t+1} + \alpha_1 \epsilon_{t+1}^1 \\ v_{t+1} \end{bmatrix} \tag{31}$$

or

$$\mathbf{x}_{t+1} = A\mathbf{x}_t + \mathbf{v}_{t+1}^0. \tag{32}$$

Then firms in market 2 see

$$\begin{bmatrix} P_t^1 \\ z_t^2 \end{bmatrix} = \begin{bmatrix} -b_1 f_0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} K_t^1 \\ M_t^1 \\ \theta_t \end{bmatrix} + \begin{bmatrix} \epsilon_t^1 \\ \epsilon_t^2 \end{bmatrix} \tag{33}$$

or

$$\mathbf{y}_t = C\mathbf{x}_t + \mathbf{w}_t^0. \tag{34}$$

The point is that (34) and (32) represent a system to which the Kalman filtering algorithm may again be applied directly. In this application, then, firms in market 2 attempt to keep track of the capital stock and mean forecasts in the first market, as well as the persistent shock θ_t. Again, all these variables are unobserved.

Now, in this case a less trivial Kalman filtering theorem[12] ensures the convergence of the variance-covariance matrix of the conditional distribution of the state vector $\mathbf{x}_t$, defined in (31) and (32), with means satisfying the recursive relationship of (23) or

$$\begin{aligned} E\left(\begin{bmatrix} K_{t+1}^1 \\ M_{t+1}^1 \\ \theta_{t+1} \end{bmatrix} \middle| \Omega_{t+1}^2 \right) &= (A - \bar{\Sigma} C' N^{-1} CA) E\left(\begin{bmatrix} K_t^1 \\ M_t^1 \\ \theta_t \end{bmatrix} \middle| \Omega_t^2 \right) \\ &\quad + \bar{\Sigma} C' N^{-1} \left(\begin{bmatrix} P_{t+1}^1 \\ z_{t+1}^2 \end{bmatrix} \right). \end{aligned} \tag{35}$$

[12] This theorem (see Kwakernaak and Sivan 1972, p. 535) requires that the variance-covariance matrix M be written $M = GM^*G'$, where $M^* \geqslant \alpha I$, $N \geqslant \beta I$, M^* positive definite, where α, β are positive constants, and the rows of

$$\begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{bmatrix}$$

and columns of $(G, AG, A^2G, \ldots, A^{n-1}G)$ span the n-dimensional space. For the application here, let G and M^* be defined by

$$G = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad M^* = \begin{bmatrix} \alpha_1^2\sigma_v^2 + \alpha_1^2\sigma_\epsilon^2 & \alpha_1\sigma_v^2 \\ \alpha_1\sigma_v^2 & \sigma_v^2 \end{bmatrix},$$

where C, A, and N are defined implicitly in the text.

(Note that here the matrices of coefficients weighting the prior mean and the current observation depend on the parameters of technology and demand, through the definitions of A and C in [31]–[32] and [33]–[34], respectively.) Substituting (33) at $t + 1$ and then (31) into (35) yields

$$E\left(\begin{bmatrix} K_{t+1}^1 \\ M_{t+1}^1 \\ \theta_{t+1} \end{bmatrix} \middle| \Omega_{t+1}^2\right) = (A - \bar{\Sigma}C'N^{-1}CA)E\left(\begin{bmatrix} K_t^1 \\ M_t^1 \\ \theta_t \end{bmatrix} \middle| \Omega_t^2\right)$$
$$+ \bar{\Sigma}C'N^{-1}CA\begin{bmatrix} K_t^1 \\ M_t^1 \\ \theta_t \end{bmatrix} + \bar{\Sigma}C'N^{-1}C\mathbf{v}_{t+1}^0 + \bar{\Sigma}C'N^{-1}\mathbf{w}_{t+1}^0. \tag{36}$$

Moreover, substitution of (31)–(32) into (36) and the obvious manipulation yield

$$E\left(\begin{bmatrix} K_t^1 \\ M_t^1 \\ \theta_t \end{bmatrix} \middle| \Omega_t^2\right) - \begin{bmatrix} K_t^1 \\ M_t^1 \\ \theta_t \end{bmatrix} \tag{37}$$
$$= [I - (A - \bar{\Sigma}C'N^{-1}CA)L]^{-1}[\bar{\Sigma}C'N^{-1}CA(I - AL)^{-1}\mathbf{v}_{t-1}^0$$
$$+ \bar{\Sigma}C'N^{-1}C\mathbf{v}_t^0 + \bar{\Sigma}C'N^{-1}\mathbf{w}_t^0] - (I - AL)^{-1}\mathbf{v}_t^0.$$

This is the market 2 analogue of (30), and it is again apparent that the forecast error in market 2 will be serially correlated as a moving average of shocks ϵ_t^1, ϵ_t^2, and v_t. Note also that forecast errors will be serially correlated across markets 1 and 2 due to the influence of the v_t *and* the ϵ_t^1 terms.

In summary, then, this economy with two markets has laws of motion

$$\begin{bmatrix} K_{t+1}^2 \\ E(K_{t+1}^1 | \Omega_{t+1}^2) \\ E(M_{t+1}^1 | \Omega_{t+1}^2) \\ M_{t+1}^2 \\ K_{t+1}^1 \\ M_{t+1}^1 \\ \theta_{t+1} \end{bmatrix} = \begin{bmatrix} h_1 & (0 \quad 0 \quad h_2) & 0 \\ 0 & A - \bar{\Sigma}C'N^{-1}CA & \bar{\Sigma}C'N^{-1}CA \\ 0 & 0 & A \end{bmatrix} \tag{38}$$

$$\cdot \begin{bmatrix} K_t^2 \\ E(K_t^1 \mid \Omega_t^2) \\ E(M_t^1 \mid \Omega_t^2) \\ M_t^2 \\ K_t^1 \\ M_t^1 \\ \theta_t \end{bmatrix} + \begin{bmatrix} 0 \\ \\ \bar{\Sigma} C' N^{-1} C \mathbf{v}_{t+1}^0 + \bar{\Sigma} C' N^{-1} \mathbf{w}_{t+1}^0 \\ \\ \\ \mathbf{v}_{t+1}^0 \\ \\ \end{bmatrix} .$$

The recursive manner in which this hierarchical system is determined is apparent from (38): apart from the shocks, motion in the first market determines beliefs in the second market which determine motion in the second market. Note also that this system is stable because the matrices along the diagonal are stable. So all variables are bounded in mean. Thus, prices in each market are bounded in mean, and the existence of an equilibrium is established.

VII. Time-Series Dynamics for the Hierarchical Structure

Now suppose an econometrician were given time series of capital stocks and prices in both markets. We might ask what implications the theory has for these series, at least for various possible parameter values. One way to answer that question is to look at the implied covariance structure of the data, and indeed that route will be taken in the discussion of formal estimation in Section X. Here we wish to characterize the data, perhaps in a more revealing way, by examining the response of the system from an initial steady state (of zero) to a positive, one-standard-deviation impulse in specified economic shocks at date 1, holding all other economic shocks at zero, following the suggestions of Lucas (1975) and Sims (1980*a*, 1980*b*).

Consider first the capital stock series in the first market. Note in particular that (31) can be written after repeated substitutions as

$$\begin{aligned} &(1 - \rho L)(1 - h_1 L)(1 - \alpha_0 L) K_{t+1}^1 \\ &= h_2 \alpha_1 \sigma_v v_t^* + h_2 \alpha_1 (1 - \rho L) \sigma_\epsilon \epsilon_t^{*1}, \end{aligned} \tag{39}$$

where the * shocks are normalized to have unit variance. Equation (39) completely describes the statistical properties of the capital stock

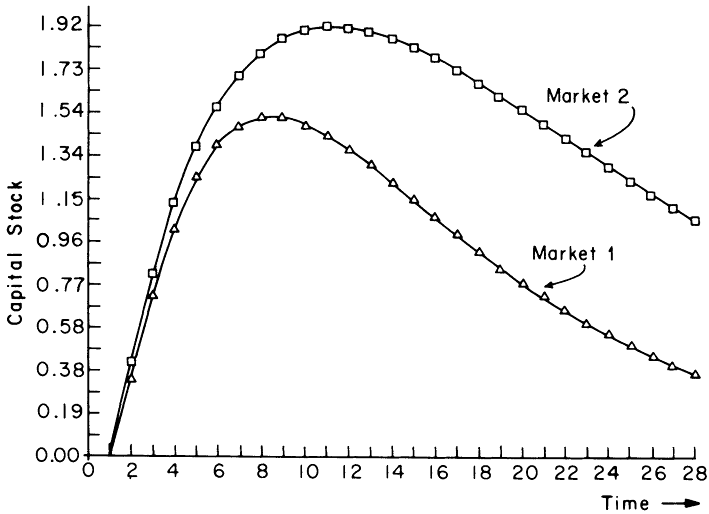

FIG. 1.—Hierarchical structure response to persistent common shock (high adjustment costs).

in market 1. These properties become especially clear when a comparison is made to the model with no learning (full information), namely,

$$(1 - \rho L)(1 - h_1 L)K^1_{t+1} = h_2 \sigma_v v^*_t. \tag{40}$$

In both (39) and (40), intrinsic persistence will be higher and the capital stock series will be smoother the higher is the serial correlation parameter ρ in the driving stochastic process θ_t, since demand stays high longer, and the higher is the ratio of adjustment costs to revenue (high h_1 is associated with high f_2/f_0) because it is costly to adjust more quickly. And the effects of economic shocks can not only persist in (39) and (40), they can also cumulate in the sense of Lucas (1975). That is, the full effect of the shock is not immediate—there can be a rising portion in the response function (see fig. 1). But we might note that in (39), relative to (40), there is additional persistence induced by the learning, persistence which varies directly with the noise in current observations (high ratio of variances σ_ϵ/σ_v is associated with high α_0). On the other hand, (39) allows a response to entirely transitory shocks, the ϵ^1_t, as these are partially mistaken for relatively persistent shocks v_t. This adds more volatility to the capital stock series.

The statistical process for price in the first market is linearly related to the statistical process for the capital stock via the market-clearing hypothesis. There is in addition a stochastic intercept, representing the period-by-period shocks to demand. That is,

$$P_t^1 = \sum_{j=0}^{\infty} \rho^j v_{t-j} + \epsilon_t^1 - b_1 f_0 K_t^1. \tag{41}$$

The statistical process for the capital stock in the second market is more complicated than in (39) since, as noted, (41) is taken into account in the inference problems of firms in the second market.[13] It may be noted again that firms in the second market will respond to the relatively persistent shocks v_t as well as to the transitory shocks ϵ_t^1 and ϵ_t^2 in both markets. In this sense, the second market is more volatile than the first. But possessing more observations, firms in the second market are relatively better informed. Thus market 2 responds more readily to persistent common shocks (see table 1 and fig. 1) and less readily to transitory local shocks in the first market (see table 3 and fig. 3). In this sense, then, the statistical process for K_t^2 moves away from the limited information system (39) toward the full-information system (40).

There are, however, some interesting dynamics in the second market which have no parallel in either the full-information system (40) or the self-contained system with learning (39). Consider the response of the second market to transitory shocks in the first market. Under high adjustment costs (see table 2), the best forecast in the second market of the parameter θ_t (which remains zero) first increases, then goes negative, then becomes positive again, and then seems to decrease toward zero. Thus, forecast errors on θ_t display (sharply) damped oscillations around the steady state. This "volatility" may be contrasted with the usual exponential decay of forecast errors in the first market under (39). Clearly, firms in the second market are drawing inferences from the price in the first market, which is first positive and then negative. However, the negative inferences are eventually corrected. (Recall that all variables should be interpreted as variations from steady-state values.)

When adjustment costs are low (see table 3), one again observes the strikingly different behavior forecasts in the two markets (see fig. 2). Here forecasts in the second market remain negative enough to pull the capital stock in the second market below its steady-state value (see fig. 3).

As between tables 2 and 3, the qualitatively different nature of the response to a transitory shock, for high and low adjustment costs, respectively, may be attributed in part to the different steady-state

[13] From system (35), let $D = A - \bar{\Sigma}C'N^{-1}CA$ and $F = \bar{\Sigma}C'N^{-1}$. Then $M_t^2 = E(\theta_t|\Omega_t^2)$ is determined by the third row of (35), namely, $E(x_t|\Omega_t^2) = (I - DL)^{-1}F(P_t^1 z_t^2)'$. The law of motion (13) for K_t^2 then completes the picture.

TABLE 1

HIERARCHICAL STRUCTURE RESPONSE TO PERSISTENT COMMON SHOCK
(High Adjustment Costs)

	Mean Beliefs		Capital Stock		
Time	In Island 1	In Island 2	In Island 1	In Island 2	Price in Island 1
1	.597406781	.723919251	.000000000	.000000000	1.000000000
2	.754126830	.830056978	.344032644	.416888227	.831193471
3	.757145206	.815213292	.715557441	.818848934	.666888512
4	.709848917	.786750604	1.021045993	1.138934858	.524790801
5	.649160913	.759951506	1.243569749	1.384238931	.407386050
6	.587975734	.734319136	1.390551374	1.569361062	.312379725
7	.530529996	.708619299	1.475485134	1.705951679	.236343973
8	.477966812	.682456960	1.511843409	1.802825166	.175928218
9	.430347608	.655889972	1.511299207	1.866960475	.128207369
10	.387377153	.629113683	1.483431489	1.904096755	.090734191
11	.348662738	.602335458	1.435901784	1.919038704	.061498083
12	.313804907	.575738991	1.374747835	1.915833951	.038861029
13	.282427475	.549478622	1.304675924	1.897897531	.021494352
14	.254185836	.523681405	1.229317171	1.868110381	.008323149
15	.228767654	.498450495	1.151441842	1.828901060	−.001520444
16	.205891034	.473868351	1.073134982	1.782314506	−.008735864
17	.185301983	.449999559	.995938937	1.730070100	−.013885769
18	.166771804	.426893300	.920968428	1.673610764	−.017421869
19	.150094631	.404585512	.849003123	1.614144522	−.019705989
20	.135085170	.383100788	.780561909	1.552679721	−.021027210
21	.121576654	.362454051	.715962309	1.490054926	−.021615807
22	.109418989	.342652015	.655367885	1.426964325	−.021654588
23	.098477090	.323694470	.598825928	1.363979332	−.021288095
24	.088629381	.305575409	.546297307	1.301567008	−.020630080
25	.079766443	.288284008	.497680019	1.240105746	−.019769561
26	.071789799	.271805480	.452827655	1.179898682	−.018775732
27	.064610819	.256121823	.411563805	1.121185137	−.017701942
28	.058149737	.241212463	.373693200	1.064150407	−.016588903
29	.052334763	.227054816	.339010256	1.008934138	−.015467288
30	.047101287	.213624767	.307305551	.955637496	−.014359000

NOTE.—Assigned parameters: ρ = .900, b_1 = 1.000, β = .960, f_0 = .200, f_1 = .000, f_2 = .800, σ_v^2 = 1.000, σ_ϵ^2 = 1.000, σ_η^2 = 1.000. Computed parameters: h_1 = .818, h_2 = .576, α_0 = .362, α_1 = .597.

$\bar{\Sigma}$ matrix (island 1)	$\bar{\Sigma}$ matrix (island 2)		
.597406781	1.362374115	.312149664	.169669081
	.312149664	.616612548	.265536615
	.169669081	.265536615	.378926534

covariance matrix of the conditional distribution of beliefs in the second market, reported in the tables as the sigma ($\bar{\Sigma}$) matrix, a matrix which is quite sensitive to the adjustment cost parameters f_0 and f_2. In neither case, however, are there any zero entries in that matrix. In particular, under the assumed parameter specifications, firms in market 2 never learn with certainty the capital stock in market 1. This is what generates the interesting dynamics. That is, the space spanned by current and past prices in market 1 is not equivalent with the space spanned by current and past market 1 shocks, the z_t^1.

TABLE 2

HIERARCHICAL STRUCTURE RESPONSE TO TRANSITORY MARKET 1 SHOCK (High Adjustment Costs)

	Mean Beliefs		Capital Stock		
Time	In Island 1	In Island 2	In Island 1	In Island 2	Price in Island 1
1	.597403342	.344998702	.000000000	.000000000	1.000000000
2	.216461330	.064259251	.344030663	.198676713	−.068806133
3	.078431947	.005355471	.405926789	.199439029	−.081185358
4	.028418796	−.002231772	.377043768	.166140967	−.075408754
5	.010297181	−.001425525	.324628249	.134547903	−.064925650
6	.003731050	−.000429915	.271338681	.109182420	−.054267736
7	.001351897	.000000616	.223989043	.089017519	−.044797809
8	.000489842	.000131948	.183906943	.072779082	−.036781389
9	.000177488	.000157482	.150640280	.059578531	−.030128056
10	.000064310	.000153722	.123262325	.048800761	−.024652465
11	.000023302	.000143964	.100813547	.039986932	−.020162709
12	.000008443	.000133894	.082436314	.032775325	−.016487263
13	.000003059	.000124516	.067402943	.026873477	−.013480589
14	.000001108	.000115860	.055108896	.022042857	−.011021779
15	.000000402	.000107841	.045056436	.018088467	−.009011287
16	.000000146	.000100389	.036837362	.014850828	−.007367472
17	.000000053	.000093456	.030117485	.012199516	−.006023497
18	.000000019	.000087002	.024623410	.010027870	−.004924682
19	.000000007	.000080994	.020131559	.008248664	−.004026312
20	.000000003	.000075401	.016459115	.006790565	−.003291823
21	.000000001	.000070194	.013456605	.005595235	−.002691321
22	.000000000	.000065346	.011001819	.004614961	−.002200364
23	.000000000	.000060834	.008994840	.003810720	−.001798968
24	.000000000	.000056633	.007353980	.003150592	−.001470796
25	.000000000	.000052722	.006012449	.002608467	−.001202490
26	.000000000	.000049081	.004915643	.002162985	−.000983129
27	.000000000	.000045691	.004018920	.001796672	−.000803784
28	.000000000	.000042536	.003285779	.001495232	−.000657156
29	.000000000	.000039599	.002686379	.001246963	−.000537276
30	.000000000	.000036864	.002196323	.001042293	−.004392646

NOTE.—Assigned parameters: ρ = .900, b_1 = 1.000, β = .960, f_0 = .200, f_1 = .000, f_2 = .800, σ_v^2 = 1.000, σ_ϵ^2 = 1.000, σ_η^2 = 1.000. Computed parameters: h_1 = .818, h_2 = .576, α_0 = .362, α_1 = .597.

$\bar{\Sigma}$ matrix (island 1)	$\bar{\Sigma}$ matrix (island 2)		
.597403342	1.361775834	.312092067	.169608920
	.312092067	.616601903	.265529661
	.169608920	.265529661	.378920486

VIII. Equilibrium in a Symmetric Information Structure with Confounding

In the hierarchical structure, firms in the second market are forming inferences on the basis of an endogenous time series, prices in the first market. These prices depend on the aggregate capital stock in the first market, which the firms do not see. The law of motion of this capital stock is known, however, though it depends on beliefs in the first market, which the firms in the second market do not see. The law

TABLE 3

HIERARCHICAL STRUCTURE RESPONSE TO TRANSITORY MARKET 1 SHOCK
(Low Adjustment Costs)

	Mean Beliefs		Capital Stock		
Time	In Island 1	In Island 2	In Island 1	In Island 2	Price in Island 1
1	.597406781	.275942299	.000000000	.000000000	1.000000000
2	.216460727	−.029342976	.486052485	.224507730	−.388841988
3	.078431059	−.012405283	.276674876	.022575855	−.221339901
4	.028418231	−.005139347	.121054325	−.005422166	−.096843460
5	.010296888	−.004498733	.048166658	−.005303209	−.038533326
6	.003730912	−.003886124	.018343005	−.004757391	−.014674404
7	.001351836	−.003262932	.006830551	−.004146043	−.005464441
8	.000489816	−.002739521	.002513062	−.003512528	−.002010450
9	.000177477	−.002301869	.000918455	−.002955607	−.000734764
10	.000064306	−.001934158	.000334419	−.002484308	−.000267535
11	.000023300	−.001625151	.000121509	−.002087628	−.000097207
12	.000008442	−.001365512	.000044097	−.001754148	−.000035277
13	.000003059	−.001147354	.000015992	−.001473909	−.000012794
14	.000001108	−.000964049	.000005797	−.001238435	−.000004638
15	.000000402	−.000810030	.000002101	−.001040580	−.000001681
16	.000000146	−.000680617	.000000761	−.000874334	−.000000609
17	.000000053	−.000571880	.000000276	−.000734648	−.000000221
18	.000000019	−.000480515	.000000100	−.000617278	−.000000080
19	.000000007	−.000403747	.000000036	−.000518660	−.000000029
20	.000000003	−.000339243	.000000013	−.000435798	−.000000011
21	.000000001	−.000285044	.000000005	−.000366173	−.000000004
22	.000000000	−.000239505	.000000002	−.000307673	−.000000001
23	.000000000	−.000201241	.000000001	−.000258518	−.000000000
24	.000000000	−.000169090	.000000000	−.000217216	−.000000000
25	.000000000	−.000142076	.000000000	−.000182513	−.000000000
26	.000000000	−.000119377	.000000000	−.000153354	−.000000000
27	.000000000	−.000100305	.000000000	−.000128854	−.000000000
28	.000000000	−.000084280	.000000000	−.000108268	−.000000000
29	.000000000	−.000070815	.000000000	−.000090971	−.000000000
30	.000000000	−.000059502	.000000000	−.000076437	−.000000000

NOTE.—Assigned parameters: $\rho = .900$, $b_1 = 1.000$, $\beta = .960$, $f_0 = .800$, $f_1 = .000$, $f_2 = .200$, $\sigma_{\nu}^2 = 1.000$, $\sigma_{\epsilon}^2 = 1.000$, $\sigma_{\eta}^2 = 1.000$. Computed parameters: $h_1 = .207$, $h_2 = .814$, $\alpha_0 = .362$, $\alpha_1 = .597$.

$\bar{\Sigma}$ matrix (island 1)	$\bar{\Sigma}$ matrix (island 2)		
.597406781	.490756468	.316838956	.211021629
	.316838956	.689790614	.347483817
	.211021629	.347483817	.444759602

of motion of these beliefs is known, however, and depends on the underlying demand shock θ_t, with known law of motion. In summary, firms in the second market are forming expectations on the expectations in the first market, but the laws of motion of the market 1 expectations are well defined and can be expressed in terms of a finite number of state variables. Thus, the Kalman filter can be applied. In fact, we could well add on another layer to the hierarchical information structure, supposing that firms in some third market see prices in the second market and so on, building on system (38).

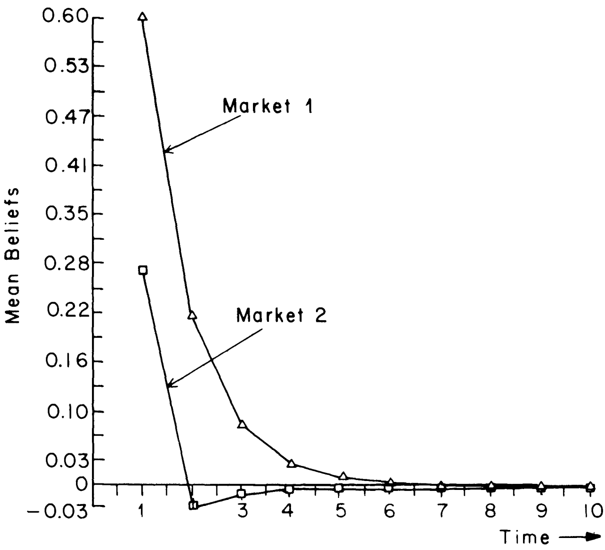

FIG. 2.—Hierarchical structure response to transitory market 1 shock (low adjustment costs).

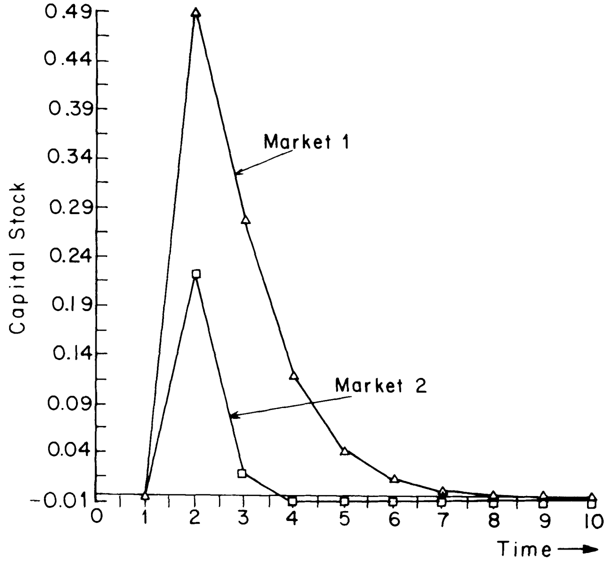

FIG. 3.—Hierarchical structure response to transitory market 1 shock (high adjustment costs).

One might well ask, though, what would happen in a symmetric information structure rather than a hierarchical one. Thus, suppose firms in market 1 (firms 1) see P_t^2. Then to make inferences, firms 1 must have expectations on M_t^2, say, $E_t^1(M_t^2)$. But firms in market 2 (firms 2) see P_t^1 also. So firms 2 must have expectations on M_t^1. But M_t^1 is related to $E_t^1(M_t^2)$. Thus firms 2 need to have expectations on M_t^1 *and* $E_t^1(M_t^2)$, say, $E_t^2(M_t^1)$ and $E_t^2E_t^1(M_t^2)$. Returning to market 1, it becomes apparent that there is an infinite regress problem here (see also Townsend 1978). In the space of mean beliefs, at least, we seem to have need of an infinite number of state variables, but then it is no longer possible to make use of *standard* Kalman filtering formulas. This section of the paper is devoted to a discussion of this problem and possible solutions.

The infinite regress problem emerges also in a model with an infinity of markets where it is somewhat easier to analyze. Suppose firms in each market see economy-wide average price $\bar{P}_t$ with error η_t, that is, see $P_t = \bar{P} + \eta_t$ where $\bar{P}_t = \lim_{I\to\infty}\Sigma_{i=1}^I(P_t^i/I)$ and η_t is distributed normally with mean zero and variance σ_η^2, independent of v_t and of ϵ_t^i, for all i. Of course, averaging over markets, $\bar{P}_t = -b_1 f_0 \bar{K}_t + \theta_t$ where economy-wide average capital stock is $\bar{K}_t = \lim_{I\to\infty}\Sigma_{i=1}^I(K_t^i/I)$ and the economy-wide average local shock is zero, that is, $\lim_{I\to\infty}\Sigma_{i=1}^I(\epsilon_t^i/I) = 0$ by the law of large numbers. Then firms in each market i attempt to make inferences about the economy-wide common demand shock θ_t on the basis of noisy economy-wide price data P_t and the market-specific shocks z_t^i. But to do this they must in effect make inferences about the economy-wide average capital stock (output), which they do not see. The law of motion of the latter is known to depend on economy-wide average mean beliefs. So firms in each market would need to make inferences about such beliefs, and the infinite regress is started.

The general problem is that laws of motion are needed for the inference problems, but the inferences in turn determine the laws of motion. This section captures this simultaneity by returning to the space of contemporary and past economic innovations. The model is then written down in the form of moving-average representations, and the method of undetermined coefficients is used to handle the simultaneity.

To begin, let $\bar{M}_t = \lim_{I\to\infty}\Sigma_{i=1}^I(M_t^i/I)$ denote economy-wide average mean beliefs. Now suppose economy-wide average mean beliefs follow

$$\bar{M}_t = C(L)v_t + D(L)\eta_t, \tag{42}$$

where $C(L)$ and $D(L)$ are polynomials in the lag operator with coefficients yet to be determined. These polynomials need not be of finite order. Then, averaging across markets, the law of motion for the economy-wide average capital stock is

$$\bar{K}_t = h_1\bar{K}_{t-1} + h_2\bar{M}_{t-1}. \tag{43}$$

Substitution of (42) at $t - 1$ into (43) and the obvious manipulation yield

$$\bar{K}_t = \frac{h_2C(L)L}{1 - h_1L}v_t + \frac{h_2D(L)L}{1 - h_1L}\eta_t \tag{44}$$

on the assumption of an infinite past. And thus one may write the observations of firms in market i at time t as

$$z_t^i = \frac{v_t}{1 - \rho L} + \epsilon_t^i \tag{45}$$

$$P_t = \left[\frac{1}{1 - \rho L} - \frac{b_1f_0h_2C(L)L}{1 - h_1L}\right]v_t + \left[1 - \frac{b_1f_0h_2D(L)L}{1 - h_1L}\right]\eta_t. \tag{46}$$

As in Section VI, one might hope to make use of the observer equations (45)–(46) at t, $t - 1$, and so on with formulas for conditional means and variances of normal random variables to form posterior distributions at date t on the innovations $v_t, v_{t-1}, \ldots$, with means $E(v_t|\Omega_t^i)$, $E(v_{t-1}|\Omega_t^i), \ldots$, and so on. These means might in turn be expressed in terms of current and past economic innovations $v_t, v_{t-1}, \ldots, \eta_t, \eta_{t-1}, \ldots, \epsilon_t^i, \epsilon_{t-1}^i, \ldots$. Then these means might be used to form conditional mean forecasts of θ_t at time t, that is, $M_t^i = E(\theta_t|\Omega_t^i) = \Sigma_{j=0}^{\infty}\rho^jE(v_{t-j}|\Omega_t^i)$. One might then average over markets i and deliver an equation of the form (42). One might then hope the coefficients in $C(L)$ and $D(L)$ could be determined.

A potential problem with this approach is the appearance of an infinite number of innovations. To circumvent that problem, one might suppose an initial starting date, an initial prior distribution, and an appropriate (and possible time-dependent) specification of (42). But that approach is awkward at best, and there remains the problem of increasing dimensionality over time, as in Section VI.

An alternative approach, which keeps the dimensionality finite but allows the advantage of an infinite past, is to alter the information structure somewhat. Suppose that all economic innovations dated $t - j$ and backward are known at date t (see also Chari 1979). For example, suppose $j = 2$. Then,

$$\Omega_t^i = \{P_t^i, P_{t-1}^i, \ldots, K_t^i, K_{t-1}^i, \ldots, M_t^i, M_{t-1}^i, \ldots, P_t, P_{t-1}, \ldots, \eta_{t-2}, \eta_{t-3}, \ldots, v_{t-2}, v_{t-3}, \ldots, \epsilon_{t-2}^i, \epsilon_{t-3}^i, \ldots\}.$$

Then there are a *finite* number of observation variables in market i at date t containing information on the unknown innovations, namely,

$$z_t^{i*} = v_t + \rho v_{t-1} + \epsilon_t^i, \tag{47}$$

$$z_{t-1}^{i**} = v_{t-1} + \epsilon_{t-1}^i, \tag{48}$$

$$P_t^* = v_t + \rho v_{t-1} - b_1 f_0 h_2 D_0 \eta_{t-1} - b_1 f_0 h_2 C_0 v_{t-1} + n_t, \tag{49}$$

$$P_{t-1}^{**} = v_{t-1} + n_{t-1}, \tag{50}$$

where the * and ** variables on the left-hand side of (47)–(50) are defined implicitly and include all past *known* innovations. (Note here from [43] that $\bar{K}_{t-1}$ and $\bar{K}_{t-2}$ are known at time t.) Thus one can make use of the standard formulas for conditional means and variances.[14] In (17)–(18), for example, let

$$\mathbf{x}' = [v_t \; v_{t-1}] \qquad \bar{\mathbf{x}}' = [0 \;\; 0]$$

$$\mathbf{y}' = [z_t^{i*} \; z_{t-1}^{i**} \; P_t^* \; P_{t-1}^{**}] \quad \mathbf{w}' = [\epsilon_t^i \; \epsilon_{t-1}^i \; \eta_t - b_1 f_0 h_2 D_0 \eta_{t-1} \; \eta_{t-1}],$$

and so on. Then (17) delivers $E(v_t | \Omega_t^i)$ and $E(v_{t-1} | \Omega_t^i)$ as linear functions of v_t, v_{t-1}, ϵ_t^i, ϵ_{t-1}^i, η_t, and η_{t-1}. That is,

$$M_t^i = E(v_t | \Omega_t^i) + \rho E(v_{t-1} | \Omega_t^i) + \sum_{j=2}^{\infty} \rho^j v_{t-j}, \tag{51}$$

which takes on the form

$$\begin{aligned} M_t^i = {} & \psi_0 v_t + \psi_1 v_{t-1} + \phi_0 \eta_t + \phi_1 \eta_{t-1} \\ & + \Lambda_1 \epsilon_t^i + \Lambda_1 \epsilon_{t-1}^i + \sum_{j=2}^{\infty} \rho^j v_{t-j}. \end{aligned} \tag{52}$$

Averaging over markets i then yields

$$\bar{M}_t = \psi_0 v_t + \psi_1 v_{t-1} + \phi_0 \eta_t + \phi_1 \eta_{t-1} + \sum_{j=2}^{\infty} \rho^j v_{t-j}. \tag{53}$$

This is of the same form as (42), so one can hope to match coefficients. Namely,

$$\psi_0 = C_0, \; \phi_0 = D_0. \tag{54}$$

We shall return to (54) momentarily.

Of course, the procedure just described can be applied to any finite, full-information lag specification, that is, with economic innovations dated $t - j$ and backward known at date t, j finite. But it is also apparent that the dimensionality problem begins to reemerge as j increases; a matrix of dimension $2j$ must be inverted.

[14] This procedure was suggested to me by Ian Bain and Edward C. Prescott as an alternative to the procedure described below. The *entire* paper is now written in a way, I hope, which makes the present procedure obvious and straightforward.

This observation alone might motivate the search for an alternative, more recursive procedure. And as it turns out, there exists an alternative procedure which is of intrinsic interest given the topic of this paper, since it makes clear and exploits the distinction between shocks to driving stochastic processes, the economic shocks, and shocks which are innovations relative to decision makers' information sets.

To begin to describe the alternative procedure, suppose one had under consideration a system of the form

$$\mathbf{y}_t = M(L)\boldsymbol{\xi}_t, \tag{55}$$

where $M(L)$ is a polynomial which is one-sided in nonnegative powers of L. Then the observation variables $\mathbf{y}_t$ are just moving averages of past, observable shocks $\boldsymbol{\xi}_t$ (the $\boldsymbol{\xi}_t$ are assumed to be serially uncorrelated and have an identity variance-covariance matrix). Then one makes use of the powerful Weiner-Kolmogorov prediction formulas, namely,

$$E_t(\mathbf{y}_{t+1}) = \left|\frac{M(L)}{L}\right|_+ \boldsymbol{\xi}_t \tag{56}$$

and

$$E_t(\mathbf{y}_{t+2}) = \left|\frac{M(L)}{L^2}\right|_+ \boldsymbol{\xi}_t \tag{57}$$

and so on, where + means ignore negative powers of L. Clearly, these predictions are easily calculated.

Returning to equations (47) and (49) and supposing full information at $t - 2$, for example, one may write

$$\begin{bmatrix} v_{t-2} \\ z_t^{i*} \\ P_t^* \end{bmatrix} = \begin{bmatrix} L^2 & 0 & 0 \\ (1 + \rho L) & 1 & 0 \\ (1 + \alpha L) & 0 & (1 - \delta L) \end{bmatrix} \begin{bmatrix} v_t \\ \epsilon_t^i \\ \eta_t \end{bmatrix}, \tag{58}$$

where $\alpha = \rho - b_1 f_0 h_2 C_0$ and $\delta = b_1 f_0 h_2 D_0$. System (58) is similar in appearance to (55), that is, it can be written as $\mathbf{y}_t = M(L)\,\boldsymbol{\xi}_t$, but there is an obvious problem: the shocks on the right-hand side of (58) are *not* observable or deducible (note that in [58], $M[L]$ does not have an inverse which is one-sided in nonnegative powers of L). Equivalently, the moving-average system (58) is not fundamental relative to the information sets of firms in market i. To make use of the Weiner-Kolmogorov prediction formulas, one must express the variables on the right-hand side of (58) in terms of observable or deducible shocks.[15] That is, one must find a representation which is fundamental relative to the information sets of firms.

To motivate the search for this fundamental moving-average representation, consider the covariance generating function for the $\mathbf{y}_t$ in (58), namely, $g(z) = \Sigma_{h=-\infty}^{+\infty}\gamma(h)z^h$, where $\gamma(h) = E[\mathbf{y}_t\mathbf{y}'_{t+h}]$, a function which thus describes the entire variance-covariance structure of the observables $\mathbf{y}_t$. On the assumption of unit variances, $\sigma_\epsilon^2 = \sigma_\eta^2 = \sigma_v^2 = 1$, which is not essential, and following Sargent (1979), the function $g(z)$ may be written

$$g(z) = M(z)M(z^{-1})' \tag{59}$$

or, alternatively,

$$g(z) = M(z)WB(z)\hat{W}B(z)B(z^{-1})'\hat{W}'B(z^{-1})'W'M(z^{-1})', \tag{60}$$

where

$$B(z)B(z^{-1})' = I, \quad WW' = I, \quad \text{and} \quad \hat{W}\hat{W}' = I. \tag{61}$$

Here $B(z)$ is the Blaschke factor discussed in Hansen and Sargent (1981). Thus the variance-covariance structure of the data does not pin down a unique moving-average representation. The idea, then, from (60) is to convert the representation (58) to a fundamental representation

$$\mathbf{y}_t = M^{**}(L)\boldsymbol{\xi}_t^{**}, \tag{62}$$

that is, with known innovation $\boldsymbol{\xi}_t^{**}$ where

$$M^{**}(L) = M(L)WB(L)\hat{W}B(L), \tag{63}$$

$$\boldsymbol{\xi}_t^{**} = B(L^{-1})'\hat{W}'B(L^{-1})'W'\boldsymbol{\xi}_t. \tag{64}$$

The reader is referred to Appendix A for details of the calculations.

[15] I would like to thank Lars Hansen for suggesting systems like (58) and the use of Weiner-Kolmogorov prediction formulas, and for teaching me about Blaschke factors (see below). Since η_{t-2} is in the information set of firms at time t, it may appear that system (58) should be augmented to include η_{t-2} in the vector on the left-hand side. But the procedure described in the text below can be applied to the augmented system with the result that the predictions of v_t and v_{t-1} remain unaltered. It might also seem that parameter δ must be less than unity to justify the procedure described in the text, but the augmented system makes clear that this need not be the case. (I am indebted to Lars Hansen for this insight.)

Again, the Weiner-Kolmogorov prediction formulas are

$$E\left(\begin{bmatrix} v_{t-1} \\ z^{i*}_{t+1} \\ P^*_{t+1} \end{bmatrix} \middle| \Omega^i_t\right) = \left[\frac{M^{**}(L)}{L}\right]_+ \xi^{**}_t \tag{65}$$

$$E\left(\begin{bmatrix} v_t \\ z^{i*}_{t+2} \\ P^*_{t+2} \end{bmatrix} \middle| \Omega^i_t\right) = \left[\frac{M^{**}(L)}{L^2}\right]_+ \xi^{**}_t \tag{66}$$

for the one-step-ahead and two-step-ahead forecasts. So, to forecast v_t and v_{t-1}, one needs only the first row of $M^{**}(z)$:

$$\frac{1}{\sqrt{3}}\hat{W}_{11} + \frac{2z}{\sqrt{6}}\hat{W}_{21} \quad \frac{z}{\sqrt{3}}\hat{W}_{12} + \frac{2z^2}{\sqrt{6}}\hat{W}_{22} \quad \frac{2z^2}{\sqrt{6}}\hat{W}_{23}.$$

Thus,

$$\begin{aligned} E(v_{t-1}|\Omega^i_t) = {} & \frac{2\hat{W}_{21}}{\sqrt{6}}\xi^{**}_{1t} + \frac{\hat{W}_{12}}{\sqrt{3}}\xi^{**}_{2t} + \frac{2\hat{W}_{22}}{\sqrt{6}}\xi^{**}_{2,t-1} \\ & + \frac{2\hat{W}_{23}}{\sqrt{6}}\xi^{**}_{3,t-1} \end{aligned} \tag{67}$$

and

$$E(v_t|\Omega^i_t) = \frac{2\hat{W}_{22}}{\sqrt{6}}\xi^{**}_{2t} + \frac{2\hat{W}_{23}}{\sqrt{6}}\xi^{**}_{3t}. \tag{68}$$

These prediction formulas are linear in the ξ^{**} with coefficients linear in the $\hat{W}_{ij}$'s.

To search for the undetermined coefficients C_0 and D_0, (67) and (68) must be cast in terms of the economic innovations, v_t, η_t, and ϵ^i_t. The relationship is expressed in (64). For example,

$$\begin{aligned} \xi^{**}_{1t} = {} & \left(\frac{2}{\sqrt{6}}\hat{W}_{21}L + \frac{1}{\sqrt{3}}\hat{W}_{11}L^2\right)v_t \\ & + \left(\frac{1}{\sqrt{2}}\hat{W}_{31}L + \frac{1}{\sqrt{6}}\hat{W}_{21}L - \frac{1}{\sqrt{3}}\hat{W}_{11}L^2\right)\epsilon^i_t \\ & + \left(-\frac{1}{\sqrt{2}}\hat{W}_{31}L + \frac{1}{\sqrt{6}}\hat{W}_{21}L - \frac{1}{\sqrt{3}}\hat{W}_{11}L^2\right)\eta_t. \end{aligned} \tag{69}$$

One can thus see how the information set innovations are linear functions of current and past economic shocks. Substituting expressions like (69) into

(67) and (68) yields the forecasts $E(v_{t-1}|\Omega_t^i)$ and $E(v_t|\Omega_t^i)$ as linear functions of v_t, v_{t-1}, ϵ_t^i, ϵ_{t-1}^i, η_t, and η_{t-1}, with coefficients which are *quadratic* in the $\hat{W}_{ij}$'s. This again solves the forecasting problem.[16]

We turn now to the problem of finding the undetermined C_0 and D_0 from equation (54). We might note in particular that C_0 and D_0 enter on the left-hand side of (54) in the parameters ψ_0 and ϕ_0. That is, parameters C_0 and D_0 were taken as given by the firms in solving their inference problems. But these inferences in turn determine the law of motion in (42) and thus affect the parameters ψ_0 and ϕ_0. This is the confounding problem. One may note, moreover, the fundamentally nonlinear nature of this problem. The parameters C_0 and D_0 determine (in part) the coefficients $\hat{W}_{ij}$, and ψ_0 and ϕ_0 are quadratic functions of the latter. Still, equations (54) turn out to be not all that awesome. Existence and uniqueness of solutions are established in Appendix B, making use of the Blaschke procedure.

The mapping described in Appendix B also suggests a scheme for computing solutions to (54): the method of successive approximations. And though there is no formal basis for supposing that solutions can be computed by this method—that is, the mapping is not shown to be a contraction—in practice this method generates the solution quickly for a broad range of initial parameter values (e.g., accurate to eight decimal places after 10 iterations). The economics behind this method is somewhat reassuring. Suppose firms were to take as given some (arbitrary) moving-average representation for economy-wide average forecasts of the common demand shock. Then the inference problem in each market becomes well defined and produces a best forecast in each market of the common shock. Averaging over markets, one arrives at a new moving-average representation. Repeating the process, there is convergence to a situation in which average forecasts in the population are consistent with best forecasts based on that average.[17] (This is reminiscent of a convergence result in Lucas [1978].)

In concluding this section, it should be noted from (52) that the response of M_{t+j}^i to innovations η_t, ϵ_t^i, and v_t is determined entirely by the coefficient ρ^j on v_t after j periods for j sufficiently large. Thus, M_t^i

[16] As Subramanian (1982) has established, the Blaschke procedure described above is numerically tractable for *any* finite, full-information lag system, with innovations known at $t-j$. (Lags up to $j = 50$ have been computed.) The dimension of the matrices remains at $2 + 1 = 3$ and does not increase with j. Of course, the Blaschke algorithm requires more and more iterations; that is, more and more matrices W, $\hat{W}$, and so on must be computed. But there is a sense in which the procedure is recursive, as is apparent from (63) and (64). The reader is urged to consult Subramanian for a more detailed account.

[17] Subramanian (1982) has established that this procedure converges, and converges to apparently unique solutions, for arbitrarily large information lag j systems.

is stable, and so K_t^i and P_t^i are stable and therefore bounded in mean. Thus the existence of an equilibrium has been established.

IX. Time-Series Dynamics for the Symmetric Structure

Now suppose an econometrician were to observe economy-wide average capital stock and price series for the economy of Section VIII. To characterize the restrictions which the theory puts on these series, we shall examine responses to measurement errors and to innovations in economic shocks at date 1 for various possible parameter values.

To begin, note from (42) and the law of motion for economy-wide average capital stock that

$$(1 - h_1 L)\overline{K}_{t+1} = h_2[C(L)v_t + D(L)\eta_t] \tag{70}$$

with economy-wide average price determined by

$$\overline{P}_t = -b_1 f_0 \overline{K}_t + \sum_{j=0}^{\infty} \rho^j v_{t-j}. \tag{71}$$

The response of the system (70)–(71) to shocks v_t in the common demand component is very much as expected with the extent of cumulation and persistence varying directly with the magnitude of the autocorrelation parameter ρ and the degree of costly capital adjustment f_2/f_0. More interesting experiments concern the response of the system to measurement errors η_t. Though there are no oscillations in forecast errors (or output), a somewhat surprising feature is the magnitude and extent of the response to such bogus shocks, as displayed in figure 4. For the case of relatively high adjustment costs, the capital stock cumulates and takes considerable time to slide back toward its steady-state value. And this is so even if firms are given full information after two periods ($j = 2$). In fact the response patterns are not much different for the case of the longer information lags; the case $j = 20$ is computed here using Subramanian's (1982) program and displayed in figure 4. For the case of relatively low adjustment costs, the response patterns again are not much different. There is more persistence when the lag to full information increases, the $j = 20$ case (see fig. 5). But the initial response to the measurement error in the $j = 20$ case is now slightly *less* than in the $j = 2$ case, though firms have relatively less information overall. Again the steady-state variance-covariance matrix of beliefs is sensitive to the adjustment cost parameters f_0 and f_2.

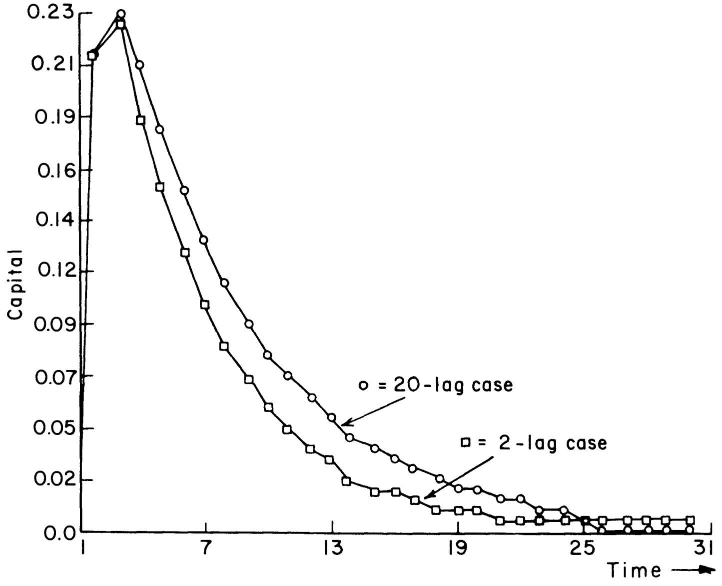

FIG. 4.—Symmetric structure response to measurement error (high adjustment costs).

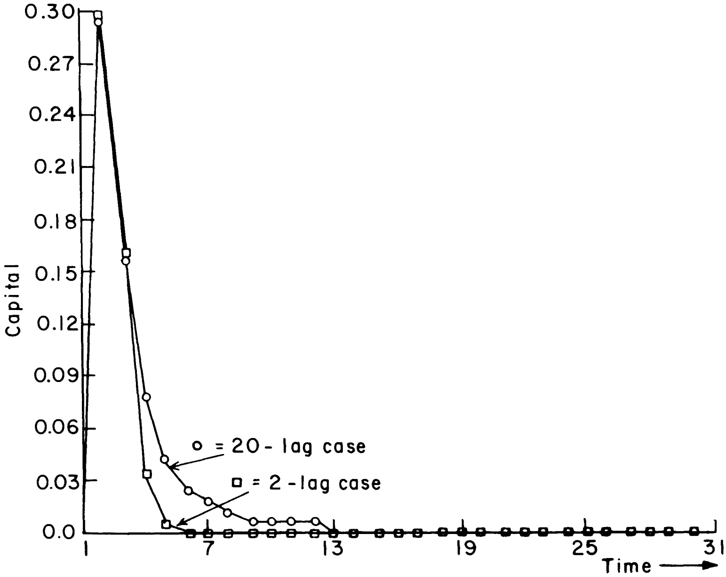

FIG. 5.—Symmetric structure response to measurement error (low adjustment costs)

X. Fitting the Theory to Data (in Principle)

We now ask if an econometrician could determine in principle whether the theories with learning and somewhat sophisticated expectation formation being proposed in this paper are consistent with actual data. One way one might hope to do this, consistent with the time-series dynamics already presented and with the work of Sims (1980*a*, 1980*b*), is to look in the data for patterns of response to economic shocks or innovations and to try to match these with response patterns predicted by the theory. We shall comment on this procedure at the end of this section. But for now we merely note that these response patterns are a convenient way, but not the only way, of describing the variance-covariance structure of the variables observed by the econometrician. Formal estimation or "fitting" procedures try to match directly the variance-covariance structure of observables with the variance-covariance structure of corresponding variables of the theory.

To be more precise, we shall adopt here, with suitable modifications, a technique of Sargent (1978, 1981) and Hansen and Sargent (1980*a*, 1980*b*), a technique which lets the theory deliver the entire econometric model, including the error term. The econometric model is not exact, because the econometrician treats both forecasts and underlying economic shocks as unobservable, latent variables. And here, as in the Hansen-Sargent papers, the theory implies certain cross-equation restrictions on the coefficients of the econometric model and certain lag (zero) restrictions on its error terms as well. One then uses these restrictions to estimate and identify the underlying parameters of the theory, the parameters of preferences, technology, and driving stochastic processes. There follows a caveat on the "exact" nature of such error term models. This will provoke a discussion of the possibilities and limitations of more standard econometric procedures, on the one hand, and to a suggestion of Kydland and Prescott (1982), on the other. As noted, this section concludes with some comments on Sims's suggestion for summarizing data and a caveat on that method.

To begin, suppose an econometrician had available time series on the capital stock and price in the first market in the hierarchical structure of Section VI. (We shall focus on the first market alone in order to concentrate on method.) Suppose, however, that the underlying shocks, v_t and ϵ_t^1, are unobserved by the econometrician. Thus, market beliefs M_t^1 and these shocks play the role of unobserved variables and, following Hansen and Sargent, deliver the error term in the vector time series, a mixed autoregressive moving-average process:

$$\begin{bmatrix} (1-\rho L)(1-h_1L)(1-\alpha_0L) & 0 \\ 0 & (1-\rho L)(1-h_1L)(1-\alpha_0L) \end{bmatrix} \begin{bmatrix} K^1_{t+1} \\ P^1_t \end{bmatrix}$$

$$= \begin{bmatrix} h_2\alpha_1\sigma_v \\ [(1-h_1L)(1-\alpha_0L) - b_1f_0h_2\alpha_1L]\sigma_v \end{bmatrix} \tag{72}$$

$$\begin{bmatrix} [h_2\alpha_1(1-\rho L)]\sigma_\epsilon \\ [(1-h_1L)(1-\rho L)(1-\alpha_0L) - b_1f_0h_2\alpha_1L(1-\rho L)]\sigma_\epsilon \end{bmatrix} \begin{bmatrix} v^*_t \\ \epsilon^{*1}_t \end{bmatrix}.$$

System (72) is the econometric model. One notes immediately the various zero restrictions on the coefficients of the lag structure and the cross-equations restrictions on coefficients in terms of the parameters of the model: $f_0, f_1, f_2, \rho, \beta, \sigma^2_\epsilon, \sigma^2_v, b_1$. (Note that these parameters determine h_1, h_2, α_0, and α_1.)

Now suppose an econometrician were given data over T periods, $\mathbf{K}^1(T)' = [K^1_1K^1_2 \ldots K^1_T]$ and $\mathbf{P}^1(T)' = [P^1_0P^1_1 \ldots P^1_{T-1}]$. On the assumption (consistent with the theory) that these data are generated from a normal distribution, the likelihood function is

$$l(T) = -\tfrac{1}{2}(2T)\log 2\pi \quad \tfrac{1}{2}\log\det[\Gamma(T)]$$
$$- \tfrac{1}{2}[\mathbf{K}^1(T)'\mathbf{P}^1(T)']\Gamma(T)^{-1}\begin{bmatrix}\mathbf{K}^1(T)\\ \mathbf{P}^1(T)\end{bmatrix}, \tag{73}$$

where

$$\Gamma(T) = E\begin{bmatrix}\mathbf{K}^1(T)\\ \mathbf{P}^1(T)\end{bmatrix}[\mathbf{K}^1(T)'\mathbf{P}^1(T)']$$

is the variance-covariance matrix and is determined by the theory and the parameters of the model. Maximum-likelihood estimates are derived by maximizing $l(T)$ with respect to various feasible values for these parameters.

It is instructive to see how $\Gamma(T)$ is determined by theory and the parameters of the model. For convenience, system (72) may be rewritten as

$$H_1(L)\begin{bmatrix}K^1_{t+1}\\ P^1_t\end{bmatrix} = H_2(L)\begin{bmatrix}v^*_t\\ \epsilon^{*1}_t\end{bmatrix}. \tag{74}$$

Now note that $\Gamma(T)$ is completely determined under this stationary structure by the covariance matrices

$$\gamma(h) = E\begin{bmatrix}K^1_{t+1}\\ P^1_t\end{bmatrix}[K^1_{t+1+h}P^1_{t+h}],\ h = 0, \pm 1, \pm 2, \ldots,$$

and the $\gamma(h)$ are completely determined by the covariance generating function $g(z) = \Sigma_{h=-\infty}^{+\infty}\gamma(h)z^h$, where $g(z) = H_1(z)^{-1}H_2(z) \times H_2(z^{-1})'H_1(z^{-1})^{-1\prime}$. As noted, the matrices of coefficients in $H_1(z)$, $H_2(z)$ have already been determined by theory and can be expressed in terms of the underlying parameters of the model.

It is of interest to ask whether these underlying parameters can be identified, that is, whether there is any indeterminacy in the estimates for large (infinite) samples. Suppose one were armed with the true variance-covariance structure of the bivariate process K_{t+1}^1, P_t^1 so that the $\gamma(h)$ and hence $g(z)$ were actually known. We ask first whether the coefficients in the matrices of $H_1(z)$, $H_2(z)$ would be known and then whether the underlying parameters can be determined from these coefficients.

The first question can be posed another way. Given some $H_1(z)$, $H_2(z)$ pair, is there another $H_1(z)$, $H_2(z)$ pair consistent with the same $g(z)$? Writing $g(z) = H_1(z)^{-1}H_2(z)B(z)B(z^{-1})'H_2(z^{-1})'H_1(z^{-1})^{-1\prime}$, where $B(z)$ is orthogonal, that is, $B(z)B(z^{-1})' = I$, it seems the answer may be no; apparently $H_2(z)$ is only unique up to orthogonal matrices $B(z)$. Here again, $B(z)$ is the Blaschke factor discussed in Section VIII. But now invoke the restrictions implied by theory on the matrices of $H_2(z)$, in particular the zero restrictions. It can be shown that $H_1(z)$ and $H_2(z)$ are indeed identified.[18]

The second part of the identification question is whether one can uncover the underlying parameters from the coefficients in $H_1(z)$ and $H_2(z)$. It is claimed that this can be done *if* the econometrician has prior knowledge of the discount rate, β, making use of the cross-equation restrictions.[19] Thus, the vector time series from the first market will effectively reveal all parameter values.[20]

[18] The argument is due to Lars Hansen, and is available from me on request.

[19] From the second-row, first-column and second-column elements of $H_2(L)$, it is evident that σ_η and σ_ϵ are determined. From the first-row, second-column element of $H_2(L)$, $h_2\alpha_1\sigma_\epsilon$ and $\rho h_2\alpha_1\sigma_\epsilon$ are determined, so ρ is known. Then $\bar{\Sigma}$ can be computed from (24) and (25), and thus α_0 and α_1 are determined. Thus, from $h_2\alpha_1$, h_2 is determined. From $H_1(L)$, the three parameters h_1, α_0, and ρ are known up to labeling. But with ρ and α_0 known, h_1 is determined. So from the second-row, first-column element, $b_1 f_0$ is determined. Now recall from (13) that $h_1 = \gamma_1$, $h_2 = (f_0\gamma_1\beta\rho)/[f_2(1 - \gamma_1\beta\rho)]$. With γ_1, β, and ρ known, f_0/f_2 is determined. Then, since $\gamma_1\gamma_2 = 1/\beta$, $\gamma_1 + \gamma_2 = (f_1/f_2) + [(1 + \beta)/\beta] + (b_1 f_0 f_0/f_2)$, f_1/f_2 is determined. Thus the parameters of technology are determined up to some normalization, say $f_2 = 1$, as expected since multiplying the objective function through by a constant cannot change the implications of the theory. With $f_2 = 1$, f_0, f_1, and b_1 are determined.

[20] Perhaps a more interesting question to ask is whether identification would be possible if parameter values of technology and demand were allowed to vary across markets. Data from the first market would allow one to determine the coefficients in the matrices C, A, M, and N, and thus $\bar{\Sigma}$ would be determined. These might then be used to identify the parameters in the second market from the analogue of system (72). The point is that the recursive structure can be exploited.

To proceed with more formal econometrics for the symmetric information structure, system (70)–(71) can be expressed as an ARMA process, and thus maximum-

The technique described above for estimation and identification assumes that the data really are generated by the model of the theory. Note, for example, that the error term in a vector autoregression of price and capital stock in the first market is a linear combination of current and past unobserved economic shocks which constitute part of the model itself; in the model these shocks have effects on observables either directly, as on contemporary prices, or indirectly through their effect on decision makers' forecasts. Thus, movement in the error term plays a key role in identifying the underlying parameters of the model. And the variance of the error term is not necessarily small relative to the variance of observables. Consider the model of capital stock movements in the econometric model (72). There is no reason a priori to believe that uncertainty is not a fundamental part of the problem confronting decision makers (as it would be if both σ_v^2 and σ_ϵ^2 were small) or that current observations are unimportant in learning (as they would be if α_1 were small).

Unfortunately, though, this wedding of theory and econometrics is not without difficulties. The above-mentioned technique leaves no room for errors which result from theory as an abstraction or approximation. That is, all the rich detail of the data is forced on the free parameters of the model. Viewed as an observable-unobservable index model—as described, for example, in Sargent and Sims (1977)—the model is exact; there are no own error terms (with, we hope, small variances).

This leads us to a discussion of perhaps a more standard econometric procedure, deriving an exact model from theory and tacking on error terms as dictated by discretion and by the data. We shall now examine in what sense this is or is not possible if agents of the model themselves have incomplete information on exogenous driving processes and thus are engaged in learning.

Suppose, first, that the econometrician has *direct* observations, along with the agents of the model, on the stochastic processes being forecast. Then the econometrician could potentially estimate the parameters of those stochastic processes, and, with guesses about agents' initial priors, the Kalman filtering algorithm could be used to compute agents' forecasts period by period. With these constructed forecast series and the other observables, part of the theoretical model

---

likelihood techniques of Hansen and Sargent may again be used for estimation. Under such techniques the parameters C_j, D_j must be computed, say, by successive approximations. But, these computations are no more burdensome than those used in solving the recursive eq. (24) of the Kalman filter. And, needless to say, numerical methods must be used in any event under any estimation procedure. One caveat is in order, though: here identification of the parameters in technology and demand *may* be a problem because there is no recursive way to attack the cross-equation restrictions, as in the two-market hierarchical structure. This is left as an open question.

may be exact. Thus, one might consider tacking on error terms for purposes of estimation. This is virtually the procedure adopted by Crawford (1979); it could be applied here to the model of the first market if the econometrician were to get direct observations on z_t^1.

Without such direct observations, though, this recursive procedure cannot be applied. With data on the capital stock and price only, the pricing equation in the first market alone will not yield the parameters ρ, σ_v^2, and σ_ϵ^2; that is, the parameter b must be estimated jointly. But to do this consistently, one needs to make explicit use of the evolution of the capital stock series. Thus, even if the econometrician and the agents of the model see the same time series, the agents, knowing the structure, can deduce more period by period. Thus, from the point of view of the econometrician, the model delivered by theory is not exact. This leads us back toward the first estimation strategy of Hansen and Sargent discussed above.

In some contexts it may be possible to loosen the Hansen-Sargent procedure directly by tacking error terms onto the econometric model derived from theory. This is especially true if the model from theory displays stochastic singularities, as the one here apparently does. Then one might hope to retain identification. In general, though, identification becomes an open question.

Another approach which takes theory seriously and which avoids tacking on error terms is that of Kydland and Prescott (1982). Rather than asking theory to explain the entire variance-covariance structure of the data, as in maximum-likelihood techniques, those authors concentrate on key moments or "on certain statistics for which the noise introduced by approximations and measurement errors is likely to be small relative to the statistic." Under this approach, one tries to match specified sample moments from runs of the model under various possible parameter specifications with the specified moments from actual data. Thus, one again makes direct use of the Kalman filtering algorithm in computing agents' forecasts.

If one is unable or unwilling to focus on a subset of moments, then one is again faced with the problem of fitting the theory to the entire variance-covariance structure of the data. Sims (1980*a*, 1980*b*) has suggested that a nice way to summarize this structure is to estimate the coefficients in relatively unrestricted vector autoregressions and examine the response of the system to innovations in the error terms of these regressions. This brings us back to the response to economic shocks in Section VII as one way to summarize the implications of the theory. A caveat is in order here, however: the error term in the theoretical model (72), for example, has no natural economic interpretation, as it is a linear combination of current and *past* economic shocks. That is, the theoretical model (72) can be inconsistent with the assumptions Sims makes on error terms in his vector autoregressions (in other models, of course) and with his interpretation of error terms as surprises or forecast errors. Putting this another way, the error term in a vector autoregression can be sensitive to the assumed information structure for decision makers in the model itself. The error term in the full-information model (40) has a ready interpretation as a permanent shock. The error term in the learning model (39) is some mongrel. The full implication of this for (highly desirable) exploratory data analysis is left as a subject for future research.

Appendix A

An Algorithm for Uncovering the Fundamental Moving-Average Representation

Note first that the matrix $M(z)$ is lower triangular and hence has a determinant of $z^2(1 - \delta z)$ with two zeros at $z = 0$. Thus $M(z)$ has two zeros *inside* the unit circle, and this is precisely why $M(z)$ does not have a one-sided inverse in nonnegative powers of L. We shall search for matrices W, $\hat{W}$, and $B(z)$ consistent with the orthogonality conditions (61) which "flip" these zeros outside the unit circle.

The algorithm is as follows. First, evaluate the matrix $M(z)$ at $z = 0$. Next, postmultiply by some matrix W in such a way as to put zeros in the first column of the product matrix. In this case

$$W = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} & 0 \\ -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \end{bmatrix}.$$

Next, get rid of the zeros in the first column by use of the Blaschke matrix,

$$B(z) = \begin{bmatrix} z^{-1} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

That is, let $M^*(z) = M(z)WB(z)$, so that

$$M^*(z) = \begin{bmatrix} \frac{z}{\sqrt{3}} & \frac{2z^2}{\sqrt{6}} & 0 \\ \left(\frac{1+\rho z}{\sqrt{3}} - \frac{1}{\sqrt{3}}\right)z^{-1} & \frac{(1+\rho z)2}{\sqrt{6}} + \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \left(\frac{1+\alpha z}{\sqrt{3}} - \frac{1-\delta z}{\sqrt{3}}\right)z^{-1} & \frac{(1+\alpha z)2}{\sqrt{6}} + \frac{1-\delta z}{\sqrt{6}} & \frac{-(1-\delta z)}{\sqrt{2}} \end{bmatrix}.$$

We have, in effect, flipped one of the zeros outside the unit circle. But note that the determinant of $M^*(z)$ still has a zero at $z = 0$. So, apply the same procedure, again starting from $M^*(z)$. This time the matrix analogue of W is $\hat{W}$ defined by terms like

$$\hat{W}_{11} = \left(\frac{18}{18 + k_1^2 + 3k_2^2}\right)^{1/2},$$

and so on, where $k_1 = \rho + \alpha + \delta$ and $k_2 = \rho - \alpha - \delta$, so that the $\hat{W}_{ij}$ are nonlinear functions of α and δ. The relevant Blaschke matrix is the same $B(z)$ as before. This procedure yields $M^{**}(z) = M^*(z)\hat{W}B(z)$. Thus, the fundamental moving-average representation has been achieved with moving-average coefficients $M^{**}(z)$ as defined in (63) and with innovations ξ_t^{**} as defined in (64).

Appendix B

Existence and Uniqueness of Solutions for the Undetermined Coefficients

Consider first the following mapping. Start with arbitrary parameters C_0 and D_0. These determine numbers α and δ, and these in turn determine numbers k_1 and k_2. The numbers k_1 and k_2 determine the $\hat{W}_{ij}$'s, but it is important to note that the matrix $\hat{W}$ must be orthogonal. Thus, each $\hat{W}_{ij}$ is bounded in absolute value by unity, for all initial parameters C_0 and D_0. And, from system (54), the computed C_0, D_0 values on the left-hand side of (54) must satisfy

$$C_0 = \frac{2\hat{W}_{22}^2}{3} + \frac{2\hat{W}_{23}^2}{3} + \frac{2\rho\hat{W}_{12}\hat{W}_{22}}{\sqrt{18}},$$

$$D_0 = \frac{2\hat{W}_{22}}{\sqrt{6}}\left(\frac{1}{\sqrt{6}}\hat{W}_{22} - \frac{1}{\sqrt{2}}\hat{W}_{32}\right) + \frac{2\hat{W}_{23}}{\sqrt{6}}\left(\frac{1}{\sqrt{6}}\hat{W}_{23} - \frac{1}{\sqrt{2}}\hat{W}_{33}\right) \quad \text{(B1)}$$

$$+ \rho\frac{\hat{W}_{12}}{\sqrt{3}}\left(\frac{1}{\sqrt{6}}\hat{W}_{22} - \frac{1}{\sqrt{2}}\hat{W}_{32}\right).$$

Thus, computed C_0 and D_0 take on values in a nonempty, compact, convex space. Moreover, the mapping defined above is well defined, assuming $\rho \neq 0$, and continuous on this space. So, by the Brouwer fixed-point theorem, there exists a solution to equations (54).

It remains to formally establish the uniqueness of solutions to (54). Substitution into (B1) and cancellation of like terms yield

$$C_0 = \frac{12 + 4\rho^2 - 2\rho b_1 f_0 h_2(C_0 - D_0) + 2(b_1 f_0 h_2)^2(C_0 - D_0)^2}{4\rho^2 - 4\rho b_1 f_0 h_2(C_0 - D_0) + 4(b_1 f_0 h_2)^2(C_0 - D_0)^2 + 18}, \quad \text{(B2)}$$

$$D_0 = \frac{6 + 2\rho^2 - 2\rho b_1 f_0 h_2(C_0 - D_0) + 2(b_1 f_0 h_2)^2(C_0 - D_0)}{4\rho^2 - 4\rho b_1 f_0 h_2(C_0 - D_0) + 4(b_1 f_0 h_2)^2(C_0 - D_0)^2 + 18}, \quad \text{(B3)}$$

and

$$C_0 - D_0 = \frac{6 + 2\rho^2}{4\rho^2 - 4\rho b_1 f_0 h_2(C_0 - D_0) + 4(b_1 f_0 h_2)^2(C_0 - D_0)^2 + 18}. \quad \text{(B4)}$$

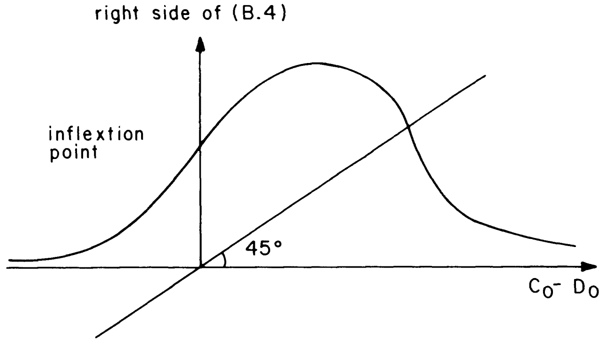

FIG. 6

Plotting the right-hand side of equation (B4) in figure 6, we see clearly that there is a unique solution to (B4) in (C_0-D_0) space. Then C_0 and D_0 are determined from (B2) and (B3), respectively. (Unfortunately, attempts to sign the slope of the function in fig. 6 have been unsuccessful, foiling attempts to formally establish why computing solutions by the method of successive approximations works.)

References

Allen, Beth E. "Generic Existence of Completely Revealing Equilibria for Economies with Uncertainty When Prices Convey Information." *Econometrica* 49 (September 1981): 1173–99.

Arrow, Kenneth J., and Green, Jerry R. "Notes on Expectations Equilibria in Bayesian Settings." Working Paper no. 35, Stanford Univ., Inst. Math. Studies Soc. Sci., August 1973.

Barro, Robert J. "A Capital Market in an Equilibrium Business Cycle Model." *Econometrica* 48 (September 1980): 1393–1417.

Bertsekas, Dimitri P. *Dynamic Programming and Stochastic Control.* New York: Academic Press, 1976.

Blume, Lawrence E., and Easley, David. "Learning to Be Rational." *J. Econ. Theory* 26 (April 1982): 340–51.

Bray, Margaret. "Learning, Estimation, and the Stability of Rational Expectations." *J. Econ. Theory* 26 (April 1982): 318–39.

Bray, Margaret, and Kreps, David M. "Rational Learning and Rational Expectations." Manuscript. Stanford, Calif.: Stanford Univ., Graduate School Bus., June 1981.

Brunner, Karl; Cukierman, Alex; and Meltzer, Allan H. "Stagflation, Persistent Unemployment and the Permanence of Economic Shocks." *J. Monetary Econ.* 6 (October 1980): 467–92.

Chari, V. V. "Linear Stochastic Models with Decentralized Information." Manuscript. Pittsburgh: Carnegie-Mellon Univ., Graduate School Indus. Admin., 1979.

Crawford, R. G. "Expectations and Labor Market Adjustments." *J. Econometrics* 11 (October/December 1979): 207–32.

Cyert, Richard M., and DeGroot, Morris H. "Rational Expectations and Bayesian Analysis." *J.P.E.* 82 (May/June 1974): 521–36.

DeCanio, Stephen J. "Rational Expectations and Learning from Experience." *Q.J.E.* 93 (February 1979): 47–57.

Eisner, Robert, and Strotz, Robert H. "Determinants of Business Investment." In *Impacts of Monetary Policy*, by the Commission on Money and Credit. Englewood Cliffs, N.J.: Prentice-Hall, 1963.

Friedman, Benjamin M. "Optimal Expectations and the Extreme Information Assumptions of 'Rational Expectations' Macromodels." *J. Monetary Econ.* 5 (January 1979): 23–41.

Frisch, Ragnar. "Propagation Problems and Impulse Problems in Dynamic Economics." In *Economic Essays in Honor of Gustav Cassel.* London: Allen & Unwin, 1933. Reprinted in *Readings in Business Cycles*, edited by Robert A. Gordon and Lawrence R. Klein. Homewood, Ill.: Irwin (for American Econ. Assoc.), 1965.

Futia, Carl A. "Stochastic Business Cycles." Manuscript. Murray Hill, N.J.: Bell Laboratories, 1980.

———. "Rational Expectations in Stationary Linear Models." *Econometrica* 49 (January 1981): 171–92.

Gould, John P. "Adjustment Costs in the Theory of Investment of the Firm." *Rev. Econ. Studies* 35 (January 1968): 47–55.

Green, Jerry R. "The Non-Existence of Informational Equilibria." *Rev. Econ. Studies* 44 (October 1977): 451–63.

Grossman, Sanford J., and Weiss, Laurence. "Heterogeneous Information and the Theory of the Business Cycle." Manuscript. Chicago: Univ. Chicago, Dept. Econ., August 1980.

Haavelmo, Trygve. *A Study in the Theory of Investment.* Chicago: Univ. Chicago Press, 1960.

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

———. "Linear Rational Expectations Models for Dynamically Interrelated Variables." In *Rational Expectations and Econometric Practice*, edited by Robert E. Lucas, Jr., and Thomas J. Sargent. Minneapolis: Univ. Minnesota Press, 1980. (*b*)

———. "Exact Rational Expectations Models: Specification and Estimation." Revised. Minneapolis: Federal Reserve Bank, January 1981.

Holt, Charles C.; Modigliani, Franco; Muth, John F.; and Simon, Herbert A. *Planning Production, Inventories, and Work Force.* Englewood Cliffs, N.J.: Prentice-Hall, 1960.

Jordan, J. S. "Expectations Equilibrium and Informational Efficiency for Stochastic Environments." *J. Econ. Theory* 16 (December 1977): 354–72.

Jorgenson, Dale W. "Capital Theory and Investment Behavior." *A.E.R. Papers and Proc.* 53 (May 1963): 247–59.

Jovanovic, Boyan. "Job Matching and the Theory of Turnover." *J.P.E.* 87, no. 5, pt. 1 (October 1979): 972–90.

Kalman, R. E. "A New Approach to Linear Filtering and Prediction Problems." *Transactions American Soc. Mechanical Engineers*, Ser. D., *J. Basic Engineering* 82 (1960): 35–45.

Keynes, John Maynard. *The General Theory of Employment, Interest and Money.* New York: Harcourt Brace, 1936.

King, Robert G. "Asset Markets and the Neutrality of Money." Unpublished paper. Rochester, N.Y.: Univ. Rochester, Dept. Econ., 1978.

Kreps, David M. "A Note on 'Fulfilled Expectations' Equilibria." *J. Econ. Theory* 14 (February 1977): 32–43.

Kwakernaak, Huibert, and Sivan, Raphael. *Linear Optimal Control Systems.* New York: Wiley, 1972.

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

Lucas, Robert E., Jr. "Adjustment Costs and the Theory of Supply." *J.P.E.* 75, no. 4, pt. 1 (August 1967): 321–34.

———. "Expectations and the Neutrality of Money." *J. Econ. Theory* 4 (April 1972): 103–24.

———. "An Equilibrium Model of the Business Cycle." *J.P.E.* 83 (December 1975): 1113–44.

———. "Econometric Policy Evaluation: A Critique." In *The Phillips Curve and Labor Markets,* edited by Karl Brunner and Allan H. Meltzer. Carnegie-Rochester Conference Series on Public Policy. Amsterdam: North-Holland, 1976.

———. "Asset Prices in an Exchange Economy." *Econometrica* 46 (November 1978): 1429–45.

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

Lucas, Robert E., Jr., and Sargent, Thomas J. "After Keynesian Macroeconomics." In *After the Phillips Curve: Persistence of High Inflation and High Unemployment.* Federal Reserve Bank of Boston, Conference Series no. 19. Boston: Federal Reserve Bank, 1978.

Mortensen, Dale T. "Generalized Costs of Adjustment and Dynamic Factor Demand Theory." *Econometrica* 41 (July 1973): 657–65.

Muth, John F. "Rational Expectations and the Theory of Price Movements." *Econometrica* 29 (July 1961): 315–35.

Phelps, Edmund S. "Introduction: The New Microeconomics in Employment and Inflation Theory." In *Microeconomic Foundations of Employment and Inflation Theory,* edited by Edmund S. Phelps et al. New York: Norton, 1970.

Pigou, Arthur C. *Industrial Fluctuations.* 2d ed. London: MacMillan, 1929.

Prescott, Edward C., and Townsend, Robert M. "Equilibrium under Uncertainty: Multiagent Statistical Decision Theory." In *Bayesian Analysis in Econometrics and Statistics: Essays in Honor of Harold Jeffreys,* edited by Arnold Zellner. Amsterdam: North-Holland, 1980.

Radner, Roy. "Rational Expectations Equilibrium: Generic Existence and the Information Revealed by Prices." *Econometrica* 47 (May 1979): 655–78.

Sargent, Thomas J. "Estimation of Dynamic Labor Demand Schedules under Rational Expectations." *J.P.E.* 86 (December 1978): 1009–44.

———. *Macroeconomic Theory.* New York: Academic Press, 1979.

———. "Interpreting Economic Time Series." *J.P.E.* 89 (April 1981): 213–48.

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: Proceedings from a Conference,* edited by Christopher A. Sims. Minneapolis: Federal Reserve Bank, 1977.

Sims, Christopher A. "Macroeconomics and Reality." *Econometrica* 48 (January 1980): 1–48. (*a*)

———. "International Evidence on Monetary Factors in Macroeconomic Fluctuations." Manuscript. Minneapolis: Univ. Minnesota, Dept. Econ., September 1980. (*b*)

Subramanian, Lakshmi. Summer Paper II. Pittsburgh: Carnegie-Mellon Univ., Graduate School Indus. Admin., September 1982.

Townsend, Robert M. "Market Anticipations, Rational Expectations, and Bayesian Analysis." *Internat. Econ. Rev.* 19 (June 1978): 481–94.

———. "Forecasting the Forecasts of Others." Manuscript. Pittsburgh: Carnegie-Mellon Univ., Graduate School Indus. Admin., 1981.

Treadway, Arthur B. "On Rational Entrepreneurial Behavior and the Demand for Investment." *Rev. Econ. Studies* 36 (April 1969): 227–39.