Townsend R. "Equilibrium Theory with Disparate Expectations and Learning: Some Issues and Methods" in *Individual Forecasting and Aggregate Outcomes: 'Rational Expectations' Examined*, eds. Edmund S. Phelps and Roman Frydman (Cambridge University Press, 1983), 169-202.

Therefore, it seems to me that the principal question left open in the wake of Kirman's chapter is whether active nonmyopic learning rules exist that can actually be beneficial to the participants when the model set forth is given a genuine dynamic optimization structure. If so, one must view the results of this chapter with some caution, as the misspecification becomes more likely to be discovered.

CHAPTER 9

Equilibrium theory with learning and disparate expectations: some issues and methods

ROBERT M. TOWNSEND

9.1. Introduction

Motivated by certain stylized facts or observations, and perhaps by a desire to develop models that offer a reliable guide for policy, economists have developed equilibrium models that are more and more sophisticated in their treatment of time and imperfect information. The profession has witnessed a movement from static models to dynamic models with perfect foresight and, more recently, with the advances of Muth (1961), Lucas and Prescott (1971), and Lucas (1972, 1975), among others, to dynamic models with uncertainty and rational expectations.[1] Muth's work (1961) was motivated by two major conclusions from studies of expectations data: "1. Averages of expectations in an industry are more accurate than naive models and as accurate as elaborate equation systems, although there are considerable cross-sectional differences of opinion. 2. Reported expectations generally underestimate the extent of changes that actually take place." Muth goes on to invoke his rational expectations hypothesis, "that the expectations of firms (or, more generally, the subjective probability distribution of outcomes) tend to be distributed, for the same information set, about the prediction of the theory (or the 'objective' probability distributions of outcomes)" and to establish that both the periodic movement of time series displayed in so-called corn–hog cycles and the second preceding observation can be explained

Helpful comments from Lars Peter Hansen, Thomas J. Sargent, John Taylor, Roswell G. Townsend, and the participants of the conference in New York and a seminar at the University of Pennsylvania, and support from the National Science Foundation and the Alfred P. Sloan Foundation, are all gratefully acknowledged. The author assumes full responsibility for any errors, as well as for the views expressed here.

[1] Needless to say, this chapter does not attempt to trace the development of the rational expectations equilibrium notion or to tie together the microeconomic and macroeconomic literatures.

with that hypothesis. Similarly, Lucas and Prescott's work (1971) was motivated by econometric considerations, a desire to develop a guide as to how variables measuring anticipated future demand should be placed in industry investment rules; in making forecasting considerations explicit, they invoked the more general rational expectations hypothesis that the entire subjective and objective probability distributions be equal. Finally, Lucas's work (1972, 1975) on business cycles was motivated by the desire to explain certain key cross-correlation and serial-correlation properties of prices, unemployment, and output and to develop models with implications about the way in which reduced-form equations might change with changes in the systematic component of stabilization policy. Thus, Lucas developed formal equilibrium analysis consistent with rational but disparate expectations across decision makers.

Despite these advances, there remains some concern with the use of the rational expectations hypothesis. It is perhaps unwise to attempt to speak for others, but this concern itself may be motivated by economic observations. Rational expectations, it is thought, require essentially uniform beliefs at some level and imply more stationarity or stability of time series than is consistent with even casual observations. Put rather crudely, decision makers do not agree about things that keep moving around. Indeed, it is easy to read Keynes (1936) on the role of expectations in stock markets and Pigou (1929) on the role of expectations in industrial fluctuations and sustain the view that equilibrium analysis is inappropriate. It was Keynes's 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 the market prices themselves. Thus, conventional valuation will be "established as the outcome of the mass psychology of a large number of ignorant individuals which is liable to change violently as the result of a sudden fluctuation in opinion due to factors which do not really make much difference to prospective yield." In this environment, Keynes believed that even expert professionals "possessing judgment and knowledge beyond that of the average private investor" would be forced to concern themselves with average opinion. Similarly, Pigou (1929) argued that the key feature of cycles is that decision makers are unaware of what others are doing and are thus bound to make mistakes, that is, to make forecast errors; of course, this is also a key idea in Lucas (1975). Further, Pigou argued that these errors can be mutually generated, that is, with errors of optimism generated by errors of pessimism, and so on.

The research program of which this chapter is a part continues in the tradition of enriching equilibrium theory, extending it to the point that it might begin to address some of the observations that Keynes and Pigou

had in mind. Thus, it follows Muth, Lucas, and Prescott, among others, adopting a strategy described by Hayek (1939), offering precise equilibrium constructs to explain relatively complicated dynamic phenomena. Of course, the equilibrium constructs of this chapter should not be taken as literal descriptions of the way the world works; as in Muth, it is not asserted here that the system of equations resembles the scratchwork of entrepreneurs in any way. Moreover, one can argue, as did Muth (1961), that rationality of expectations is a logical (albeit somewhat extreme) specification to examine, an assumption that can be modified: "Systematic biases, incomplete or incorrect information, poor memory, etc. can be examined with analytical methods based on rationality."

For this research program to be successful, equilibrium models with learning and with disparate but rational expectations must be tractable. Unfortunately, such models are not currently regarded as tractable by many in the economics profession, perhaps for a variety of reasons. First, models with learning are frequently used to address the question whether or not there will be convergence to rational expectations equilibria. Indeed, Muth (1961) and Lucas and Prescott (1971) suggested that rational expectations should be regarded as the natural outcome of some unspecified process of learning and adapting on the part of decision makers. But attempts to model that process have produced mixed conclusions; see Arrow and Green (1973), Cyert and DeGroot (1974), Townsend (1978), DeCanio (1979), Bray and Kreps (1981), Blume and Easley (1982), Bray (1982), and Frydman (1982). In general, convergence to (strong-form) rational expectations seems problematical, and there remain many open questions concerning how best to model learning processes and to pose stability questions. Second, equilibrium models with rational but disparate expectations often seem to require that decision makers not only forecast underlying economic state variables but also forecast the forecasts of others. This can be viewed as quite demanding. [See also B. Friedman (1979).] Third, economy-wide average forecasts can become an object of speculation in these models, and this in turn can lead to an infinite regress problem in which decision makers forecast the forecasts of forecasts of others, and so on. Finally, and related, Lucas (1975), Chari (1979), and Futia (1981) have indicated analytical difficulties in solving for closed-form solutions in models in which decision makers forecast from *endogenous* time series, encountering the infinite regress problem, nonlinear equations, or an infinite number of state variables.[2]

Thus, on the one hand, there is the hope that one might explain the

[2] This reference is somewhat misleading, because in one way or another these authors *do* solve their respective problems. The solution procedures offered here are not unrelated.

movement of economic time series in equilibrium models with learning and with disparate but rational expectations; on the other hand, there is the view that such models are difficult to solve. This chapter takes a step toward resolving this dilemma by devoting much of its attention to the formulation-analysis issue. That is, the basic intent of the research of which this chapter is a part is to find constructs and techniques that will prove useful in a variety of economic models. In this regard, these efforts have been somewhat successful; it is argued here that the class of linear dynamic models with learning and with disparate but rational expectations is indeed tractable under a wide variety of information specifications. Moreover, these models do deliver qualitatively interesting time series, as hoped. In particular, even in a simple setting, they deliver waves of optimism and pessimism and cyclical fluctuations in output.

This chapter proceeds by fixing a simple, partial equilibrium, industry model that might be suggested by a reading of Muth (1961), a basic structure with shocks to demand that are the sum of a (relatively) permanent component and a transitory component. There follows an investigation of what problems arise and what formulation-solution techniques are needed as the information structure is varied in the simple model. With arbitrary prior beliefs on the initial demand shock, there *is* an infinite regress problem in forecasts. But that problem has a natural solution, at least in principle, in an infinite-dimensional space (see Section 9.3). Moreover, statistical decision theory frequently can be used to pin down the source of prior beliefs, and in that way one can return to finite-dimensional state space methods, such as the Kalman filtering algorithm – Bayes's rule for linear dynamic systems (see Sections 9.4, 9.5, and 9.6). In particular, recursive structures with informed and uninformed traders are tractable. And even when simple Kalman filtering is inapplicable, the infinite regress problem can be solved by a nonlinear technique in a finite-dimensional space, the space of undetermined coefficients for moving average representations of economy-wide average forecasts (see Section 9.7).

In all these information specifications, decision makers take as given linear laws of motion of state variables, such as shocks and forecasts, that they do and/or do not see, and more often than not these laws have time-invariant coefficients. That is, as in Muth's early formulation (1961), state variables take on period-by-period realizations as if generated from well-defined stationary stochastic processes. These "self-fulfilling" stochastic processes are the analogues of the "self-fulfilling" price distributions in relatively static rational expectations models. It should be noted also that to the extent that the state variables are unobserved, the information specifications here contain unobserved components or latent variables. Such latent variable constructs have become increasingly useful to economists studying social phenomena; see, for example, Friedman (1957) and Muth (1960) on permanent income, Crawford (1976) and Jovanovic (1978) on labor turnover and job match, and Kydland and Prescott (1981) and Meltzer (1981) on fluctuations. For a more comprehensive discussion of the relationship between economic models with uncertainty and statistical decision theory see Prescott and Townsend (1980).

Finally, it may be emphasized again that in this chapter the interest in equilibrium models with learning and with disparate but rational expectations is motivated by the search for abstractions with testable implications, that is, abstractions capable of explaining the movement of economic time series. On this account, dynamic linear equilibrium models show some promise, and on this account (but certainly not others) the issue of whether or not there is convergence to strong-form rational expectations outcomes loses some of its interest (see Section 9.5). In particular, there can be convergence in such models to a steady state suitable for econometric purposes – the variance–covariance matrices of beliefs settle down, so that coefficients in the laws of motion are time-invariant. But in such a steady state, means of beliefs are subject to economic and measurement error shocks and thus move around in an interesting way, a movement that is not predicted in strong-form rational expectations equilibrium (of the same model). In fact, as noted earlier, such shocks can induce oscillations and the waves of optimism and pessimism that Pigou (1929) had in mind. More generally, the research program of which this chapter is a part builds on the contributions of Hansen and Sargent (1980*a*, 1980*b*, 1981) and Sargent (1978, 1981*a*) on the formulation and estimation of dynamic linear rational expectations models, incorporating learning and information discrepancies across decision makers. Of course, all this work is designed to circumvent Lucas's critique (1976) of conventional econometric policy evaluation.

9.2. The basic model

Following Muth (1961), a linear partial equilibrium framework is adopted. Each of a set of firms maximizes expected profit, a function that is quadratic in output. This produced commodity is sold in a competitive market with stochastic demand. Production decisions must be made prior to the realization of the demand. Subsequent to the realization of demand, the market clearing price is the one that would be determined by a Walrasian auctioneer. There are no contingent commodity markets. Under these assumptions, the production decision of each firm is a linear function of the expected price. In terms of generality, neither the linearity nor the partial equilibrium approach nor the exogenous restriction on markets nor the risk neutrality is satisfactory.

It is assumed that there are many firms, so that each regards his own contribution to market output and his influence on market beliefs as negligible. Formally, this is accomplished by supposing that the set of firms I is the unit interval. That is, firm i has label $i \in [0,1]$. Let q_{it} denote the output of firm i at time t. Let P_t denote the market price of the commodity output at time t. Output q_{it} must be chosen prior to the realization of P_t. Prior to its realization, suppose each firm i believes that P_t is a real-valued random variable with mean $E_{it}(P_t)$. Each firm acts to maximize $E_{it}(P_t)q_{it} - (1/2a)(q_{it})^2$, where $a > 0$, with respect to q_{it}. This yields the linear decision rule $q_{it} = aE_{it}(P_t)$. Also, let $Q_{it} = \int q_{it}\, di$ denote aggregate or average output at time t, where it is supposed, both here and later, that such Lebesgue integration is well defined.

In conjunction with market clearing (i.e., supply equals demand), price P_t is determined by the linear demand schedule $P_t = \theta_t + \epsilon_t - bQ_t$, where $b > 0$. Here $\{\epsilon_t\}$ is a sequence of independent and identically distributed random variables, completely transitory shocks, each of which is normally distributed with mean zero and variance σ_ϵ^2. Parameter θ_t, the relatively permanent shock, follows the first-order Markov process $\theta_{t+1} = \rho\theta_t + v_{t+1}$, where $|\rho| < 1$ and $\{v_t\}$ is a sequence of independent and identically distributed random variables, each with mean zero and variance σ_v^2, independent of the $\{\epsilon_t\}$. Finally, it should be noted that the basic model and all of its information variants that follow have steady states of zero. Thus, variables should be interpreted as deviations from means both here and later.

9.3. Arbitrary prior beliefs and the infinite regress problem in expectations: a solution[3]

Now suppose that at the end of each period t, firms see P_t but do not see (directly) aggregate output Q_t. Suppose also that both the relatively permanent and transitory components of demand, θ_t and ϵ_t, are unobserved. Then, in choosing output $q_{i,t+1}$ at the beginning of period $t+1$, firm i will attempt to forecast P_{t+1}, which, under market clearing, is a linear combination of the forecast of θ_{t+1} and Q_{t+1}, namely, $E_{i,t+1}(P_{t+1}) = E_{i,t+1}(\theta_{t+1}) - bE_{i,t+1}(Q_{t+1})$. But how does the firm forecast Q_{t+1}? Taking account of the symmetry of the situation, we might suppose that each firm recognizes that all the other firms are attempting to solve a similar forecasting problem. Hence, if firm i's output decision $q_{i,t+1}$ depends at least on $E_{i,t+1}(\theta_{t+1})$, aggregate output Q_{t+1} must depend at least on market anticipations of θ_{t+1}, namely, $\int E_{j,t+1}(\theta_{t+1})\, dj$. So firm i will attempt to forecast this market anticipation, $E_{i,t+1} \int E_{j,t+1}(\theta_{t+1})\, dj$. But again, by the symmetry of the situation, other firms form such anticipations. So again, to be one step ahead of the market, so to speak, firm i will attempt to forecast these latter forecasts $E_{i,t+1} \int E_{j,t+1} \int E_{k,t+1}(\theta_{t+1})\, dk\, dj$, and so on. Thus described is an infinite regress problem in expectations.

It is argued in this section that this infinite regress problem necessarily emerges in the basic model if one is only willing to specify arbitrary initial prior beliefs on the part of firms about the parameter θ_t and about the beliefs of other firms. But, perhaps contrary to appearance, the infinite regress problem does have a solution, at least in principle, in an infinite-dimensional space. That is, there exists a dynamic competitive equilibrium with rational but disparate expectations. Taken up in the next section is a discussion of how to resolve the infinite regress problem, perhaps in a more natural way, by truncating the infinite regress.

The discussion of this section will be aided considerably by a small investment in notation. So let $\phi_{0t} = \theta_t$ for all t. Let the prior of firm i on θ_t at time t be termed its zero-order belief on θ_t with expectation $m_{0t}(i)$. Regard the economy-wide average of these forecasts $\int m_{0t}(j)\, dj$ as another variable ϕ_{1t}. Let the prior of firm i on ϕ_{1t} at time t be termed its first-order belief on θ_t with expectation $m_{1t}(i)$. Regard the economy-wide average of these forecasts $\int m_{1t}(j)\, dj$ as another variable ϕ_{2t}. Let the prior of firm i on ϕ_{2t} be termed its second-order belief on θ_t with expectation $m_{2t}(i)$. Continuing *recursively* in this manner, let $\phi_{nt} = \int m_{n-1,t}(j)\, dj$ for all integers $n \geqslant 1$.

$E^i\phi_0$ $E^i\theta_t$ $\int E^j\theta$ $E^i\phi_1$ $\int m_{1t}$

wide average

Now define an infinite-dimensional vector of variables $\phi_t = \{\phi_{nt}, n \geqslant 0\}$. Let $m_t(i) = \{m_{nt}(i), n \geqslant 0\}$ be an infinite-dimensional vector of expectations on ϕ_t, and let $\Sigma(\phi_t)$ denote a doubly infinite dimensional matrix of covariances with $(k+1)$th row and $(n+1)$th column element $\sigma_{kn}(\phi_t)$. We shall suppose that firm i's beliefs about θ_t and about the beliefs of other firms are completely described by the infinite-dimensional normal distribution with mean $m_t(i)$ and covariance matrix $\Sigma(\phi_t)$. The covariance matrix is common knowledge. Thus, in this specification, means can vary across firms i, but covariances cannot. But both means and covariances can evolve over time. Beliefs at $t=0$ are taken as an initial condition.

resp. of

cov. of

[3] This section builds on the work of Townsend (1978), where the solution procedure was first described. However, the discussion here is marked by a strikingly simpler and more conventional definition of equilibrium – one that removes strategic considerations among decision makers (as described later). With this definition, it is easier to describe the solution algorithm, one that handles the added generality here of a Markov process on θ_t.

With this notation, the definition of a competitive equilibrium is now relatively straightforward.

Definition 1: A dynamic competitive equilibrium with rational but disparate expectations is a specification of a decision rule for each firm $i \in I$,

exp(i) of wide average.

$$q_{it} = \sum_{n=0}^{\infty} \alpha_n m_{nt}(i) = \alpha_0 m_{0t}(i) + \alpha_1 m_{1t}(i) + \text{——} \tag{1}$$

a rule for economy-wide average output,

$$Q_t = \sum_{n=0}^{\infty} \beta_n \phi_{n+1,t} \qquad \theta_{t+1} = \rho\theta_t + \vartheta_{t+1} \tag{2}$$

laws of motion for the variables ϕ_t, $\quad \phi_{0t} = \theta_t$

$$\phi_{0,t+1} = \rho\phi_{0t} + v_{t+1} \qquad \phi_{0,t+1} = \theta_{t+1} \tag{3}$$

$$\phi_{k,t+1} = \delta_{k0}(t)P_t + \sum_{n=1}^{\infty} \delta_{kn}(t)\phi_{nt} \quad (k \geqslant 1) \tag{4}$$

laws of motion for the forecasts of the ϕ_t for each firm $i \in I$,

$$m_{k,t+1}(i) = \gamma_{k+1,0}(t)P_t + \sum_{n=0}^{\infty} \gamma_{k+1,n+1}(t)m_{nt}(i) \geqslant 0 \quad (k \geqslant 0) \tag{5}$$

and market clearing equation,

$$P_t = \phi_{0t} + \epsilon_t - bQ_t \tag{6}$$

such that the following hold.

(i) *Maximization:* The decision rule (1) is maximizing for each firm i given the average output rule (2) and market clearing condition (6).

(ii) *Statistically correct forecasting:* The laws of motion for forecasts (5) are statistically correct for each firm i given the average output rule (2), market clearing condition (6), and parameter laws (3) and (4).

(iii) *Consistent aggregation in output:* The average output rule (2) is consistent with the individual decision rules (1); that is, $\alpha_n = \beta_n$, $n \geqslant 0$.

(iv) *Consistent aggregation in forecasts:* The laws of motion for forecasts (5) over all i generate the laws of motion for variables (4); that is,

$$\gamma_{kn}(t) = \delta_{kn}(t) \quad (k \geqslant 1,\ n \geqslant 0)$$

The definition is self-explanatory in large part, but some additional comments seem appropriate. One notes in particular the separation between individual and average laws, both in output and in forecasts. That is, each i firm takes the average output rule (2) and laws of motion of parameters (3) and (4) as given in determining its own output and in making its own forecast. This keeps strategic interaction at a minimum. In fact, substituting (2) into (6) completely removes the dependence of the output decision of firm i on the average decision Q_t; just as in a standard rational expectations equilibrium, there is a kind of self-fulfilling distribution in prices.

An equilibrium may seem difficult to compute, but it is not. Essentially, all one needs to do is make the substitutions suggested in consistency requirements (iii) and (iv), and then search for the undetermined coefficients $\{\alpha_n\}$ in (1) and (2) and the undetermined coefficients $\{\gamma_{kn}(t)\}$ in (4) and (5). Moreover, these searches can be conducted recursively, as is now indicated. First, use (1), (2), and (6) to determine the $\{\alpha_n\}$. Under some mild regularity conditions, there exists a solution of the form $\alpha_n = (-1)^n a(ab)^n$, $n \geqslant 0$ (see Townsend, 1978, proof of Proposition 2). Second, consider the updating problem of firm i. Using (2) and (6) again, the posterior of firm i on the variables ϕ_t at the end of period t is a weighted average of its prior and the observation P_t; see Townsend [1978, equations (12) and (13)], for a similar calculation. The weights are determined by ratios of linear combinations of covariances, all in terms of the (known) parameters of $\Sigma(\phi_t)$, σ_v^2, σ_ϵ^2, a, and b. Now attack the system recursively. The law of motion for ϕ_{0t} is given in (3). So the posterior on ϕ_{0t} (just determined) and the law (3) give the updating formula of firm i for $\phi_{0,t+1}$, that is, (5) at $k=0$. Averaging over firms i produces law of motion for $\phi_{1,t+1}$, that is, (4) at $k=1$. Now take the expectation of firm i over $\phi_{1,t+1}$ and substitute the posterior formulas for all the ϕ_t. This yields (5) at $k=1$. Then average over all firms i to produce (4) at $k=2$. One can proceed in principle to compute all the infinite laws in this manner.

Thus, the infinite regress problem has a direct solution, at least in principle. But various potential drawbacks of this setup and its solution should be noted. First, it must be verified ~~that~~ the many infinite sums are well defined. Second, and related, in computing a solution one has to hope for closed-form solutions or hope that iterative procedures can be truncated after some point. Third, the coefficients in the equilibrium representation vary over time, an aspect that can only complicate econometric work; one hopes that these converge to constants over time. Fourth, and related, one hopes that the influence of the initial priors dissipates over time. In subsequent sections, all these hopes are fulfilled with alternative setups and solution techniques.

9.4. Truncating the infinite regress: the gain from statistical decision theory

As noted, we might be happier for conceptual and computational reasons if the regress of the previous section could be truncated after some point. [Even Keynes (1936) stopped with third-order expectations in his observations on the stock market.] A sufficient way to truncate is for there to be *common knowledge of expectations at some (finite) order.* This section discusses various ways in which such common knowledge might occur.

To begin the discussion, imagine in the setup of the previous section that there is diversity at the beginning of period t regarding the forecasts of θ_t, that is, that the $m_{0t}(i)$ vary over i. But suppose for some reason or other that there is agreement on the extent of the disagreement, that is, that the average of these forecasts is known, known to be known, and so on, that is, $\int m_{0t}(j)\,dj$ is common knowledge. Then it can be established as a special case of the previous analysis that there exist equilibrium decision rules of the form

$$q_{it} = am_{0t}(i) - \frac{a(ab)}{1+ab}\int m_{0t}(j)\,dj$$

Moreover, in this case, because economy-wide output Q_t is known, statistical updating produces the result that $\int m_{0t}(j)\,dj$ will remain common knowledge in every period; see Townsend (1978, p. 487) for details.

It is problematical how common knowledge of some finite-order forecasts might be achieved. Surveys of expectations would surely help, but of course there are sampling and survey errors. In general, this is an area for future research efforts (Townsend, 1978, Section 6).

This brings us to perhaps the most compelling *modeling strategy* for truncating the infinite regress: statistical decision theory. Imagine that each decision maker's beliefs are not arbitrary but rather are the result of conditioning on a joint distribution on some specific observations. Then decision makers do not necessarily share common forecasts (conditional distributions). But they will have common knowledge of the *way* others are making their own forecasts (i.e., the statistical updating formulas), and *in some contexts this can be enough to truncate the infinite regress.* The following three sections, each of which is intended to illustrate a separate point, illustrate this gain to statistical decision theory in successively more complicated information structures.

9.5. Learning, convergence to rational expectations, and the existence of steady-state distributions: some conceptual issues

In introducing and developing the concept of rational expectations, Muth (1961) and Lucas and Prescott (1971) have argued that the notion of rational expectations is an *equilibrium* concept and should be regarded perhaps as the outcome of some unspecified process of learning and adapting on the part of economic decision makers. Recently, various authors have taken up this issue, asking whether or not models with learning converge to the rational expectations equilibrium outcome. This section addresses the convergence question.

To pose the issue of learning here, a limiting special case of the general model is considered first. Imagine that each of the parameters θ_t is equal to some constant θ, that is, that $\rho=1$ and $\sigma_v^2=0$. In this context, a rational expectations equilibrium is a self-fulfilling price distribution; that is, a distribution of prices that, if taken as given by firms and used in their maximum problems, implies output decisions that, in turn, with market clearing under the random demand schedule, imply the initial distribution of prices. More specifically,

$$P_t(\epsilon_t) = \frac{\theta}{1+ab} + \epsilon_t \tag{7}$$

is such a self-fulfilling distribution. We may now ask whether or not there will be convergence to this distribution over time if for some reason initial guesses about the distribution are wrong but learning is allowed.

One way to proceed is to imagine that decision makers, firms in this instance, have in mind a relatively simple model that they use to make decisions and update beliefs, that is, learn about one or a number of parameters. Suppose, borrowing here and later from Cyert and DeGroot (1974), that each firm believes

$$P_t = \rho P_{t-1} + \epsilon_t \tag{8}$$

where $\{\epsilon_t\}$ is a sequence of independent and identically distributed (i.i.d.) variables, each with mean zero and variance $\sigma_\epsilon^2 = 1/r$. Each firm regards ρ as a fixed but unknown parameter about which it attempts to learn. That is, ρ is regarded as a normal random variable, with mean m_t and variance $1/h_t$ at the beginning of each period t. Thus, with contemporary observations on price, P_t,

$$m_{t+1} = \frac{h_t m_t + rP_{t-1}P_t}{h_t + rP_{t-1}^2}, \quad h_{t+1} = h_t + rP_{t-1}^2 \tag{9}$$

from standard statistical updating formulas (DeGroot, 1970). To close the model, note that for the representative firm, from (8),

$$E_t(P_t) = m_t P_{t-1} \tag{10}$$

But from the actual market clearing equation with $\theta = 0$ (as is now assumed),

$$P_t = -abE_t(P_t) + \epsilon_t \tag{11}$$

Thus, substituting (10) into (11),

$$P_t = -abm_t P_{t-1} + \epsilon_t \tag{12}$$

Here, then, the relatively simple model used by firms, namely (8), is not generally consistent with the actual distribution of prices (12); in the learning process, m_t moves around, and so parameter ρ in (8) is not fixed.

We might hope for consistency in the limit. In fact, if $m_t \to 0$, then we would have convergence to a strong-form rational expectations equilibrium. But the results reported by Cyert and DeGroot (1974) on their model are not comforting in this regard; the model oscillates and even explodes under some reasonable parameter specifications in various Monte Carlo runs.

What inferences can be drawn from the failure of Cyert and DeGroot's model to converge? Perhaps little. The failure to converge might be taken to indicate that decision makers were doomed at the outset by an incorrect view of the world and a limited statistical procedure: The model firms used was wrong and was never tested by them. This, of course, raises the question of how best to model inconsistent learning processes and to conduct stability analysis. As noted, a number of authors are attacking these issues.

An alternative to postulating relatively simple (incorrect) models that decision makers use in learning is to suppose that decision makers' models are consistent with the structure of the economy. It must be emphasized immediately that this is not so demanding an assumption as it might first seem, for one can still entertain all the rich variety of information structures and learning that statistical decision theory allows. There can be uncertainty about the nature of demand, the information and expectations of other agents, the decisions taken by them, and so on.

In the context of such a consistent structure, the convergence question can again be posed. In fact, it might be asked whether or not the period-by-period price distributions of the model of Section 9.3, with the *equilibrium* definition proposed there, would converge to the self-fulfilling

price distribution described at the outset of this section. That is, would there be convergence to a (strong-form) rational expectations equilibrium? This question was answered affirmatively by Townsend (1978) for a special case, when industry output is observed period by period. More recently, Bray and Kreps (1981) have indicated that convergence (if not existence) results may be quite general. But it is argued here, as in Bray and Kreps (1981), that this particular convergence question is no longer of interest as a potential justification for rational expectations given that *rational expectations are imposed in the learning process.* To put it as bluntly as possible, such an imposition is tantamount to assuming what one wants to prove. The failure of the model to converge to a special configuration would seem to indicate little.

Of what use, then, are rational expectations models with rational learning? Again, the hope is that such models might mimic certain key features or stylized facts of the data. That is, do rational expectations models with learning have potentially testable implications, and, if so, are they consistent with data? This alternative issue leads in turn to yet another convergence question, one that *can* be answered in many contexts.

It seems unlikely a priori that a model with fixed shocks or parameter values can explain much of the interesting movement in economic time series even if there is learning and an absence of convergence. Thus, the model of this section, with θ_t shocks to demand fixed at θ for all t, loses its appeal. So we return to the more general structure in which demand is perpetually buffeted by new shocks v_t. That is, with $\theta_t = v_t + \rho v_{t-1} + \rho^2 v_{t-2} + \cdots + \rho^t \theta_0$, firms attempt to learn at the end of period t about the parameter θ_0 *and* past parameters $v_1, v_2, \ldots, v_t$. Equivalently, firms attempt to learn about the random variable θ_t at the end of each period t. See Townsend (1982) for a more detailed discussion of this equivalence.

Imagine that initially at $t = 0$ each firm has a common normal prior on the variable θ_0 with mean m_0 and variance Σ_0. Moreover, this prior belief is common knowledge. In addition, at the end of each period t each firm sees the market clearing price P_t *and* observation u_t, which is linearly related with noise to the demand shocks θ_t, namely,

$$u_t = \theta_t + w_t$$

where w_t is an i.i.d. normal random variable with mean zero and variance σ_w^2. In this context, we might hope to find an equilibrium that is consistent with considerable and interesting movement in economic time series and that is potentially testable.

The discussion will be aided considerably by redefining an equilibrium for this particular information structure. In doing so, it will prove helpful to let q_t and m_t denote the output decision and mean forecast of the representative firm, with Q_t and M_t the corresponding economy-wide averages.

Definition 2: A dynamic competitive equilibrium with rational and homogeneous expectations is a law of motion for θ_t,

$$\theta_{t+1}=\rho\theta_t+v_{t+1} \tag{13}$$

and observer equation,

$$u_t=\theta_t+w_t \tag{14}$$

a decision rule for the representation firm,

$$q_t=\alpha_0 m_t+\alpha_1 M_t \tag{15}$$

a rule for economy-wide average output,

$$Q_t=\beta M_t \tag{16}$$

a law of motion for the mean forecast of the representative firm,

$$m_{t+1}=\gamma_0 m_t+\gamma_1 P_t+\gamma_2 M_t+\gamma_3 u_t \tag{17}$$

a law of motion for the economy-wide average mean forecast,

$$M_{t+1}=\delta_1 P_t+\delta_2 M_t+\delta_3 u_t \tag{18}$$

and market clearing equation,

$$P_t=\theta_t+\epsilon_t-bQ_t \tag{19}$$

such that the following hold.

(i) *Maximization:* The decision rule (15) is maximizing given the average output rule (16) and market clearing equation (19).

(ii) *Statistically correct forecasting:* The law of motion (17) is statistically correct given the law of motion (13), observer equation (14), market clearing equation (19), average output rule (16), and average forecasting rule (18).

(iii) *Consistent output aggregation:* Rules (15) and (16) are consistent.

(iv) *Consistent forecast aggregation:* Rules (17) and (18) are consistent.

It is hoped that what strikes one about this definition is its apparent simplicity. In particular, here we have a *finite number of state variables* and *time-invariant coefficients.* Both these specifications are warranted; that is, an equilibrium with these properties can be constructed, as is now indicated.

First, note, from (18), that because M_0 is assumed to be known, and P_t and u_t are common observations, M_1 is known, and so on. This is what keeps the dimensionality finite. It only remains then to compute the undetermined time-invariant coefficients. Now equations (15), (16), and (19) are easily solved as before to yield $\alpha_0=a$, $\alpha_1=-a(ab)/(1+ab)$, $\beta=\alpha_0+\alpha_1$. To compute the coefficients in (17), use is made of the incredibly powerful Kalman filtering algorithm (e.g., Bertsekas, 1976).[4] The system described here is a special case of a model with a linear law of motion for state variables

$$x_{t+1}=Ax_t+\bar{v}_{t+1}$$

and observer equations

$$y_t=Cx_t+\bar{w}_t$$

Here $x_t=\theta_t$, $A=\rho$, $\bar{v}_t=v_t$,

$$\bar{w}_t=\begin{bmatrix}\epsilon_t\\ w_t\end{bmatrix},\quad y_t=\begin{bmatrix}P_t+b\beta M_t\\ u_t\end{bmatrix},\quad\text{and}\quad C=\begin{bmatrix}1\\1\end{bmatrix}$$

where

$$E\bar{v}_t=0,\quad E(\bar{w}_t)=\begin{bmatrix}0\\0\end{bmatrix},\quad E\bar{v}_t\bar{v}_t'=M=\sigma_v^2,\quad E\bar{w}_t\bar{w}_t'=N=\begin{bmatrix}\sigma_\epsilon^2 & 0\\ 0 & \sigma_w^2\end{bmatrix}$$

Under mild regularity conditions that are satisfied here (Bertsekas, 1976; Kwakernaak and Sivan, 1977), for arbitrary prior distributions on x_0 with mean $E_0(x_0)$ and variance $\Sigma_{0|-1}$, the *updated posterior distributions converge,* so that the means satisfy the recursive relationship

$$E_t(x_t)=AE_{t-1}(x_{t-1})+\bar{\Sigma}C'N^{-1}[y_{t-1}-CAE_{t-1}(x_{t-1})] \tag{20}$$

and the covariance matrices satisfy the recursive relationship

$$\Sigma_{t|t-1}=A[\Sigma_{t-1|t-2}-\Sigma_{t-1|t-2}C'(C\Sigma_{t-1|t-2}C'+N)^{-1}C\Sigma_{t-1|t-2}]A'+M \tag{21}$$

where

[4] Again, one could write out the entire history of observables and use standard formulas for conditional means and variances of normal random variables to forecast the entire history of innovations, and hence forecast θ_t. The Kalman filter is a recursive procedure that circumvents the dimensionality problems associated with this alternative.

$$\Sigma_{t|t-1}=E_t\{[x_t-E_t(x_t)][x_t-E_t(x_t)]'\}$$

and where $\Sigma_{t|t-1}$ converges termwise to a *time-invariant* matrix Σ. Then $\bar{\Sigma}$ in (20) satisfies

$$\bar{\Sigma}=\Sigma-\Sigma C'(C\Sigma C'+N)^{-1}C\Sigma \tag{22}$$

and its terms are variances and covariances of beliefs on the state vector x_t at the end of date t. Clearly, (20), (21), and (22) can be used to deliver the time-invariant coefficients in (17). Then (18) is determined by $\delta_1=\gamma_1$, $\delta_3=\gamma_3$, and $\delta_2=\gamma_0+\gamma_2$. Finally, it will be noted that (20) can be viewed as in place even at $t=0$ on the assumption that the system has an infinite past. This device of assuming an infinite past as the limiting case will be used again in a subsequent section. It is highly convenient for analytic and econometric purposes.

It is clear, of course, that the steady-state matrix $\bar{\Sigma}$ *in general* will not converge to zero, its value in a *period-by-period strong-form* rational expectations equilibrium, with θ_t known at the end of period t, so that $E_{t+1}(\theta_{t+1})=\rho\theta_t$. At best, we can estimate or learn past innovations v_τ, $\tau<t$, arbitrarily well if τ is arbitrarily far in the past. But contemporary θ_t is continually buffeted by new shocks v_t about which little is known. The additional information in u_t will help in forecasting (i.e., reduce $\bar{\Sigma}$) but will not cause a degeneracy unless $\sigma_w^2=0$.

Thus, the failure of the system to converge to a period-by-period strong-form rational expectations equilibrium should not come as a surprise and is certainly not damning. In fact, the freedom that the model allows in specifying indirectly the steady-state forecasting variance $\bar{\Sigma}$ may be *useful for econometric purposes,* because in principle it allows a better fit of the model to actual data. It determines the extent to which forecasts respond to contemporary observations, and this in turn determines the extent to which the system itself moves around, say in response to relatively permanent shocks. In addition, a nontrivial forecasting variance $\bar{\Sigma}$ allows the system to respond to completely transitory shocks ϵ_t and to measurement errors u_t. Finally, it should be emphasized that this linear-normal system *can,* in principle, be taken to data by the methods described in detail by Townsend (1982, Section 10).

9.6. Exploiting recursive structures with finite state space methods[5]

The information structure of the previous section allowed no diversity in forecasts; equilibrium was characterized by rational and homogeneous expectations. This section and the next consider progressively more complicated information structures that do allow diversity. In this section we have a group of informed firms and a group of uninformed firms. It is established that such structures can be successfully attacked with statistical decision theory and the use of recursive procedures. Yet, despite the diversity, and in contrast to the analysis of Section 9.3, *there is no infinite regress problem.* Moreover, such structures can produce interesting time series.

To begin, imagine that at the beginning of period zero, all firms have a common normal prior on the variable θ_0, with mean $E(\theta_0)$ and variance $\text{var}(\theta_0)$. This prior is common knowledge. Now, at the end of each period t, $t\geqslant 0$, each of a set I of informed firms, fraction ρ_I of the set of all firms, observes a random variable $u_t=\theta_t+w_t$, where w_t is an i.i.d. normal random variable with mean zero and variance σ_w^2. Each of a set U of uninformed firms, fraction ρ_U, does not observe u_t. However, *all* firms do see the market clearing price P_t at the end of each period t. Industry output, on the other hand, is unobserved directly, though it may be inferred by some.

To attack this structure, we shall first uncover a *finite* list of state variables that will allow us to write down the decisions of firms and the evolution of their beliefs. All these state variables will be known in each period to the informed firms, but uninformed firms will know only a subset. Of course, θ_t remains unknown to all.

To begin, let $E_{It}(\theta_t)$ and $\Sigma_{It}(\theta_t)$ denote the beginning-of-period prior mean and variance on θ_t of informed firms. Now, for uninformed firms, $E_{It}(\theta_t)$ is unknown, because the u_t are unobserved by them. But we may suppose they regard $E_{It}(\theta_t)$ as an unknown variable π_t that is jointly normally distributed with the unknown variable θ_t; that is, $[\theta_t,\pi_t]$ is normally distributed, with mean $[E_{Ut}(\theta_t),E_{Ut}(\pi_t)]$ and covariance matrix $\Sigma_{Ut}(\theta_t,\pi_t)$. Then the state variables mentioned earlier are π_t, $\Sigma_{It}(\theta_t)$, $E_{Ut}(\theta_t)$, $E_{Ut}(\pi_t)$, and $\Sigma_{Ut}(\theta_t,\pi_t)$. Finally, note that $\Sigma_{It}(\theta_t)$ and $\Sigma_{Ut}(\theta_t,\pi_t)$ are assumed to be common knowledge.

We are now in a position to redefine an equilibrium with this state space.

Definition 3: A dynamic competitive equilibrium with rational but disparate expectations is a law of motion for θ_t,

$$\theta_{t+1}=\rho\theta_t+v_{t+1} \tag{23}$$

an observer equation,

[5] This specification was considered by Prescott and Townsend (1980). But again, the discussion here is marked by an alternative definition of equilibrium in the more general setting in which θ_t's move around. Here, also, new Kalman filtering results are obtained, and a general method for attaching such structures is described.

$$u_t = \theta_t + w_t \tag{24}$$

a law of motion for π_t,

$$\pi_{t+1} = \xi_{1t}\pi_t + \xi_{2t}\theta_t + \xi_{3t}w_t + \xi_{4t}\epsilon_t \tag{25}$$

a decision rule for uninformed firms,

$$q_{Ut} = \alpha_1 E_{Ut}(\theta_t) + \alpha_2 E_{Ut}(\pi_t) \tag{26}$$

a decision rule for informed firms,

$$q_{It} = \beta_0 E_{It}(\theta_t) + \beta_1 E_{Ut}(\theta_t) + \beta_2 E_{Ut}(\pi_t) \tag{27}$$

laws of motion for forecasts and covariance matrix of beliefs of uninformed firms,

$$E_{U,t+1}(\theta_{t+1}) = \gamma_{1t}E_{Ut}(\theta_t) + \gamma_{2t}E_{Ut}(\pi_t) + \gamma_{3t}P_t \tag{28}$$

$$E_{U,t+1}(\pi_{t+1}) = \gamma_{4t}E_{Ut}(\theta_t) + \gamma_{5t}E_{Ut}(\pi_t) + \gamma_{6t}P_t \tag{29}$$

$$\Sigma_{U,t+1}(\theta_{t+1}, \pi_{t+1}) = \psi[\Sigma_{Ut}(\theta_t, \pi_t), \sigma_v^2, \sigma_w^2, \sigma_\epsilon^2] \tag{30}$$

laws of motion for forecasts and variance of beliefs of informed firms,

$$E_{I,t+1}(\theta_{t+1}) = \delta_{1t}E_{It}(\theta_t) + \delta_{2t}P_t + \delta_{3t}u_t + \delta_{4t}E_{Ut}(\theta_t) + \delta_{5t}E_{Ut}(\pi_t) \tag{31}$$

$$\Sigma_{I,t+1}(\theta_{t+1}) = \Phi[\Sigma_{It}(\theta_t), \sigma_v^2, \sigma_w^2, \sigma_\epsilon^2] \tag{32}$$

and market clearing equation,

$$P_t = \theta_t + \epsilon_t - b[\rho_I q_{It} + \rho_U q_{Ut}] \tag{33}$$

such that the following hold.

(i) *Maximization:* The decision rules (26) and (27) are each maximizing given the rules (26) and (27) and the market clearing equation (33).

(ii) *Statistically correct forecasting for informed firms:* The laws of motion (31) and (32) are statistically correct given the decision rules (26) and (27), market clearing equation (33), parameter law (23), and observer equation (24).

(iii) *Statistically correct forecasting for uninformed firms:* The laws of motion (28), (29), and (30) are statistically correct given the decision rules (26) and (27), market clearing equation (33), parameter laws of motion (23) *and* (25), and observer equation (24).

One comment is immediately in order. Unlike previous definitions, there is no distinction here between individual and aggregate decisions and forecasts. But it should be understood that the representative informed firm takes as given that the output decisions and forecasts of other informed firms are identical with its own, and similarly for uninformed firms.

An equilibrium is not difficult to compute. First, the coefficients α_1, α_2, β_0, β_1, and β_2 in (26) and (27) are computed by the method of undetermined coefficients, using property (i) of the equilibrium.[6] Second, the forecasting formulas (31) and (32) for informed firms are readily computed by Bayesian considerations; essentially, each informed firm is getting two independent observations on θ_t at the end of period t, through u_t and $P_t - b[\rho_{It}q_{It} + \rho_{Ut}q_{Ut}]$, and the law of motion (23) is known. For this purpose it is important to remember that q_{Ut} (and q_{It}) is expressed in terms of state variables that informed firms know. Now (25) follows immediately from (31), with $\pi_t = E_{It}(\theta_t)$. Finally, one has to attack the forecasting formulas for the uninformed firms. Using (33) with (31) and (27) [and (26)], uninformed firms are getting one observation θ_t and π_t through P_t - recall again $\pi_t = E_{It}(\theta_t)$. So the posterior $E_{U,t+1}(\theta_t)$ and $E_{U,t+1}(\pi_t)$ may be formed in the usual Bayesian way. Now the parameter laws of motion (23) and (25) for θ_{t+1} and π_{t+1} are known; so priors on π_{t+1} and θ_{t+1} at the beginning of period $t+1$ are readily calculated. This yields (28), (29), and (30).

An unfortunate aspect of the foregoing definition is the presence of time-varying coefficients - the coefficients are functions of period-by-period covariances that are allowed to move around. As it turns out, this is not a problem. First, steady-state Kalman filtering considerations can be used to remove the time dependence of Σ_{It}. This removes time dependence from (31) and hence (25). It then becomes apparent that the inference problem of uninformed firms can be cast as a classic optimal observer problem as well. To see this, consider the state vector $x_t = [\theta_t, \pi_t]'$. Then uninformed firms see $y_t = P_t + b\rho_U q_{Ut} + b\rho_I\beta_1 E_{Ut}(\theta_t) + b\rho_I\beta_2 E_{Ut}(\pi_t) = \theta_t + \epsilon_t - b\rho_I\beta_0\pi_t$, so that

$$y_t = Cx_t + \bar{w}_t \tag{34}$$

where $C = [1 - b\rho_I\beta_0]$ and $\bar{w}_t = \epsilon_t$. Of course, x_t has the linear law of motion

[6] These undetermined coefficients are determined by the equations

$$\beta_0 = a[1 - b\rho_I\beta_0]$$
$$\beta_1 = a[-b\rho_U\alpha_1 - b\rho_I\beta_1]$$
$$\beta_2 = a[-b\rho_I\beta_2 - b\rho_U\alpha_2]$$
$$\alpha_1 = a[1 - b\rho_I\beta_1 - b\rho_U\alpha_1]$$
$$\alpha_2 = a[-b\rho_I\beta_0 - b\rho_I\beta_2 - b\rho_U\alpha_2]$$

It can be verified directly that these linear equations do have a solution.

$$x_{t+1} = Ax_t + \bar{v}_{t+1} \tag{35}$$

from (23) and (25), where

$$A = \begin{bmatrix} \rho & 0 \\ \xi_2 & \xi_1 \end{bmatrix}, \quad \bar{v}_{t+1} = \begin{bmatrix} v_{t+1} \\ \xi_3 w_t + \xi_4 \epsilon_t \end{bmatrix}$$

In the optimal observer system (34)–(35), the measurement error noise $\bar{w}_t$ and state excitation noise $\bar{v}_{t+1}$ are correlated. But Kalman filtering can handle this correlation – the covariance matrix of beliefs for uninformed firms converges to some constant, removing the time dependence from Σ_{Ut} and hence from (28) and (29).

It may now be noted how the hierarchical information structure and statistical decision theory allow a recursive attack on the system. First, it is supposed that there is a common prior distribution that is common knowledge. Informed firms know that the forecast of uninformed firms is the mean of their conditional distribution, the common distribution conditioned on variables that the uninformed firms *and* the informed firms see. So informed firms know the forecast of uninformed firms. Uninformed firms do not know the actual forecast of informed firms. But they do know that the forecast of informed firms is the mean of their conditional distribution, the common distribution conditioned on variables that the informed firms see, but that uninformed firms do not see completely. By Bayes's rule, then, uninformed firms know how the unobserved forecast of informed firms moves period by period, namely, as a function of the underlying unobserved state variables, observed and unobserved information variables, and the forecast of uninformed firms, which, of course, uninformed firms know. So uninformed firms can forecast the forecast of informed firms. There is no infinite regress problem: *Recursive structures are tractable by finite state space methods.*

As a final note, we might remark on the potential of recursive structures to generate interesting time series, as displayed in Table 9.1 and Figures 9.1 and 9.2. Suppose, for example, the industry is in a steady state at time t $(t=0)$ and is subjected to a unit-variance shock ϵ_t, holding all other contemporary and future shocks at zero. All firms' forecasts of θ_t (which actually remains zero) will increase, thereby increasing industry output. Sustained increases of output (above the steady state of zero) for the first three periods keep prices low (below zero) for the first three periods, and uninformed firms come to believe that θ_t is actually negative. Consequently, uninformed firms reduce output (below zero), causing prices to go positive for periods four through seven. As a result, their forecasts go positive again for periods eight through ten, and prices go negative. This oscillatory pattern continues for some time, eventually

Table 9.1. *Response to a transitory ϵ_t shock*[a]

t	P_t	q_{It}	q_{Ut}	θ_t	$\pi_t = E_{It}(\theta_t)$	$E_{Ut}(\theta_t)$	$E_{Ut}(\pi_t)$
1	−0.30353	0.05970	0.31636	0.00000	0.36322	0.61102	0.17714
2	−0.26419	−0.16979	0.28703	0.00000	0.09439	0.56990	0.48685
3	−0.08535	−0.06082	0.09305	0.00000	0.02453	0.19823	0.44104
4	0.06403	0.07040	−0.07110	0.00000	0.00637	−0.12748	0.16707
5	0.10680	0.10845	−0.11813	0.00000	0.00166	−0.22883	−0.08042
6	0.06504	0.06547	−0.07191	0.00000	0.00043	−0.14487	−0.16611
7	0.00235	0.00246	−0.00260	0.00000	0.00011	−0.01026	−0.11149
8	−0.03357	−0.03354	0.03710	0.00000	0.00003	0.07003	−0.01330
9	−0.03282	−0.03281	0.03628	0.00000	0.00000	0.07142	0.04884
10	−0.01237	−0.01237	0.01367	0.00000	0.00000	0.02858	0.05319
11	0.00643	0.00643	−0.00710	0.00000	0.00000	−0.01243	0.02311
12	0.01284	0.01281	−0.01419	0.00000	0.00000	−0.02739	−0.00754
13	0.00853	0.00853	−0.00942	0.00000	0.00000	−0.01889	−0.01979
14	0.00095	0.00095	−0.00105	0.00000	0.00000	−0.00269	−0.01444
15	−0.00380	−0.00380	0.00420	0.00000	0.00000	0.00787	−0.00268
16	−0.00409	−0.00409	0.00452	0.00000	0.00000	0.00887	0.00542
17	−0.00175	−0.00175	0.00194	0.00000	0.00000	0.00400	0.00657
18	0.00060	0.00060	−0.00067	0.00000	0.00000	−0.00112	0.00319
19	0.00153	0.00153	−0.00169	0.00000	0.00000	−0.00325	−0.00062
20	0.00111	0.00111	−0.00122	0.00000	0.00000	−0.00244	−0.00233

[a] $\sigma_v^2 = 1.00000$; $\sigma_\epsilon^2 = 1.00000$; $\sigma_w^2 = 1.00000$; $\rho = 0.95000$; %inf $= 0.05000$; $a = 1.00000$; $b = 1.00000$.

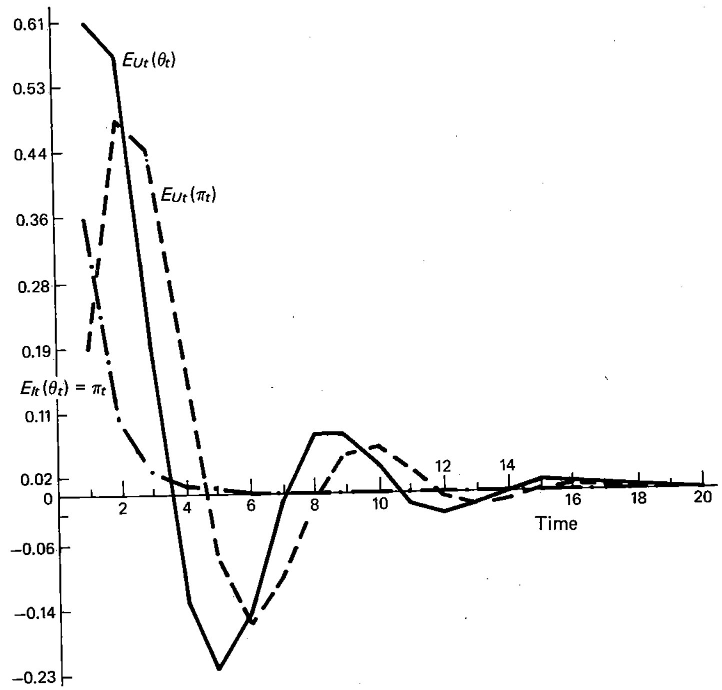

Figure 9.1. Response to transitory shock – forecasts.

converging to zero. Now, if one takes the difference between θ_t and $E_{Ut}(\theta_t)$ as a forecast error, then in this scenario, forecast errors oscillate, delivering the waves of optimism and pessimism that Pigou (1929) had in mind. This may be contrasted with the more conventional exponential decay of forecasts of informed firms (who figure out that the shock is zero within five periods, virtually).

It should be emphasized here that uninformed firms *do* take into account their own quantity movements in forecasting, but they remain uncertain about *informed* firms' outputs. It is thus that a series of low prices can cause them to put weight on θ_t being low. Of course, these forecasts themselves *are* the best possible given the uninformed firms' information set and their knowledge of the economy; that is, the forecast error of θ_t that uninformed firms expect or forecast at the beginning of period t given their information then *is* zero. It should also be emphasized in this regard that the scenario described here is not the only possible sequence of events that can take place in the economy, and uninformed firms, as good Bayesians, put positive weight on other possible scenarios generating the random variable they see (in fact, innovation accounting experiments like the one described take place with probability zero, though they serve us well in characterizing the time series, here picking up a degree of volatility, for example).

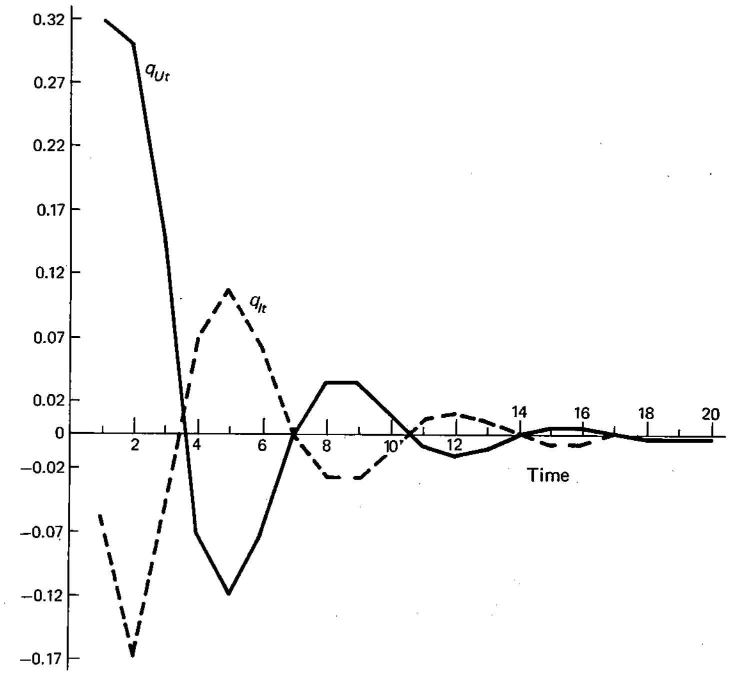

Figure 9.2. Response to transitory shock – outputs.

Despite its simplicity, the present industry model can thus be used to *begin* an analysis of some of the observations made by Keynes (1936). In particular, factors that make no real difference to prospective profitability, such as completely transitory shocks, trigger oscillations in less informed opinion, this in turn causing fluctuations in the market price. Moreover, informed firms do take less informed opinion into account, though in this industry setting the latter influence is stabilizing. This can be seen from Table 9.1 and Figure 9.2. Again, after four periods, informed firms know the shock is completely transitory, but they continue to respond in a negative way to output decisions of the uninformed, ending up on the other side of the steady state, as it were.

9.7. Equilibrium in a symmetric but disparate information structure: the infinite regress problem again with a new solution[7]

As noted, the information structure of the previous section allows a recursive attack on the system. But imagine a symmetric information structure in which *each* firm receives its own information on the underlying state variable of interest, namely, θ_t. That is, suppose that, in addition to P_t, each firm i sees at the end of period t a random variable

$$u_{it}=\theta_t+w_{it} \tag{36}$$

where the w_{it} are regarded as normally distributed random variables, independent both across firms and over time, each with mean zero and variance $\sigma_w^2 \neq 0$. With industry output unobserved, each firm is getting observations on θ_t that are confounded through P_t with the forecasts other firms made on θ_t last period, which under (36) will not be known. Thus, to correctly forecast θ_t, each firm must forecast the economy-wide average forecasts of θ_t last period. Thus, each firm's forecast of θ_t depends on its forecast of the economy-wide average forecast of θ_t last period. Thus, the economy-wide average forecast of θ_t depends on the economy-wide average forecast of the economy-wide average forecast of θ_t last period. Now the economy-wide average forecast of θ_t last period must have a similar property. Thus, to correctly forecast θ_t, each firm must forecast the economy-wide average forecast last period of the economy-wide average forecast of θ_t two periods ago, and so on. It is apparent that there is an infinite regress problem here, despite the use of statistical decision theory.

This section describes a technique for overcoming this difficulty, a technique of undetermined coefficients, as in Townsend (1982). It is argued, moreover, that the information specification here, like the one in the previous section, has interesting implications for economic time series.

We begin by setting out the definition of equilibrium for this section in rather general terms. For notation, let Ω_{it} denote the information set of firm i at the beginning of period t. Let Ω_{it} contain at least past prices $P_{t-1}, P_{t-2}, \ldots$ and past information variables $u_{i,t-1}, u_{i,t-2}, \ldots$. Also, let $E(P_t \mid \Omega_{it})$ denote firm i's forecast of price P_t conditioned on the information set Ω_{it}, and let $\bar{P}_t$ denote the economy-wide average of these forecasts. This yields the following definition.

[7] This section draws heavily on some recently developed techniques described by Townsend (1982). They are applied here to a model with *within-market interactions,* a feature missing from the work of Townsend (1982). Helpful comments from Lars Hansen on the procedures of this section are gratefully acknowledged.

Definition 4: A dynamic competitive equilibrium with rational but possibly disparate expectations is a law of motion for the parameter θ_t,

$$\theta_{t+1}=\rho\theta_t+v_{t+1} \tag{37}$$

an observer equation for each firm i,

$$u_{it}=\theta_t+w_{it} \tag{38}$$

a decision rule for each firm i,

$$q_{it}=aE(P_t \mid \Omega_{it}) \tag{39}$$

a rule for economy-wide average output,

$$Q_t=a\bar{P}_t \tag{40}$$

a forecasting formula for each firm i,

$$E(P_t \mid \Omega_{it})=C(L)v_{t-1}+D(L)\epsilon_{t-1}+E(L)w_{i,t-1} \tag{41}$$

a specification of the economy-wide average forecast $\bar{P}_t$,

$$\bar{P}_t=A(L)v_{t-1}+B(L)\epsilon_{t-1} \tag{42}$$

and market clearing equation,

$$P_t=\theta_t+\epsilon_t-bQ_t \tag{43}$$

such that the following hold.

(i) *Statistically correct forecasting:* The forecast (41) is statistically correct given the parameter law (37), observer equation (38), information set Ω_{it}, the market clearing equation (43), average output rule (40), and the rule for economy-wide average forecast (42).

(ii) *Consistent aggregation in forecasts:* Formulas (41) and (42) are consistent; that is, the one-sided polynomials in the lag operator $A(L)$, $B(L)$, $C(L)$, $D(L)$, $E(L)$ satisfy $C(L)=A(L)$ and $D(L)=B(L)$.

Some comments are immediately in order. First, individual firm maximization and consistent aggregation in output have already been imposed, as is clear from (39) and (40). Second, the main task in the construction of an equilibrium is the simultaneous and consistent determination of individual and average forecasts. Note, in particular, that the average rule (42) is used by individual firms in the determination of (41).

The general procedure for the determination of an equilibrium is the method of undetermined coefficients. One begins with an arbitrary specification of (42), that is, with coefficients in $A(L)$ and $B(L)$ unknown to the modeler but assumed known to the firms. Then, in some manner, the forecasting problems of individual firms are solved. This yields (41) in terms of the undetermined coefficients. Then (41) is averaged over firms, with the w_{it} disappearing (for this purpose, one should view the set of firms as countably infinite). The average is then matched up with (42) to determine the undetermined coefficients.

There is a potential dimensionality problem with the foregoing procedure. To see this, u_{it} and P_t can be written in terms of the economic innovations ϵ_t and v_t, as well as the measurement error, w_{it}, namely,

$$u_{it}=\sum_{j=0}^{\infty}\rho^j v_{t-j}+w_{it} \tag{44}$$

$$P_t=\sum_{j=0}^{\infty}\rho^j v_{t-j}-abA(L)v_{t-1}-abB(L)\epsilon_{t-1}+\epsilon_t \tag{45}$$

Cast in this form, it appears that each firm should be forecasting *all* current and past shocks to demand. So, to simplify the problem, following Chari's (1979) solution to his own (infinite) dimensionality problem, suppose in addition that the entire history of shocks and measurement errors from $t-j$ on is known at time t. For example, let $j=2$; then v_{t-2}, $v_{t-3},\ldots,\epsilon_{t-2},\epsilon_{t-3},\ldots$ are included in Ω_{it}. Then equations (44) and (45) may be written as

$$u_{it}^*=v_t+\rho v_{t-1}+w_{it} \tag{46}$$

$$P_t^*=v_t+\rho v_{t-1}-abA_0 v_{t-1}-abB_0\epsilon_{t-1}+\epsilon_t \tag{47}$$

where the variables with asterisks are defined implicitly.

There are two methods that can be used to forecast the innovations v_t, v_{t-1}, ϵ_t, and ϵ_{t-1} from (46) and (47). Both are described by Townsend (1982) and summarized here. For the first method, write out all current and past observables that contain information on the unknown innovations; that is, write out (46) and (47) at t *and* $t-1$. Then use standard formulas for conditional means and variances of normal random variables. This delivers the forecasts of the innovations in terms of observables and hence, by substitution, in terms of the innovations themselves. This method is straightforward, but it is not recursive; a dimensionality problem begins to arise as the lag to full information increases, that is, as $j\to\infty$.

For the second method, the observer system (46)–(47) for each firm i may be written as

$$\begin{bmatrix}\epsilon_{t-2}\\ v_{t-2}\\ u_{it}^*\\ P_t^*\end{bmatrix}=\begin{bmatrix}0 & 0 & L^2\\ L^2 & 0 & 0\\ (1+\rho L) & 1 & 0\\ (1+\alpha L) & 0 & (1-\delta L)\end{bmatrix}\begin{bmatrix}v_t\\ w_{it}\\ \epsilon_t\end{bmatrix} \tag{48}$$

or

$$y_t=M(L)\xi_t \tag{49}$$

where $\alpha=\rho-abA_0$ and $\delta=abB_0$. It is from this system that each firm i will attempt to forecast v_t, v_{t-1}, ϵ_t, and ϵ_{t-1} at the end of period t. (Note that only two undetermined coefficients, A_0 and B_0, are being carried into this forecasting problem.)

System (48) cannot be used for forecasting because the representation is not fundamental relative to firm i's information set. That is, the variables on the right-hand side are not in the space spanned by the current and past values of variables on the left-hand side. To see this, note that $M(z)$ is not of full rank at $z=0$, a point inside the unit circle. But it is possible to convert (49) to a moving average representation that is fundamental, say

$$y_t=M^{**}(L)\xi_t^{**} \tag{50}$$

where

$$M^{**}(L)=M(L)WB(L)\hat{W}B(L) \tag{51}$$

$$\xi_t^{**}=B(L^{-1})'\hat{W}'B(L^{-1})'W'\xi_t \tag{52}$$

where the matrices W, $\hat{W}$, and $B(z)$ are orthogonal. This procedure is described in detail by Townsend (1982) and may be applied directly here. In fact, the matrices W, $\hat{W}$, and $B(z)$ are identical with those of Townsend (1982). System (50) can be used directly for forecasting via the Wiener-Kolmogorov formulas. This delivers forecasts of v_t, v_{t-1}, ϵ_t, and ϵ_{t-1} as linear functions of current and past ξ_t^{**}. Then (52) can be used to deliver these forecasts in terms of the unobserved shocks ξ_t. One can then proceed as indicated earlier.

It should be noted that *this procedure does not eliminate the infinite regress in expectations.* Forecasts of $\bar{P}_t$ will vary across firms, and thus one could compute an economy-wide average of these latter forecasts, and so on, perhaps indefinitely. *But it is no longer necessary to determine these forecasts to compute the equilibrium.*

A nice aspect of the foregoing procedure is that one can guarantee the existence of solutions for the undetermined coefficients via a fixed-point argument even though the equations in the undetermined coefficients are nonlinear. And one hopes, from Townsend (1982), that such solutions trary coefficients in the moving average representation (42) produce coefficients for firms' forecasts that in turn can be averaged to yield another set of coefficients in the representation (42).

the model was somewhat different.

As a special case of the model here, let $\sigma_\epsilon^2=0$, so that there is no transitory component to demand. Still, with economy-wide average output

sional extraction problem. But on the assumption of an infinite past,

one can search for a moving average representation for the economy-wide average price forecast of the form (42). In fact, one can solve for the undetermined coefficients in the representation, namely,

by current and past P_t's is identical with the space spanned by current and past v_t's; so the entire (infinite) history of the v_t's, and thus θ_t, is known to all firms at the end of period t [the observer equations (38) are redun-

$E(P_t)=\rho\theta_t/(1+\rho b)$.

More generally, though with $\sigma_\epsilon^2\neq0$, and of course $\sigma_w^2\neq0$, there will be diversity in forecasts, and economy-wide average output will not be

describing aspects of reality. Of course, no model should pretend to describe reality completely. Thus, rather than being deterred by assumptions that can cause analytical difficulties, the emphasis here has been on

tractable. It turns out that the time series of these models do display interesting oscillations and can in principle be fit to data. So, consistent with the basic premise, an empirical application may soon be warranted.

front the "change-of-regime" issue and ask what determines decision makers' expectations; see Phelps (1980) and Sargent (1981*b*). To the extent that decision makers' expectations are arbitrary, one is led to an

general political sentiment, and past experience.

tings." Working paper No. 33, Institute for Mathematical Studies in the Social Sciences, Stanford University.

demic Press.

Blume, L. E., and D. Easley (1982). "Learning to Be Rational." *Journal of Economic*

Bray, M. (1982). "Learning, Estimation and the Stability of Rational Expectations." *Journal of Economic Theory,* 26:318–39.

Research paper No. 616, Graduate School of Business, Stanford University.

Linear Stochastic Models with Decentralized Information." Unpublished manuscript,

Crawford, R. G. (1976). "An Empirical Investigation of a Dynamic Model of Labor Turn-

manufacturing industries." Dissertation, Carnegie-Mellon University.

Journal of Political Economy, 82:521–36.

Expectations and Learning from Experience." *Quarterly Journal of Economics,*

DeGroot, M. H. (1970). *Optimal Statistical Decisions.* New York: McGraw-Hill.

and the Extreme Information Assumptions of 'Rational Expectations' Macro Models." 5:25–42.

for the NBER.

Frydman, R. (1981). "Towards an Understanding of Market Processes: Individual Expec-

Unpublished manuscript, New York University.

Futia, C. A. (1981). "Rational Expectations in Stationary Linear Models." *Econometrica,*

Hansen, L. P., and T. J. Sargent (1980*a*). "Formulating and Estimating Dynamic Linear Rational Expectations Models." *Journal of Economic Dynamics and Control,* 2:7–46.

ables." In: *Rational Expectations and Econometric Practice,* edited by R. E. Lucas, Jr., and T. J. Sargent, pp. 127–56. University of Minnesota Press.

Reserve Bank of Minneapolis, Staff Report No. 71.

Price Expectations, Monetary Disturbances and Malinvestments."

in Business Cycle Theory, pp. 351–65. Philadelphia: American Economic Association.

Matching and the Theory of Turnover." Ph.D. dissertation, University of Chicago.

Keynes, J. M. (1936). *The General Theory of Employment, Interest, and Money.* New

are unique and can be computed quickly via an iterative procedure. Arbitrary coefficients in the moving average representation (42) produce coefficients for firms' forecasts that in turn can be averaged to yield another set of coefficients in the representation (42). For Townsend (1982), such a procedure converged quickly to a unique solution, though the model was somewhat different.

As a special case of the model here, let $\sigma_\epsilon^2=0$, so that there is no transitory component to demand. Still, with economy-wide average output unobserved, at least not directly, firms face an apparently nontrivial signal extraction problem. But on the assumption of an infinite past, one can search for a moving average representation for the economy-wide average price forecast of the form (42). In fact, one can solve for the undetermined coefficients in the representation, namely, $A_j=\rho^{j+1}/(1+ab)$, $B_j=0$, for all j. With these coefficients, the space spanned by current and past P_t's is identical with the space spanned by current and past v_t's; so the entire (infinite) history of the v_t's, and thus θ_t, is known to all firms at the end of period t [the observer equations (38) are redundant]. It is easy to verify then that this information specification yields a period-by-period strong-form rational expectations equilibrium, with $E(P_t)=\rho\theta_{t-1}/(1+ab)$.

More generally, though with $\sigma_\epsilon^2\neq0$, and of course $\sigma_w^2\neq0$, there will be diversity in forecasts, and economy-wide average output will not be inferred. Then the effect of ϵ_t shocks can produce interesting time series.

9.8. Concluding remarks

This chapter began with the premise that equilibrium models with learning and with disparate but rational expectations will prove to be useful in describing aspects of reality. Of course, no model should pretend to describe reality completely. Thus, rather than being deterred by assumptions that can cause analytical difficulties, the emphasis here has been on specification strategies and solution techniques that make these models tractable. It turns out that the time series of these models do display interesting oscillations and can in principle be fit to data. So, consistent with the basic premise, an empirical application may soon be warranted.

One caveat is in order: Economic policy issues have not been considered. When considering the effect of a "new" policy, one must confront the "change-of-regime" issue and ask what determines decision makers' expectations; see Phelps (1980) and Sargent (1981*b*). To the extent that decision makers' expectations are arbitrary, one is led to an infinite regress problem. On the other hand, it may well be that statistical decision theory will prove to be a useful construct here as well, with priors of decision makers linked to government policy announcements, general political sentiment, and past experience.

References

Arrow, K. J., and J. R. Green (1973). "Notes on Expectations Equilibria in Bayesian Settings." Working paper No. 33, Institute for Mathematical Studies in the Social Sciences, Stanford University.

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

Blume, L. E., and D. Easley (1982). "Learning to Be Rational." *Journal of Economic Theory,* 26:340–51.

Bray, M. (1982). "Learning, Estimation and the Stability of Rational Expectations." *Journal of Economic Theory,* 26:318–39.

Bray, M., and D. M. Kreps (1981). "Rational Learning and Rational Expectations." Research paper No. 616, Graduate School of Business, Stanford University.

Chari, V. V. (1979). "Linear Stochastic Models with Decentralized Information." Unpublished manuscript, Carnegie-Mellon University.

Crawford, R. G. (1976). "An Empirical Investigation of a Dynamic Model of Labor Turnover in U.S. Manufacturing Industries." Dissertation, Carnegie-Mellon University.

Cyert, R. M., and M. H. DeGroot (1974). "Rational Expectations and Bayesian Analysis." *Journal of Political Economy,* 82:521–36.

DeCanio, S. (1979). "Rational Expectations and Learning from Experience." *Quarterly Journal of Economics,* 93:47–57.

DeGroot, M. H. (1970). *Optimal Statistical Decisions.* New York: McGraw-Hill.

Friedman, B. (1979). "Optimal Expectations and the Extreme Information Assumptions of 'Rational Expectations' Macro-Models." *Journal of Monetary Economics,* 5:23–42.

Friedman, M. (1957). *A Theory of the Consumption Function.* Princeton University Press for the NBER.

Frydman, R. (1981). "Towards an Understanding of Market Processes: Individual Expectations, Market Behavior and Convergence to Rational Expectations Equilibrium." Unpublished manuscript, New York University.

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

Hansen, L. P., and T. J. Sargent (1980*a*). "Formulating and Estimating Dynamic Linear Rational Expectations Models." *Journal of Economic Dynamics and Control,* 2:7–46.

(1980*b*). "Linear Rational Expectations Models for Dynamically Interrelated Variables." In: *Rational Expectations and Econometric Practice,* edited by R. E. Lucas, Jr., and T. J. Sargent, pp. 127–56. University of Minnesota Press.

(1981). "Exact Rational Expectations Models: Specification and Estimation." Federal Reserve Bank of Minneapolis, Staff Report No. 71.

Hayek, F. A. (1939). "Price Expectations, Monetary Disturbances and Malinvestments." In: *Profits, Interest, and Investment.* London: Routledge; reprinted (1944) in *Readings in Business Cycle Theory,* pp. 351–65. Philadelphia: American Economic Association.

Jovanovic, B. (1978). "Job Matching and the Theory of Turnover." Ph.D. dissertation, University of Chicago.

Keynes, J. M. (1936). *The General Theory of Employment, Interest, and Money.* New York: Harcourt Brace & World.

Kwakernaak, H., and R. Sivan (1977). *Linear Optimal Control Systems.* New York: Wiley.
Kydland, F., and E. C. Prescott (1981). "Time to Build and Equilibrium Persistence of Unemployment." Unpublished manuscript, Carnegie-Mellon University.
Lucas, R. E., Jr. (1972). "Expectations and the Neutrality of Money." *Journal of Economic Theory,* 4:103–24.
(1975). "An Equilibrium Model of the Business Cycle." *Journal of Political Economy,* 83:1130–44.
(1976). "Econometric Policy Evaluation: A Critique." In: *The Phillips Curve and Labor Markets,* edited by K. Brunner and A. H. Meltzer, Carnegie-Rochester Conference Series on Public Policy. Amsterdam: North Holland.
Lucas, R. E., Jr., and E. C. Prescott (1971). "Investment Under Uncertainty." *Econometrica,* 39:659–81.
Meltzer, A. H. (1981). "Rational Expectations, Risk, Uncertainty, and Market Responses." Manuscript prepared for the NYU Conference on Crises in the Economic and Financial Structure, November, 1981.
Muth, J. F. (1960). "Optimal Properties of Exponentially Weighted Forecasts." *Journal of the American Statistical Association,* 55:299–306.
(1961). "Rational Expectations and the Theory of Price Movements." *Econometrica,* 29:315–35.
Phelps, E. S. (1980). "The Trouble with 'Rational Expectations' and the Problem of Inflation Stabilization." Unpublished manuscript, Columbia University.
Pigou, A. C. (1929). *Industrial Fluctuations,* 2nd ed. London: Macmillan.
Prescott, E. C., and R. M. Townsend (1980). "Equilibrium Under Uncertainty: Multiagent Statistical Decision Theory." In: *Bayesian Analysis in Econometrics and Statistics: Essays in Honor of Harold Jeffreys,* edited by Arnold Zellner, pp. 169–94. Amsterdam: North Holland.
Sargent, T. J. (1978). "Estimation of Dynamic Labor Demand Schedules Under Rational Expectations." *Journal of Political Economy,* 86:1109–44.
(1981*a*). "Interpreting Economic Time Series." *Journal of Political Economy,* April, pp. 213–48.
(1981*b*). "Stopping Moderate Inflations: The Method of Poincare and Thatcher." Unpublished manuscript, University of Minnesota and Federal Reserve Bank of Minneapolis.
Townsend, R. M. (1978). "Market Anticipations, Rational Expectations, and Bayesian Analysis." *International Economic Review,* 19:481–94.
(1982). "Forecasting the Forecasting of Others." Unpublished manuscript, Carnegie-Mellon University, (to be published in *Journal of Political Economy,* 83).

Comment

JOHN B. TAYLOR

Implicit in almost all practical applications of the rational expectations method are two strong assumptions. First, it is assumed that *people know the model* of the economy used in the analysis and that they form expectations using this model. Second, it is assumed that *people know that all other people know the model* and form expectations in the same way. These two assumptions seem to restrict the range of applications

of rational expectations methods. They suggest that the methods are most realistic in situations where economic events are recurrent – such as business cycles – and where policy rules are in operation for a long time. As with most hypotheses used in economic analysis, however, these assumptions should be judged not only by their apparent realism but also by how successful they are in describing and forecasting economic behavior and by how they compare with alternative assumptions. As yet, there have been few attractive alternatives available.

In this elegant and constructive chapter, Robert Townsend proposes alternative, less restrictive assumptions that have the potential of broadening the range of economic problems to which rational expectations analysis can be applied. Moreover, he develops a methodology through which tractable results can be obtained using these alternative assumptions and shows how the methods work in some representative economic applications. In my view, these alternatives deserve careful consideration by those using rational expectations in situations where the more restrictive assumptions seem inappropriate, and, as Townsend suggests, they ought to be tried out in some practical economic policy problems. In these comments I shall discuss how the Townsend assumptions represent a generalization of existing expectational assumptions and consider the types of applications where some experimentation with the methods might be useful.

Rather than assuming that economic agents know the parameters of the model, Townsend assumes that some of these parameters are unobservable and evolve over time. For example, firms are assumed to be unaware of the intercept (θ in the chapter's notation) of the demand curve that they face. Instead, they know that this intercept moves according to the probability law

$$\theta_t = \rho\theta_{t-1} + v_t$$

and can only be observed with error $u_t = \theta_t + w_t$ [see equations (23) and (24) of the Townsend chapter]. The firms use this information structure to forecast future values of the intercept and thereby form expectations of future prices and make production decisions.

It is not difficult to imagine applications where this assumption might be more appropriate than assuming θ was known. In a commodity demand equation, the parameter θ could represent tastes that change gradually and that can be estimated with error through survey methods. In a money demand function, such an assumption could represent technological change in transactions technology that can be tracked only up to some measurement error. In a fiscal or monetary policy reaction function, such an assumption could be used to capture gradually shifting economic policies that are never fully announced or believed. In this case, the probability law would represent how policy was evolving through time, and u_t could be a current policy announcement that is only imperfectly correlated with actual policy. Note that in each of these three examples Townsend's assumptions require that agents know the model that underlies these shifts: a model of taste change, a model of technological change in financial markets, or a (political?) model of policy change.

In a number of situations the Townsend assumptions might not be appropriate as an alternative to the "agents-know-the-model" assumption. For example, an important modeling task is to describe economic behavior during a transition from one policy regime to another.[1] After a change in policy regime it would be inaccurate to assume that economic agents immediately understand the new policy. Instead, they might learn about the policy gradually as they observe policy decisions over a period of time. More generally, a structural parameter of the model might change, and agents would have to learn about this change through observation. In terms of Townsend's notation, these types of problems could be represented in terms of the parameter ρ of the autoregressive process. If θ were the money supply growth rate, then a switch to $\rho = 0$ could represent a fixed monetary growth rate. People would learn about ρ only as they observed actual money growth rates. Because in Townsend's models people are assumed to know the process generating θ, this type of problem cannot be handled.

Learning about the parameters of the model in this latter sense has proved to be a quite difficult phenomenon to model adequately.[2] There are three reasons for this difficulty. First, because agents must make decisions based on estimates of parameters, their actions cannot be considered exogenous to parameter estimation. The actions form the data on which the estimates of parameters are made. Because most conventional econometric procedures require that the data be exogenous, or endogenous in particularly restrictive ways, these market interactions with data generation require different techniques for analysis. Second, there is a possibility that, as agents gradually learn about the parameters, their

[1] See, for example, J. B. Taylor (1975). "Monetary Policy during a Transition to Rational Expectations." *Journal of Political Economy,* October, pp. 1009–21, and L. H. Meyer and C. Webster, Jr. (1982). "Monetary Policy and Rational Expectations: A Comparison of Least Squares and Bayesian Learning." In: *Carnegie-Rochester Series in Public Policy,* edited by K. Brunner and A. Meltzer, Amsterdam: North Holland.

[2] Chapter 6 by Margaret Bray in this book considers such a problem in a simple one-parameter learning situation.

actions will converge to some constant value that does not generate enough new information about the parameters. In the demand-curve example, a firm might begin selling the same quantity each period based on its estimate of the expected price; this prevents quantity from varying enough to get reliable estimates of the demand curve.[3] In some instances, estimates are inconsistent, but few results are yet available. In any case, the analysis necessarily becomes quite complicated, even without the market interactions previously mentioned. The problem is much worse in a multiparameter situation, and this is one reason why many studies have focused on one-parameter examples.[4] Third, the possibility that agents might affect how much information they can obtain about the parameters changes the nature of the optimal control problems in fundamental ways. A simple example is that of a firm experimenting with its prices, temporarily deviating from its best guess of the optimal price, in order to obtain information to be used in the future. Even in one-parameter partial equilibrium problems, this "dual-control" or "joint estimation and control" problem leads to significant complications. Solutions that may have been linear in a model where the parameter was known do not even have a closed form when the parameter is unknown.[5]

Because of these computational difficulties with existing approaches to modeling learning, Townsend's approach, although assuming that the laws governing parameter movement are known, may be a satisfactory alternative. For some applications, the distinction made here between knowing parameter values (θ) and knowing the probability law generating the parameter values (ρ) may be sufficiently fine that Townsend's more tractable approach could be used.

Thus far, I have discussed situations, as in Section 9.5 of the chapter, where expectations are assumed to be homogeneous. In Sections 9.6 and 9.7, Townsend considers ways to avoid this assumption and allow for heterogeneous, or disparate, expectations. With disparate expectations,

[3] This argument can be made more general with exogenous shifts in the functions. See T. W. Anderson and J. B. Taylor (1976). "Some Experimental Results on the Statistical Properties of Least Squares Estimates in Control Problems." *Econometrica,* November, pp. 1289–302, for a discussion of the problem and a demonstration of how this problem leads to a situation of extreme "multicollinearity" and poor parameter estimates.

[4] A proof of convergence and asymptotic normality is given for a partial equilibrium one-parameter example by J. B. Taylor (1974). "Asymptotic Properties of Multiperiod Control Rules in the Linear Regression Model." *International Economic Review,* June, pp. 472–84.

[5] See E. C. Prescott (1972). "The Multiperiod Control Problem under Uncertainty." *Econometrica,* November, pp. 1043–57.

an infinite regress problem arises in which agents must not only forecast but also forecast the forecasts of others, and so on. Townsend deals with this infinite regress problem head-on, by augmenting the state variables to include forecast of forecasts – the second-order expectation, as well as third- and higher-order expectations. There is a modeling choice, however, about where to truncate the infinite regress, or whether to truncate it at all. An element of judgment is required here, but perhaps the decision could be made empirically. As Townsend has shown in an earlier study, the regress problem has implications for the serial correlation properties of the errors in statistically estimated decision rules.[6] It would be interesting to examine whether, for example, the second- or third-order expectation truncation fits the data better than the first-order truncation that is conventionally used. But, in general, because it is impossible to know which truncation to assume, this may leave an element of arbitrariness in situations where there are other reasons for serial correlation. Clearly, some empirical work is necessary before we can say whether or not Townsend's higher-order expectations model is an improvement over the first-order methods now in use.

[6] See R. M. Townsend (1982). "Forecasting the Forecasts of Others." Unpublished manuscript, Carnegie-Mellon University (to be published in *Journal of Political Economy,* 83).

CHAPTER 10

Keynesianism, monetarism, and rational expectations: some reflections and conjectures

AXEL LEIJONHUFVUD

To what extent is Keynesianism discredited? Is there anything left? Did Monetarism score a total victory? Must Rational Expectations make New Classical economists of us all? Every teacher of macroeconomics has to wrestle with these questions – hoping against hope that some new cataclysm will not let some fantastic supply-side doctrine or whatever sweep the field before he has been able to sort through the rubble of what once he knew. I am going to sort some of my rubble. The object of the exercise is to make some guesses at how the seemingly still usable pieces might fit together.

My starting points are as follows. Keynesianism foundered on the Phillips curve or, more generally, on the failure to incorporate inflation rate expectations in the model. The inflation, which revealed this critical fault for all to see, was in considerable measure the product of "playing the Phillips curve" policies. But the stable Phillips trade-off was not an integral part of Keynesian theory.[1] Its removal, therefore, should not be (rationally) expected to demolish the whole structure.

Monetarism made enormous headway in the economics profession and with the public when the misbehavior of the Phillips curve and the inflation premium in nominal interest rates became obvious for all to see. And few observers could continue to doubt the strong link between nominal income and money stock as the Great American Inflation went on and on and on. The monetarist "victory"[2] was impressive enough

I have profited from discussions with Carlos-Daniel Heymann and from comments on earlier drafts by Earlene Craver-Leijonhufvud.

[1] I know, of course, that to some people Keynesianism means little else than Phillips-curve stability, but I ask indulgence in using my own definition of the term.

[2] "Victory" and "defeat" are terms that belong, perhaps, on the sports pages rather than to the history of science. Here, however, no epistemological meaning but only sociology of knowledge connotations are intended: To "win" means to attract the best new, young talent. In this sense, the monetarism of