Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Economic Theory 127 (2006) 264–295
JOURNAL OF Economic Theory
www.elsevier.com/locate/jet

Intrinsic heterogeneity in expectation formation

William A. Branch[a,*], George W. Evans[b]

[a] *Department of Economics, University of California, 3151 Social Science Plaza, Irvine, CA 92697-5100, USA*
[b] *Department of Economics, 1285 University of Oregon, Eugene, OR 97403-1285, USA*

Received 30 May 2003; final version received 23 November 2004
Available online 4 February 2005

Abstract

We introduce the concept of *Misspecification Equilibrium* to dynamic macroeconomics. Agents choose between a list of misspecified econometric models and base their selection on relative forecast performance. A Misspecification Equilibrium is a stochastic process in which agents forecast optimally given their choices, with forecast model parameters and predictor proportions endogenously determined. Under appropriate conditions, the Misspecification Equilibrium will exhibit *Intrinsic Heterogeneity*, in which all predictors are used at all times, even in the neoclassical limit in which only the most successful predictors are used. This equilibrium is attainable under least-squares learning and dynamic predictor selection based on average profits.

JEL classification: C62; D83; D84; E30

Keywords: Cobweb model; Heterogeneous beliefs; Adaptive learning; Rational expectations

1. Introduction

Despite its dominance in dynamic macroeconomic models, the Rational Expectations Hypothesis has limitations. A frequently cited drawback to the rational expectations approach is that in effect it assumes that agents know the underlying economic structure. In response to this criticism one popular alternative is to model agents as econometricians (Evans and Honkapohja [11]). This adaptive learning approach typically assumes agents have a correctly specified model with unknown parameters. Agents then use a reasonable

* Corresponding author.
E-mail addresses: wbranch@uci.edu (W.A. Branch), gevans@uoregon.edu (G.W. Evans).

0022-0531/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.jet.2004.11.005

estimator to obtain their coefficient estimates. In many models these beliefs converge to rational expectations.

In practice, however, econometricians often misspecify their models. Economic forecasters who use VARs purposely limit the number of variables and the number of lags because of degree of freedom problems. If agents are expected to behave like econometricians then they can also be expected to misspecify their models. Evans and Honkapohja [11, Chapter 13] consider models with agents underparameterizing the law of motion, and show the existence of a Restricted Perceptions Equilibrium (RPE) in which agents form their beliefs optimally given their misspecification. [1] The issue of underparameterization is also emphasized by Evans and Ramey [14], who examine the implications of optimally chosen expectations within the simple adaptive expectations class.

In this paper, we examine expectation formation in an environment where agents must forecast using an underparameterized econometric model. More specifically we confront agents with a list of misspecified econometric models, but, given this restriction, assume that agents forecast optimally. Agents choose between these optimal underparameterized models based on their relative mean success.

We investigate this approach in a linear stochastic framework, developing the analysis in the context of the cobweb model. Because the economic model is self-referential, in the sense that expectation formation affects the law of motion for the endogenous variables, the optimal parameters of each misspecified econometric model depend on the proportions of agents using the different models. We define a new equilibrium concept, called a *Misspecification Equilibrium*, in which these proportions are consistent with optimal forecasting from each econometric model. We show that for some economic model parameters and exogenous driving variables, agents will be distributed heterogeneously between the various predictors, even as we approach the limiting case in which agents choose only between the best performing statistical models. We say that a Misspecification Equilibrium with such a property exhibits *Intrinsic Heterogeneity*.

Heterogeneity in expectations has been considered previously in papers by Townsend [23], who takes a fully rational learning approach, starting with given priors, and Haltiwanger and Waldman [16] who assume that a certain fraction of agents are not rational. In adaptive learning models Honkapohja and Mitra [17] allow agents to have different specific learning rules. The seminal least-squares learning paper by Bray and Savin [4] also allows for heterogeneity in priors. However, these papers all assume an ad hoc degree of heterogeneity, and, with least squares or Bayesian learning, the heterogeneity disappears in the limit. Evans et al. [12] allow for stochastic heterogeneity in learning rules, but again the heterogeneous expectations is only transitory.

Brock and Hommes [7] were among the first to model heterogeneous expectations as an endogenous outcome. [2] Brock and Hommes [7] examine a cobweb model in which agents choose a predictor from a set of costly alternatives. Agents base this choice on the most

[1] Sargent [20] developed the implications of policy makers estimating and forecasting using a misspecified model.

[2] Sethi and Franke [21] also find heterogeneity, as an outcome of evolution in a model of stochastic strategic complementarities, and Evans and Ramey [13] permit heterogeneous expectations due to heterogeneous calculation costs.

recent realized profits of the alternatives in a cobweb model. If agents are boundedly rational in the sense that their 'intensity of choice' between predictors is finite (that is, they do not fully optimize), then there will be heterogeneity and the degree of heterogeneity will vary in a complex manner.

Brock and Hommes illustrate these results in a particular case of rational versus myopic beliefs. Because agents always react to recent changes in profits their predictor choice will oscillate along with the equilibrium price. Our model is closely related to Brock and Hommes. Like their model, we assume that the map from predictor benefits to predictor choice resembles a multinomial logit. The multinomial logit has proven to be an important approach to modeling economic choices, [3] and has been increasingly employed in recent work in dynamic macroeconomics. Extensions of the Brock and Hommes [7] predictor selection dynamic appear in [1,2,6,8–10]. Brock and Durlauf [5] extend the framework so that agent specific choices depend on the expected choices of others.

There are three important departures in our model. First, agents do not choose between a costly accurate forecast and a costless unsophisticated forecast; rather, they are forced to choose between equally misspecified costless models. Second, in line with the econometric learning literature, each forecasting model depends on parameters, which are chosen to minimize the mean square forecast error. In equilibrium, each forecasting model is optimal, given the misspecification. Third, we assume that agents make their choices based on unconditional mean payoffs rather than on the most recent period's realized payoff. This is more appropriate in our stochastic environment since otherwise agents would frequently be misled by single period anomalies. We will show that even if agents optimally choose between these misspecified models heterogeneity can arise. Given that agents base decisions on mean profits it is not at all obvious that heterogeneity would be possible if the 'intensity of choice' is large. Indeed, we will show that instances of asymptotically homogeneous expectations also arise.

The main difference in our results is that, unlike previous work, we derive heterogeneity as a possible equilibrium outcome of a self-referential model in which agents are constrained to underparameterize. In particular, we examine the case in which agents are fully rational except that they misspecify by omitting at least one relevant variable or lag. We focus on the cobweb model for two reasons. First, we want to stay close to Brock and Hommes [7] in order to highlight the key differences. Second, the cobweb model is the simplest self-referential model that effectively illustrates the intuition of Intrinsic Heterogeneity.

We obtain conditions under which there is an equilibrium with agents heterogeneously split between the misspecified models even as the 'intensity of choice' becomes arbitrarily large. The intuition for this possibility is as follows. Suppose the cobweb price is driven by a two-dimensional vector of demand shocks. If both components of the demand shock matter for predicting prices, and if the feedback through expectations is sufficiently large, then there will be an incentive to deviate from homogeneity. If all agents coordinate on the same model the negative feedback through expectations will make the consensus model less useful for forecasting. In these instances an agent could profit by forecasting with the alternative model. With Intrinsic Heterogeneity the equilibrium is such that beliefs and predictor proportions drive expected profits to be identical.

[3] See, for example, Manski and McFadden [19].

A related issue discussed in the literature is whether speculative market forces are stabilizing or destabilizing. The conjecture that speculation is generally stabilizing was first put forth by Friedman [15] in an argument favoring a system of flexible exchange rates. This issue was considered at length by Brock and Hommes [8], who show that in some cases there could be cycles or complex dynamics around the steady state with agents shifting between forecast rules. Brock and Hommes [8] examine the "Friedman hypothesis" in the context of an asset pricing model, but the points they raise are also relevant in our setup.

To explore this issue we study the stability of a Misspecification Equilibrium under learning. In a Misspecification Equilibrium agents use only the best performing statistical models and are, thus, fully 'smart' given their underparameterization, and they choose to rely on the forecast rules with highest mean profits. If the equilibrium arises through a process of econometric learning and dynamic predictor selection, then the Friedman hypothesis holds in the sense that the equilibrium is dynamically stable. We show that if agents, in real time, estimate model parameters by least squares and mean predictor profits by sample averages, then the Misspecification Equilibrium will be attained. However, dynamic stability depends critically on the method for estimating mean profits. If agents place a high weight on recent realized profits, as in Brock and Hommes [7,8], then the system can generate dynamics around the equilibrium, with rapid switching between forecast rules.

The plan for this paper is as follows. Section 2 introduces the setup in a general cobweb model. We obtain an existence result for Misspecification Equilibria, and give conditions under which the model exhibits Intrinsic Heterogeneity. Section 3 extends and illustrates these results for the special case of a process driven by a two-dimensional VAR(1) shock with agents choosing between two underparameterized models. Section 4 studies the circumstances in which a Misspecification Equilibrium can be attained in real time under econometric learning and dynamic predictor selection. Section 5 concludes and describes future work.

2. Model

In this section we consider a self-referential stochastic process that is driven by vector autoregressive exogenous shocks. We assume that agents' expectations are based on one of a set of misspecified models of the economy, each taking the form of an underparameterization of the process. In the terminology of Brock and Hommes [7] we are in effect treating forecasts based on a fully correctly specified model as prohibitively costly, and those based on the misspecified models are equally and much less costly. (For convenience we will normalize this cost to zero.) Much previous work has assumed a particular structure of agents' misspecification. We allow the choice of the misspecified model to be endogenous.

We develop our model as a version of the Adaptively Rational Equilibrium Dynamics (ARED) of Brock and Hommes [7] in which we constrain agents to choose between underparameterized models. Agents consider the unconditional expected payoff of the various possible underparameterizations and select between them according to their relative payoffs. Using the selected model they form their expectations as the optimal linear projection given this choice. In our Misspecification Equilibrium, the projection parameters and predictor proportions are jointly determined and generate the equilibrium stochastic process.

We think that our emphasis on underparameterization is reasonable. The adaptive learning literature has argued in favor of modeling agents as econometricians as a plausible deviation from the rational expectations assumption. However, econometricians misspecify their econometric models. Computational time and limits on degrees of freedom make it impossible for an econometrician to include all economically relevant variables and lags. Our model in effect imposes such restrictions on agents, but otherwise requires them to behave optimally. A striking finding of our framework is that this can lead to the use of heterogeneous forecasting models.

We develop the model in stages. We first show that, for given predictor proportions, there exists a Restricted Perceptions Equilibrium (RPE) in which agents' misspecified beliefs are verified by the actual equilibrium process. We next allow for predictor proportions to be endogenously determined, and show the existence of a Misspecification Equilibrium. Finally, we formally define Intrinsic Heterogeneity and state a condition under which this will arise.

2.1. Setup

We consider a cobweb model of the form

$$p_t = -\phi p_t^{\mathrm{e}} + \gamma' z_t + v_t, \tag{1}$$

where v_t is white noise. Although there are several well-known economic models that fit form (1), we focus on the "cobweb" model in order to keep a close connection between our model and Brock and Hommes [7]. z_t is a vector of observable demand disturbances, which will be further specified below.

We normally expect $\phi > 0$ in the cobweb model, which corresponds to upward sloping supply curves and downward sloping demand curves. Bray and Savin [4] showed that $\phi > -1$ was the condition for the model to be stable under least-squares learning. In this paper we focus on the negative feedback case of $\phi > 0$ and leave $\phi < 0$ for future work. [4]

In the cobweb model firms have a one-period production lag. We assume that firms have quadratic costs given by $FQ_t^* + \frac{1}{2}G(Q_t^*)^2$, where Q_t^* is planned output and $F \geqslant 0$, $G > 0$. In addition we allow for exogenous productivity shocks realized after production decisions are made so that total quantity is $Q_t = Q_t^* + \kappa_t$. Here κ_t is iid with zero mean. Firms aim to maximize expected profits. [5] Thus their problem is

$$\begin{aligned}\max_{Q_t^*} E_{t-1}\pi_t &= E_{t-1}\left[p_t\left(Q_t^* + \kappa_t\right) - FQ_t^* - \frac{1}{2}G(Q_t^*)^2\right] \\ &= Q_t^* E_{t-1} p_t + E_{t-1}(p_t\kappa_t) - FQ_t^* - \frac{1}{2}G(Q_t^*)^2.\end{aligned}$$

[4] Eq. (1) with $-1 < \phi < 0$ takes the same form as a Lucas-type monetary model. In future work we will pursue the possibility of heterogeneity in that model.

[5] It would be possible to extend the model to incorporate risk by assuming agents respond to variances of profits as well as expected profits. We make the expected profits assumption to keep the model as simple as possible.

Solving this problem leads to the supply relation [6]

$$Q_t^* = G^{-1} p_t^e, \tag{2}$$

where $p_t^e = E_{t-1} p_t$. Firms differ by forecasting model. If $p_{j,t}^e$ denotes the expectation of a firm of type j, then actual supply follows $Q_t = G^{-1} \sum_j n_j p_{j,t}^e + \kappa_t$, where n_j is the proportion of firms of type j and the total number of firms has been normalized to one. Here we are treating κ_t as a market-wide shock.

Demand is given by

$$Q_t = C - Dp_t + h'\zeta_t, \tag{3}$$

where ζ_t is an $m \times 1$ vector of demand shocks that follows a zero-mean stationary VAR(n) process and $D > 0$. The ζ_t process is assumed independent of κ_t. Setting demand equal to actual supply we have the following stochastic equilibrium price process:

$$p_t = -(DG)^{-1} \sum_j n_j p_{j,t}^e + D^{-1} h'\zeta_t - D^{-1}\kappa_t, \tag{4}$$

where, for convenience, we have expressed p_t and $p_{j,t}^e$ in deviation from the mean form.

It is convenient to rewrite the model in terms of an exogenous VAR(1) process. Defining

$$z_t' = (\zeta_t', \zeta_{t-1}', \ldots, \zeta_{t-n+1}'),$$

we can write z_t in its standard VAR(1) form

$$z_t = Az_{t-1} + \varepsilon_t \tag{5}$$

for appropriately defined A and appropriately defined ε_t, which is exogenous white noise. Here z_t is $mn \times 1$ and A is $mn \times mn$. We denote the covariance matrix of z_t as $\Omega = Ezz'$, and Ω is assumed to be positive definite. Setting

$$\phi = (DG)^{-1}, \gamma' = (D^{-1}h', 0, \ldots, 0), p_t^e = \sum_j n_j p_{j,t}^e, \quad \text{and} \quad v_t = -D^{-1}\kappa_t$$

we can rewrite (4) in form (1). [7]

2.2. Model misspecification

To close the model we need to specify the determination of p_t^e. We assume that there are K econometric models available to form expectations and that model $j = 1, \ldots, K$ uses

[6] We have set, without loss of generality, $F = 0$. We are also assuming that agents treat $E_{t-1}(p_t \kappa_t)$ as a constant independent of the choice of Q_t^*. That this is a reasonable assumption can be verified by (4) below.

[7] Note that the dependence of p_t on the weighted average market expectation p_t^e arises from the dependence of p_t on aggregate supply and from the linear supply relation.

$k_j < mn$ explanatory variables. The "market expectation" is given by the weighted sum of the individual expectations

$$p_t^e = \sum_{j=1}^{K} n_j p_{j,t}^e, \tag{6}$$

where $p_{j,t}^e = b^{j\prime} x_{t-1}^j$, $x_t^j = u^j z_t$. The $k_j \times m$ matrix u^j is a selector matrix that picks out those elements of z_t used in predictor j and b^j is $k_j \times 1$. Thus, k_j is the number of elements in z_t that predictor j uses. We can rewrite (6) as

$$p_t^e = \sum_{j=1}^{K} n_j b^{j\prime} u^j z_{t-1}.$$

This setup forces agents to underparameterize the variables included in their information set and/or the number of lags of those variables. We believe this is a reasonable approximation of actual expectation formation. Cognitive and computing time constraints (as well as degrees of freedom) restrict the number of variables even the most diligent econometricians use in their models. Our form of misspecification makes agents be (at least somewhat) parsimonious in their expectation formation.

We next specify the determination of the parameters b^j. In a fully specified econometric model, and under rational expectations, all variables z_t would be included and the coefficients used to form p_t^e would be given by the least squares projection of p_t on z_t. Here each predictor is constrained to use a subset x_t^j of relevant variables, and thus each predictor differs from rational expectations. However, we will insist that the beliefs b^j are formed optimally in the sense that b^j is the least squares projection of p_t on $u^j z_{t-1}$. That is, b^j must satisfy

$$Eu^j z_{t-1}\left(p_t - b^{j\prime} u^j z_{t-1}\right) = 0.$$

Even though agents will never be "fully" accurate, they will be as accurate as possible given the variables in their information set.

2.3. Restricted perceptions equilibrium

Given the belief process (6), the VAR in (5), and the equilibrium price (1), the actual law of motion (ALM) for this economy is

$$p_t = \left[\gamma' A - \phi\left(\sum_{j=1}^{K} n_j b^{j\prime} u^j\right)\right] z_{t-1} + \gamma' \varepsilon_t + v_t$$

or

$$p_t = \xi' z_{t-1} + \gamma' \varepsilon_t + v_t, \tag{7}$$

where

$$\xi' = \gamma' A - \phi \left(\sum_{j=1}^{K} n_j b^{j\prime} u^j \right). \tag{8}$$

For convenience we write $n = [n_1, \ldots, n_K]'$ for the vector of predictor proportions and $b = [b^1, \ldots, b^K]$ for the matrix giving the forecast coefficients. Given Eqs. (7)–(8) and the parameter orthogonality condition we obtain

$$b^j = \left(u^j \Omega u^{j'} \right)^{-1} u^j \Omega \xi. \tag{9}$$

We now introduce the concept of RPE. [8] An RPE is an equilibrium process for p_t such that the parameters b^j are optimal given the misspecification. Note that, like a rational expectations equilibrium, an RPE is self-referential in that the optimal beliefs b depend on the vector of parameters ξ which depend in turn on the vector of beliefs b. Thus, an RPE can be defined as a process (7) such that ξ is a solution to (8) and (9) for fixed n.

Substituting (9) into (8) yields

$$\xi' = \gamma' A - \phi \sum_{j=1}^{K} n_j \xi' \Omega u^{j\prime} \left(u^j \Omega u^{j\prime} \right)^{-1} u^j$$

or

$$\xi = \left[I + \phi \sum_{j=1}^{K} n_j u^{j\prime} \left(u^j \Omega u^{j\prime} \right)^{-1} u^j \Omega \right]^{-1} A' \gamma. \tag{10}$$

For a given n an RPE exists (and is unique), provided the inverse in (10) exists.

In the Misspecification Equilibrium, which we define below, n is determined endogenously. Eq. (10) gives a well-defined mapping $\xi = \xi(n)$ provided the indicated inverse exists for all n in the unit simplex. We therefore assume that the following condition holds:

Condition Δ. $\Delta \neq 0$ for all n in the unit simplex $S = \{n \in \mathbb{R}^K : n_i \geqslant 0 \text{ and } \sum_{i=1}^{K} n_i = 1\}$, where

$$\Delta = \det \left(I + \phi \sum_{j=1}^{K} n_j u^{j\prime} \left(u^j \Omega u^{j\prime} \right)^{-1} u^j \Omega \right).$$

Condition Δ is a necessary and sufficient condition for the existence of a unique RPE for all $n \in S$.

[8] See Evans and Honkapohja [11] for a definition and examples. The concept introduced here extends the concept of RPE to incorporate multiple misspecified models.

We have the following result:

Proposition 1. *For $\phi \geqslant 0$ sufficiently small, Condition Δ is satisfied and hence for all n there exists a unique RPE.*

All proofs are contained in the Appendix. In the next section we demonstrate that Condition Δ holds for all $\phi \geqslant 0$ in the case of a bivariate process.

2.4. Misspecification equilibrium

We now embed the RPE into an equilibrium concept in which n is endogenously determined by the mean profits of each predictor. We will call this a *Misspecification Equilibrium*. Note that the profits of each predictor depend on the parameters ξ which in turn depend on n.

In order to discuss the mapping for predictor proportions we need the profits for predictor j, which are given by

$$\begin{aligned}\pi_t^j &= p_t\left(\phi D p_{j,t}^{\mathrm{e}} - Dv_t\right) - \frac{1}{2}\phi D\left(p_{j,t}^{\mathrm{e}}\right)^2 \\ &= \left[\xi(n)' z_{t-1} + \gamma'\varepsilon_t + v_t\right]\left[\phi D b^{j\prime} u^j z_{t-1} - Dv_t\right] - \frac{1}{2}\phi D\left(b^{j\prime} u^j z_{t-1}\right)^2,\end{aligned}$$

where, again, we have expressed profits in deviation from mean form. Taking unconditional expectations of profits yields

$$E\pi_t^j = \phi D b^{j\prime} u^j \Omega\left(\xi(n) - \frac{1}{2}u^{j\prime} b^j\right) - DEv_t^2.$$

Evaluating expected profits in an RPE (i.e. plugging in (9)) leads to

$$\begin{aligned}E\pi^j = {} & \phi D\xi(n)'\Omega u^{j\prime}(u^j\Omega u^{j\prime})^{-1}u^j\Omega\left(\xi(n) - \frac{1}{2}u^{j\prime}(u^j\Omega u^{j\prime})^{-1}u^j\Omega\xi(n)\right) \\ & - DEv_t^2. \qquad (11)\end{aligned}$$

Note that $E\pi^j$ is well-defined and finite for all n, provided Condition Δ holds so that $\xi(n)$ is well-defined. It will be convenient to denote the function given by (11) as

$$\tilde{F}_j(n) : S \to \mathbb{R} \quad \text{for } j = 1, \ldots, K$$

and to define $\tilde{F}(n) : S \to \mathbb{R}^K$ by $\tilde{F}(n) = (\tilde{F}_1(n), \ldots, \tilde{F}_K(n))'$. Note that $\tilde{F}_j(n)$ and $\tilde{F}(n)$ are continuous on S provided Condition Δ holds.

We now follow Brock and Hommes [7] in assuming that the predictor proportions follow a multinomial logit (MNL) law of motion. Brock and Hommes consider the cobweb model without noise where agents choose between rational and naive expectations. Agents adapt their choices based on the most recent relative predictor success. [9] This clearly would not be appropriate in the stochastic framework employed here, and we instead assume that

[9] Branch [1] shows that many of the qualitative properties in the model without noise carry over to a model with small demand disturbances.

agents base their decision on unconditional expected relative payoffs (however, see below for further discussion).

The MNL approach leads to the following mapping, for each predictor i,

$$n_i = \frac{\exp\{\alpha E\pi^i\}}{\sum_{j=1}^{K} \exp\{\alpha E\pi^j\}}, \tag{12}$$

where $\alpha > 0$. Note that $n_i > 0$ for α and the $E\pi^j$ finite and that $\sum_j n_j = 1$. Again, it will be convenient to denote the map defined by (12) as

$$\tilde{H}_\alpha(E\pi^1, \ldots, E\pi^K) : \mathbb{R}^K \to S$$

and clearly $\tilde{H}_\alpha$ is continuous. The parameter α is called the 'intensity of choice,' and parameterizes one dimension of agents' bounded rationality. As $\alpha \to +\infty$ we obtain the 'neoclassical' case of full optimization. We will be interested in the conditions in which heterogeneity can arise in the neoclassical case.

We remark that our choice of payoff function $E\pi^j$ will allow us to consider the fixed point of a map rather than the solution to a difference equation as in Brock and Hommes [7,8]. Another possible measure of predictor fitness is given by $\hat{\pi}_t^j = \delta \sum_{k=0}^{\infty} (1-\delta)^k \pi_{t-k}^j$. Since this measure depends on t it implies that the predictor proportions given by the MNL map also depend on t. In this setting $n_{i,t}$ and ξ_t would follow a stochastic process and the Misspecification Equilibrium concept would be defined in terms of a Markov process for (n_t, ξ_t) with an invariant distribution. Investigating this framework analytically is beyond the scope of the current paper, [10] but we do consider this numerically in Section 4. We now show that using the payoff functions $E\pi^j$ we can develop an equilibrium concept, in which predictor choices are consistent with the expected profits they determine, and in which each predictor parameter vector is optimal, given the predictor proportions.

Define the mapping

$$\tilde{T}_\alpha : S \to S, \quad \text{where } \tilde{T}_\alpha = \tilde{H}_\alpha \circ \tilde{F}.$$

Under Condition Δ this map is well-defined and continuous. $\tilde{T}_\alpha$ maps a vector of predictor choices, n, through the belief parameter mapping ξ into a vector of expected profits and then to a new predictor choice n. We are now in a position to present our central equilibrium concept:

Definition. A Misspecification Equilibrium (ME) is a fixed point, n^*, of $\tilde{T}_\alpha$.

Applying the Brouwer Fixed Point Theorem we immediately have:

Theorem 2. *Assume Condition Δ. There exists a ME.*

[10] The usual method for proving existence of such an equilibrium is illustrated by Branch and McGough [3], employing the theorems of Stokey and Lucas [22]. Such a result appears unavailable in the current setting because these theorems rely on monotone Markov processes.

In general we cannot rule out multiple equilibria. Let

$$N_\alpha = \{n^* | \tilde{T}_\alpha(n^*) = n^*\}.$$

For α finite and $E\pi^j$ finite, it is apparent that all components are positive for every fixed point n^*. Thus, heterogeneity for finite α is simply a by-product of the MNL assumption, which ensures that all predictors are used even if they differ in terms of their performance. However, it is of interest to know if heterogeneity continues to arise if agents are highly sensitive to relative performance, so that they only use predictors that are not dominated in performance. This leads to the following concept:

Definition. A model is said to exhibit Intrinsic Heterogeneity if (i) an ME exists for all $\alpha > 0$ and (ii) there exists $\bar{n} < 1$ such that $n_j^* \leqslant \bar{n}$, $j = 1, \ldots, K$, for all α and all ME $n^* \in N_\alpha$.

It can be shown that a model with intrinsic heterogeneity arises whenever the following additional condition is satisfied. [11]

Condition P. Let e_i denote the $K \times 1$ coordinate vector with 1 in position i and 0 elsewhere. Condition P is said to be satisfied if for each $i = 1, \ldots, K$ there exists $j \neq i$ such that $\tilde{F}_j(e_i) - \tilde{F}_i(e_i) > 0$.

Theorem 3. *Assume Condition Δ and also Condition* P. *Then the model exhibits intrinsic heterogeneity.*

Below we develop the details and intuition behind this result for the special case of a bivariate VAR process. A brief discussion, though, is warranted at this point. Condition P implies that there is no predictor that dominates when all agents are massed onto that predictor. Thus, under Condition P there is always an incentive for agents to deviate from homogeneity. Essentially, Condition P arises as a joint condition on the self-referential property of the model (1), parameterized by ϕ, and the asymptotic moments of the stochastic process z_t. If each component of z matters in a particular way (see below) and the negative feedback of the cobweb model is sufficiently strong, then agents will always seek to deviate from a consensus selection. When Condition P is satisfied agents must be distributed, in an equilibrium, across all forecasting models.

This result is related to Brock and Hommes [7]. In Brock and Hommes [7], in a steady-state with costless predictors, agents will be equally divided between predictors. This arises, in their deterministic setting, because each predictor returns an identical forecast in a steady-state. Our fixed point delivers an equilibrium stochastic process that is the analogue of the steady-state in their model. In our model agents will not (usually) be distributed uniformly across predictors, in equilibrium, because even in equilibrium the predictors make different forecasts. As will be made more concrete below, it is the interaction between the direct

[11] We call this "Condition P" because it gives the key condition on the profit functions that generates intrinsic heterogeneity. In the bivariate case, below, Condition P0 (resp. P1) will be used for the condition that yields an equilibrium with no (resp. all) agents incorporating z_1.

effect of the z_t process and its indirect effect acting through the self-referential feature of the cobweb model that can produce an equilibrium with agents using all predictors. The proportion assigned to a given predictor depends on the particular parameterization of the model.

The next section will present a simple example to illustrate our concepts. In particular, we present cases in which Condition P holds and the model exhibits Intrinsic Heterogeneity.

3. Example: bivariate case

To illustrate the properties of a ME we will simplify the model by considering a special case in which detailed results can be obtained. In this section we assume that z_t is a two-dimensional stationary VAR(1) $z_t = Az_{t-1} + \varepsilon_t$, where A is 2×2, with eigenvalues inside the unit circle, and $E\varepsilon_t\varepsilon_t' = \Sigma_\varepsilon$ is positive definite. Each misspecified model will omit one explanatory variable and thus $K = 2$ and $k_j = 1$ for $j = 1, 2$. This is the simplest possible illustration of our framework, and we will see that it can generate cases with Intrinsic Heterogeneity.

With bivariate demand shocks the predictors are now

$$p^{\mathrm{e}}_{1,t} = b^1 u^1 z_{t-1} = b^1 z_{1,t-1},$$
$$p^{\mathrm{e}}_{2,t} = b^2 u^2 z_{t-1} = b^2 z_{2,t-1}.$$

RPE (7) thus takes the form

$$p_t = \xi_1 z_{1,t-1} + \xi_2 z_{2,t-1} + \eta_t, \tag{13}$$

$\eta_t = \gamma'\varepsilon_t + v_t$. ξ and b can be calculated from (10) and (9). In particular, using $u^j \Omega u^{j\prime} = Ez_j^2$, $u^1\Omega = (Ez_1^2, Ez_1z_2)$ and $u^2\Omega = (Ez_1z_2, Ez_2^2)$, we obtain

$$\sum\nolimits_{j=1}^{2} n_j u_j' (u^j \Omega u^{j\prime})^{-1} u^j \Omega = \begin{bmatrix} n_1 & n_1\rho \\ n_2\tilde{\rho} & n_2 \end{bmatrix},$$

where

$$\rho = \frac{Ez_{1t}z_{2t}}{Ez_{1t}^2}, \quad \tilde{\rho} = \frac{Ez_{1t}z_{2t}}{Ez_{2t}^2}.$$

Finally, in accordance with (10) we obtain

$$\begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} = \begin{bmatrix} 1 + n_1\phi & \phi n_1\rho \\ \phi n_2\tilde{\rho} & 1 + n_2\phi \end{bmatrix}^{-1} A'\gamma \tag{14}$$

and from (9) b is given by $b' = (\xi_1 + \rho\xi_2, \tilde{\rho}\xi_1 + \xi_2)$. We remark that Ez_tz_t' is entirely governed by A and Σ_ε.

From the general results of Proposition 1 and Theorem 2 we know that a ME exists for $\phi \geqslant 0$ sufficiently small. For the bivariate case existence can be shown for all $\phi \geqslant 0$. Furthermore, we will show that this equilibrium is unique.

3.1. Misspecification equilibrium

If condition Δ is satisfied then this guarantees a unique $\xi' = (\xi_1, \xi_2)$ for each $n' = (n_1, n_2)$, and a unique RPE. Since $n_2 = 1 - n_1$, in this section we define the key functions in terms of n_1 rather than n. Thus, in particular, if Condition Δ holds then (14) defines a continuous map $\xi = \xi(n_1)$.

Proposition 4. *In the bivariate model, Condition Δ is satisfied for all $\phi \geqslant 0$. Hence there exists a unique RPE for every $n_1 \in [0, 1]$.*

From Theorem 2 it follows that there exists a ME. By developing the details we can obtain additional results. By evaluating (11) in the bivariate case, it is straightforward to verify that the profit functions are given by

$$\begin{aligned} E\pi^1 &= \frac{1}{2}\phi D\left(\xi_1^2(n_1) - \xi_2^2(n_1)\rho^2\right) Ez_{1t}^2 \\ &\quad + \phi D\left(\xi_1(n_1) + \xi_2(n_1)\rho\right)\xi_2(n_1) Ez_{1t}z_{2t} - D\sigma_v^2, \\ E\pi^2 &= \frac{1}{2}\phi D\left(\xi_2^2(n_1) - \xi_1^2(n_1)\tilde{\rho}^2\right) Ez_{2t}^2 \\ &\quad + \phi D\left(\xi_2(n_1) + \xi_1(n_1)\tilde{\rho}\right)\xi_1(n_1) Ez_{1t}z_{2t} - D\sigma_v^2, \end{aligned}$$

where $\xi_j^2(n_1)$ denotes $(\xi_j(n_1))^2$ and we define

$$F(n_1) = E\pi^1 - E\pi^2.$$

The function $F(n_1)$ gives the difference in expected profits, between forecasting models 1 and 2, for a given distribution of agents. In order to prove existence of a unique ME, we need to show that the profit difference function $F(n_1)$ is monotonic.

Lemma 5. *In the bivariate model, the function $F(n_1)$ is monotonically decreasing for all $\phi \geqslant 0$.*

We remark that it is possible to instead have a positive slope for the profit difference function $F(n_1)$ when $\phi < 0$. In this case it will be possible to have multiple equilibria. Examples with $\phi < 0$ are the focus of future research.

The predictor proportion mapping (12) can be written

$$n_1 = \frac{1}{2}\tanh\left[\frac{\alpha}{2}\left(E\pi^1 - E\pi^2\right)\right] + \frac{1}{2} \equiv H_\alpha(E\pi^1 - E\pi^2),$$

where $H_\alpha : \mathbb{R} \to [0, 1]$ is a strictly increasing function. Note that we use F and H_α in place of $\tilde{F}$ and $\tilde{H}_\alpha$ to emphasize that in contrast to the previous section the domain of F and the range of H_α is now $[0, 1]$ instead of the unit simplex S. This will simplify some of the arguments below.

Because Condition Δ is satisfied for all $\phi \geqslant 0$, there exists a well defined mapping $T_\alpha = H_\alpha \circ F$, $T_\alpha : [0, 1] \to [0, 1]$, which is continuous. From Lemma 5, and since H_α is a strictly increasing function, it follows that T_α is a continuous, decreasing function for each α. It immediately follows that there is a unique fixed point, i.e., we have:

Proposition 6. *Suppose z_t is a bivariate VAR(1). If $\phi \geqslant 0$ the model has a unique ME.*

Proposition 6 demonstrates that there is a unique equilibrium in the belief parameters b and the proportion of agents using the two misspecified models. It does not tell us how agents are distributed between the predictors. We now show that it is possible for there to be intrinsic heterogeneity. Unlike Brock and Hommes [7] who obtain heterogeneity across forecast rules as an automatic implication of assuming that α is finite, we want to show that there exists cases of heterogeneity even in the limit as $\alpha \to \infty$.[12] Because our framework is stochastic, the forecasts as well as the forecast rules will be heterogeneous. We now take up this issue.

3.2. Intrinsic Heterogeneity

The previous section established uniqueness of the ME. We now discuss more specific properties of this equilibrium. We will see that there are three possible cases depending on ϕ, γ, A and Σ_ε. In the bivariate case we are considering, Condition P of Section 2.4 holds when $F(0) > 0$ and $F(1) < 0$. Since, by Lemma 5, F is monotonically decreasing, there are two other cases: $F(0) < 0$, which we will refer to as Condition P0, and $F(1) > 0$, which we will refer to as Condition P1. We show below that each of the three possible cases can arise for appropriate values of ϕ, γ, A and Σ_ε.

Under Condition P0, $F(1) < 0$ implies that predictor 2 has higher mean profits for all $0 \leqslant n_1 \leqslant 1$. Under Condition P1, predictor 1 always has higher mean profits. In these cases we anticipate homogeneous expectations as the 'intensity of choice' $\alpha \to \infty$. However, if Condition P obtains there is an incentive to deviate from the consensus selection. We have the following result.

Proposition 7. *Consider again the model with z_t a bivariate VAR(1) and $\phi \geqslant 0$. The unique Misspecification Equilibrium n_1^* has one of the following properties*:

1. *Condition* P *implies that as $\alpha \to \infty$, $n_1^* \to \hat{n}_1 \in (0, 1)$ where $F(\hat{n}_1) = 0$. That is, the model has Intrinsic Heterogeneity.*
2. *Condition* P0 *implies that as $\alpha \to \infty$, $n_1^* \to 0$.*
3. *Condition* P1 *implies that as $\alpha \to \infty$, $n_1^* \to 1$.*

Proposition 7 establishes the possibility of Intrinsic Heterogeneity. We discuss the intuition further below. This result is novel because, for high α, rationality of agents is bounded only through their model parameterizations. Agents fully optimize given their (misspecified) model of the economy. In Brock and Hommes' ARED heterogeneity arises because of calculation costs and, most importantly, because with finite α a proportion of agents do not optimize in the sense that they do not fully respond to profit differences.

It is important to stress both the close connection to Brock and Hommes [7] steady-state ARED and the way in which our results differ. Their model focuses on the case

[12] In Brock and Hommes [7], when $\alpha = +\infty$ heterogeneity only arises in a steady state with zero costs for rational expectations, and the rational and naive predictors return the same forecasts.

of a choice between a fully rational and a naive predictor. For $\alpha \to \infty$, heterogeneity across predictors arises in their model only within a (nonstochastic) steady-state and with a costless rational predictor. Although agents use different forecast rules, both predictors make identical forecasts in a steady state, and thus there is no heterogeneity in the forecasts made across agents. In our model, agents optimize given their misspecification, all predictors are equally "sophisticated" and costless, and Intrinsic Heterogeneity can arise as part of a stochastic equilibrium. The nature of the equilibrium forces each predictor to return the same mean profit as $\alpha \to \infty$, but at each instant in time there are heterogeneous forecasts. As we stress below, the self-referential feature of the model, combined with underparameterization, is essential for generating this heterogeneity.

Heterogeneous forecast rules also can arise in the Brock and Hommes [8] asset pricing model. In that paper they focus on various examples with multiple costless predictors, including a "fundamentals" predictor and other biased or trend-setter forecast rules. They provide nonstochastic examples, for all $\alpha > 0$ and also $\alpha = +\infty$, of steady states in which the population is evenly divided between all predictors. In these examples there are heterogeneous forecasts across predictors. However, as they emphasize, the steady state is dynamically unstable when α is sufficiently large and hence there is no "long-run" intrinsic heterogeneity. This raises the question of whether our results on Intrinsic Heterogeneity are robust to extending the model to incorporate econometric learning and dynamic predictor selection. We discuss this issue at length in Section 4 below.

3.3. Connection to the Rational Expectations Equilibrium

Our equilibrium differs from the RPE in Evans and Honkapohja [11]. There agents also underparameterize, but a single perceived law of motion is imposed and all agents are homogeneous in their misspecification. These expectations differ from rational expectations by ignoring relevant information. Since all agents ignore the same information in their perceived law of motion it is clear that in equilibrium the parameters of the model will differ from a Rational Expectations Equilibrium (REE). In a ME with Intrinsic Heterogeneity, each agent uses an underparameterized model, but aggregate expectations are conditioned on all available information. In principle, it is conceivable that a ME could reproduce the REE. In this subsection we use the bivariate example to show that this is not the case: the price process in a ME will differ from the process in an REE.

Recall that

$$p_t = -\phi p_t^{\mathrm{e}} + \gamma' A z_{t-1} + \eta_t, \tag{15}$$

where γ is (2×1), A is (2×2) with elements a_{ij} for $j = 1, 2$, and $\eta_t = \gamma' \varepsilon_t + v_t$. Under rational expectations

$$p_t^{\mathrm{e}} = E_{t-1} p_t. \tag{16}$$

An REE is a stochastic process p_t that satisfies (15) and (16). The cobweb model has a unique REE given by

$$p_t = \hat{\xi}_1 z_{1,t-1} + \hat{\xi}_2 z_{2,t-1} + \eta_t,$$

where

$$\hat{\xi}_1 = (1+\phi)^{-1}(\gamma_1 a_{11} + \gamma_2 a_{21}),$$
$$\hat{\xi}_2 = (1+\phi)^{-1}(\gamma_1 a_{12} + \gamma_2 a_{22}).$$

The parameters in a ME are given by

$$\begin{bmatrix} 1+n_1^*\phi & \phi n_1^*\rho \\ \phi(1-n_1^*)\tilde{\rho} & 1+(1-n_1^*)\phi \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} = A'\gamma, \tag{17}$$

where $n_1^* \in N_\alpha$. We saw that a non-trivial solution to (17) exists for all $\phi \geqslant 0$ and is given by

$$\begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} = \frac{1}{\Delta} \begin{bmatrix} (1+(1-n_1^*)\phi)(\gamma_1 a_{11} + \gamma_2 a_{21}) - \phi n_1^*\rho(\gamma_1 a_{12} + \gamma_2 a_{22}) \\ (1+n_1^*\phi)(\gamma_1 a_{12} + \gamma_2 a_{22}) - \phi(1-n_1^*)\tilde{\rho}(\gamma_1 a_{11} + \gamma_2 a_{21}) \end{bmatrix},$$

where $\Delta = (1+n_1^*\phi)\left(1+(1-n_1^*)\phi\right) - \phi^2 n_1^*\rho\tilde{\rho}$.

Clearly the REE parameters $(\hat{\xi}_1, \hat{\xi}_2)'$ differ from the ME parameters $(\xi_1, \xi_2)'$. For example, consider the case when the random variables $z_{1,t}, z_{2,t}$ are uncorrelated. Then

$$\xi_1 = \left(1+n_1^*\phi\right)^{-1}\left(\gamma_1 a_{11} + \gamma_2 a_{21}\right),$$
$$\xi_2 = \left(1+(1-n_1^*)\phi\right)^{-1}\left(\gamma_1 a_{12} + \gamma_2 a_{22}\right).$$

3.4. Further discussion

The intuition behind Condition P and the existence of Intrinsic Heterogeneity is subtle. In a cobweb model the exogenous shocks z have both a direct and an indirect effect on price. The direct effect is simply the $\gamma' z_t$ term in (1). The indirect effect depends on the way in which agents incorporate z into their expectations p_t^e. It is the interplay between the direct and indirect effects that makes intrinsic heterogeneity possible. In this subsection we express Conditions P, P0 and P1 parametrically and illustrate them with a simple example.

From the equations for expected profit, it can be shown that [13]

$$F(1) \gtrless 0 \quad \text{iff} \quad \xi_1^2(1) \gtrless \xi_2^2(1)\,Q, \text{ and}$$
$$F(0) \gtrless 0 \quad \text{iff} \quad \xi_1^2(0) \gtrless \xi_2^2(0)\,Q,$$

where $Q = \frac{Ez_2^2}{Ez_1^2} > 0$ and, as before, $\xi_j^2(n) = (\xi_j(n))^2$, $j = 1, 2$. Furthermore, from (14) we have

$$\frac{\xi_1^2(0)}{\xi_2^2(0)} = \frac{(1+\phi)^2(\gamma_1 a_{11} + \gamma_2 a_{21})^2}{(\gamma_1 a_{12} + \gamma_2 a_{22} - \phi\tilde{\rho}(\gamma_1 a_{11} + \gamma_2 a_{21}))^2} \equiv B_0,$$
$$\frac{\xi_1^2(1)}{\xi_2^2(1)} = \frac{(\gamma_1 a_{11} + \gamma_2 a_{21} - \phi\rho(\gamma_1 a_{12} + \gamma_2 a_{22}))^2}{(1+\phi)^2(\gamma_1 a_{12} + \gamma_2 a_{22})^2} \equiv B_1.$$

[13] The Appendix contains additional details of these derivations.

These expressions assume that the denominators of the expressions are nonzero. Recall that Q, ρ, and $\tilde{\rho}$ are determined by A and Σ_ε. The above results and Lemma 5 imply:

Lemma 8. *There are three possible cases depending on ϕ, γ, A and Σ_ε.*

1. *Condition* P: $F(0) > 0$ *and* $F(1) < 0$. *Condition* P *is satisfied when* $B_1 < Q < B_0$.
2. *Condition* P0: $F(0) < 0$ *and* $F(1) < 0$. *Condition* P0 *is satisfied when* $Q > B_0$.
3. *Condition* P1: $F(0) > 0$ *and* $F(1) > 0$. *Condition* P1 *is satisfied when* $Q < B_1$.

A special case provides insight into the conditions in which each case arises. Suppose that the components $z_{1,t}$, $z_{2,t}$ are uncorrelated. Then the RPE is given by

$$\begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} = \begin{bmatrix} (1+n_1\phi)^{-1} & 0 \\ 0 & (1+(1-n_1)\phi)^{-1} \end{bmatrix} \begin{bmatrix} \gamma_1 a_{11} + \gamma_2 a_{21} \\ \gamma_1 a_{12} + \gamma_2 a_{22} \end{bmatrix}.$$

Recall that

$$p_t = \xi_1 z_{1,t-1} + \xi_2 z_{2,t-1} + \eta_t.$$

Now set $\phi = 0$. This is the case where there is no feedback from expectations to price. In this special case

$$\xi_1 = \left(\gamma_1 a_{11} + \gamma_2 a_{21}\right),$$
$$\xi_2 = \left(\gamma_1 a_{12} + \gamma_2 a_{22}\right).$$

The parameters ξ_1, ξ_2 are completely determined by the direct effect $\gamma' A$. For $\phi > 0$, the RPE parameters are

$$\xi_1 = (1+n_1\phi)^{-1}\left(\gamma_1 a_{11} + \gamma_2 a_{21}\right),$$
$$\xi_2 = (1+(1-n_1)\phi)^{-1}\left(\gamma_1 a_{12} + \gamma_2 a_{22}\right)$$

and now depend both on the direct effect $\gamma' A$ and the indirect effect of expectations through n_1 and ϕ. Note in particular that as $n_1 \to 1$ we have $|\xi_1(n_1)| \downarrow$ and $|\xi_2(n_1)| \uparrow$. For a given ϕ the indirect effect depends on n_1. As agents mass onto a particular predictor it diminishes the effect of that variable. This is because of the self-referential feature of the cobweb model that leads to an indirect effect on prices opposite to the direct effect of that variable. This makes $z_{1,t}$ a less useful predictor than before, and thus the $z_{2,t}$ component becomes more profitable. The opposite happens as $n_1 \to 0$ and consequently there is a unique n_1 in which both predictors fare equally well in terms of mean profits. This proportion is the limit point of Intrinsic Heterogeneity.

Condition P places conditions on the indirect and direct effects and on the relative importance of the two exogenous variables. In our simple example of uncorrelated shocks Condition P is equivalent to

$$\frac{\left(\gamma_1 a_{11} + \gamma_2 a_{21}\right)^2}{(1+\phi)^2\left(\gamma_1 a_{12} + \gamma_2 a_{22}\right)^2} < Q < \frac{(1+\phi)^2\left(\gamma_1 a_{11} + \gamma_2 a_{21}\right)^2}{(\gamma_1 a_{12} + \gamma_2 a_{22})^2},$$

where $Q = \frac{Ez_2^2}{Ez_1^2}$. When there is no feedback ($\phi = 0$) there does not exist a matrix A and Σ_ε which satisfies Condition P. Intrinsic Heterogeneity does not exist in this instance. Because

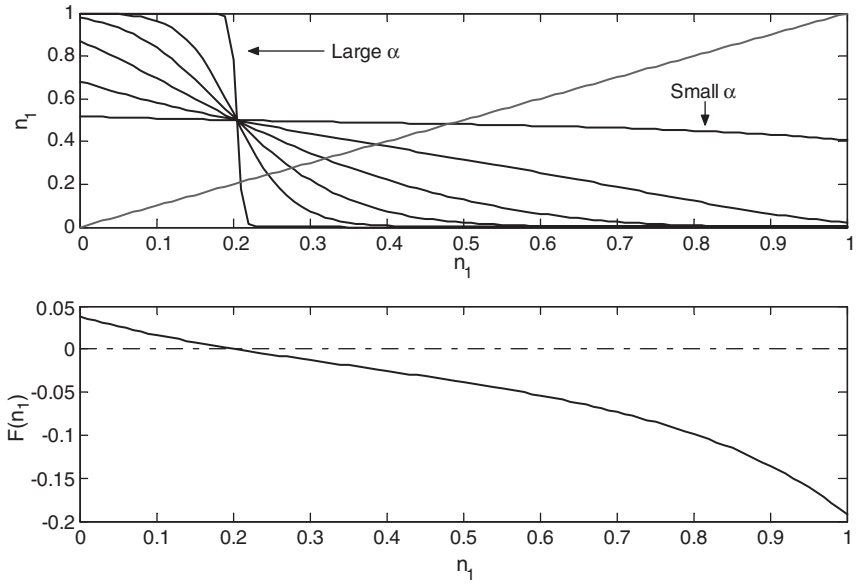

Fig. 1. T-map for various values of α and $\phi = 2$.

there is no indirect effect from expectations, and expectations have no bearing on price, agents will choose the predictor that forecasts price best. This is also true if the feedback is small: for (almost) all A, Q and Σ_ε, condition P cannot hold for $\phi > 0$ sufficiently small. However, as ϕ increases, the range of admissible Q increases. Given A, γ and Q, condition P will hold for $\phi > 0$ sufficiently large. A numerical example of this point is provided below.

3.5. Numerical examples

We illustrate our results numerically. Fig. 1 gives the T-maps for various values of α. The upper part of the figure shows the T-maps corresponding to (starting from $n_1 = 0$ and moving clockwise) $\alpha = 2, 20, 50, 100, 200, 2000$. We set

$$A = \begin{bmatrix} .3 & .10 \\ .10 & .7 \end{bmatrix},$$

$\gamma' = [.7, .5]$,

$$\Sigma_\varepsilon = \begin{bmatrix} .7 & .2 \\ .2 & .6 \end{bmatrix}$$

and $\phi = 2$. The bottom portion of the figure is the profit difference function $F(n_1)$.

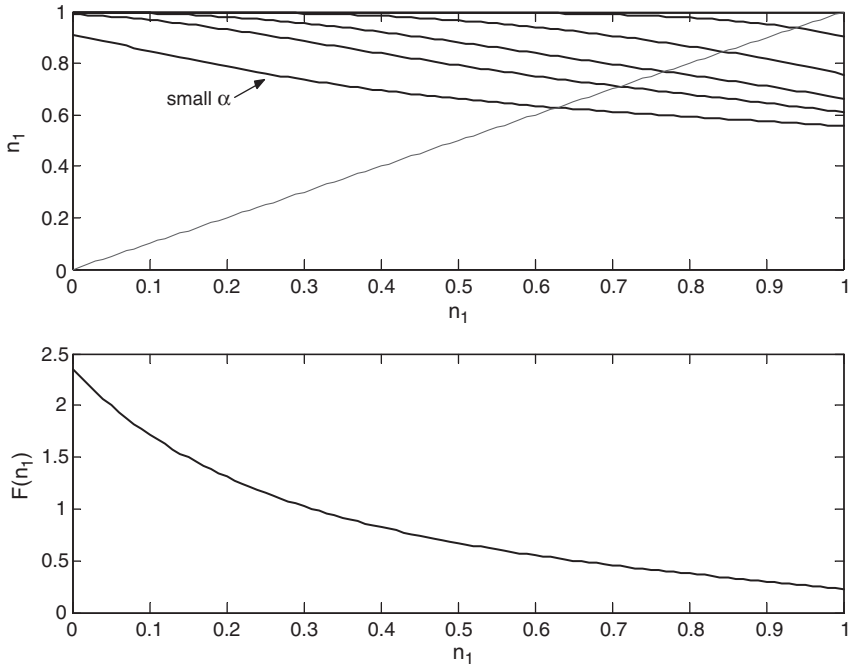

Fig. 2. T-map for various values of α and $\phi = 2$ for the case of no Intrinsic Heterogeneity and predictor 1 homogeneity.

The matrix A and parameter ϕ have been chosen so that Condition P holds, i.e. $F(1) < 0$. The proof of Proposition 7 shows that as $\alpha \to \infty$

$$H_\alpha(x) \to \begin{cases} 1 & \text{if } x > 0, \\ 0 & \text{if } x < 0, \\ 1/2 & \text{if } x = 0 \end{cases}$$

and clearly this will govern the behavior of $T_\alpha = H_\alpha \circ F$. Fig. 1 illustrates how as α increases the inverse S-shape becomes more pronounced. Fixed points of the T-map are given by intersections with the 45° line and all fixed points of the T-map will intersect this line. As α increases the fixed point declines from above .5 to about .22, which is the point at which $F(\hat{n}_1) = 0$. The ME continues to exhibit heterogeneity even as $\alpha \to \infty$.

Fig. 2 illustrates how heterogeneity may disappear as $\alpha \to \infty$. We now set

$$A = \begin{bmatrix} .93 & .10 \\ .10 & .2 \end{bmatrix},$$

so that Condition P does not hold and instead condition P1 is satisfied. For low values of α some agents continue to use z_2 even though it returns a lower expected payoff. However, as $\alpha \to \infty$ all agents behave optimally and the proportion using z_2 goes to zero.

Fig. 3 shows the role ϕ plays in the degree of Intrinsic Heterogeneity. This graph depicts the T-map for various increasing values of ϕ. Notice that as ϕ increases the fixed point of

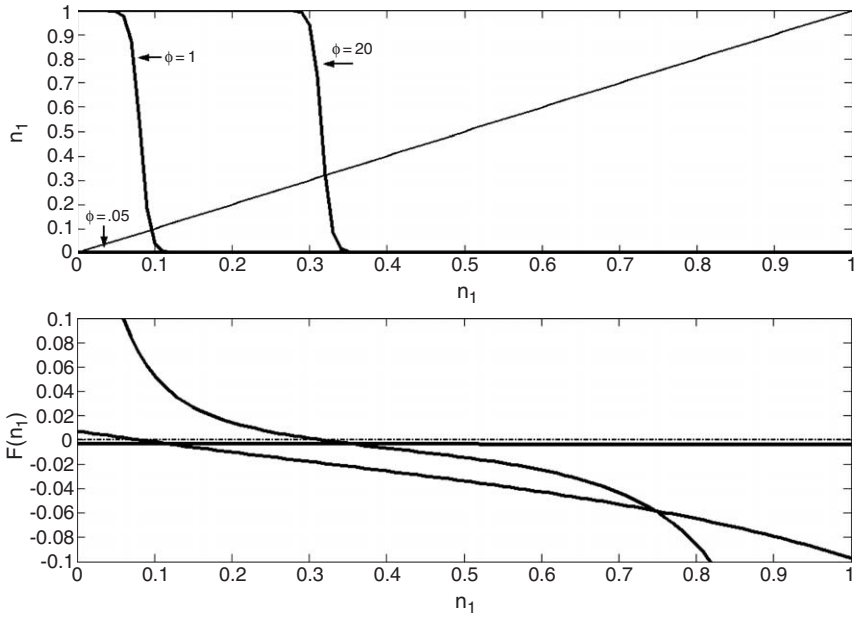

Fig. 3. T-map for $\alpha = 2000$ and $\phi = .05, 1, 20$ for the case of Intrinsic Heterogeneity. Note that as ϕ increases the fixed point of the T-map increases.

the T-map moves further to the right. In this example, z_2 has a stronger influence on the price than z_1. When z_1 has a stronger effect, the fixed point will move to the left. Fig. 3 also illustrates that for sufficiently small ϕ Condition P may not hold as the thick line running along the horizontal axis is for very small ϕ and the fixed point is at zero.

Note that in a ME, in a model with Intrinsic Heterogeneity, all predictors have the same average return as the intensity of choice α becomes large. When α is finite there can be differences in the relative performance of predictors, but as $\alpha \to \infty$ the mean returns across predictor must converge given our assumption of costless (or equally costly) predictors. Heterogeneity arises in the costless case of Brock and Hommes [7] only in the steady-state in which different predictors make identical forecasts. Our results arise in a stochastic equilibrium in which different predictors produce different forecasts, but achieve identical mean performance, as $\alpha \to \infty$.

4. Long-run Intrinsic Heterogeneity: least-squares learning and dynamic predictor selection

A ME is a fixed point defining a stochastic process in which each predictor has optimally chosen coefficients and in which the proportions of agents using each predictor depend on their mean profits. Intrinsic Heterogeneity arises in the model when, in the neoclassical limit, each predictor returns the same mean profits in the ME. As discussed in the Introduction, the attainability of the equilibrium under real-time learning is also of considerable interest.

The literature on least-squares learning emphasizes stability under econometric learning as an important criterion for whether REE are reasonable, and in the current framework there is the additional issue of dynamic predictor selection.

Stability under learning also facilitates consideration of two related issues of interest. The Friedman [15] hypothesis, outlined in the introduction, can only be analyzed in a dynamic process that takes into account real-time learning. If this adaptive process converges to a ME then this supports Friedman's hypothesis that market forces are stabilizing. Secondly, by explicitly modeling the dynamics we can sharpen the comparison with Brock and Hommes [7] and Brock and Hommes [8], who emphasize the global dynamics of predictor selection. Since their emphasis is on the possibility of complex dynamics around the steady state, we want to study the conditions in which this might emerge in our framework.

4.1. Stability under real-time learning and dynamic predictor selection

In this section we address whether the equilibrium is attainable under real-time learning of the type emphasized in Evans and Honkapohja [11]. In a ME agents misspecify, but their forecasts are the optimal linear projections given their underparameterization. Furthermore, agents choose which component of the exogenous process to omit based on unconditional mean profits. We now substitute optimal linear projections with real-time estimates formed via recursive least squares (RLS).[14] We also assume that agents choose their forecast model each period based on a real-time estimate of mean profits. Hence, there is dual learning as agents recursively update the parameters of their forecasting model and evolve their predictor choice according to a dynamic predictor selection mechanism.

Prices now depend on time-varying parameters b^j_{t-1} and time-varying predictor proportions $n_{j,t-1}$,

$$p_t = \xi_1(b^1_{t-1}, n_{1,t-1})z_{1,t-1} + \xi_2(b^2_{t-1}, n_{1,t-1})z_{2,t-1} + \eta_t$$

in which b^1_{t-1}, b^2_{t-1} are updated by RLS

$$b^1_t = b^1_{t-1} + t^{-1}R^{-1}_{1,t}z_{1,t-1}\left(p_t - b^1_{t-1}z_{1,t-1}\right),$$
$$b^2_t = b^2_{t-1} + t^{-1}R^{-1}_{2,t}z_{2,t-1}\left(p_t - b^2_{t-1}z_{2,t-1}\right),$$

where

$$R_{1,t} = R_{1,t-1} + t^{-1}(z^2_{1,t-1} - R_{1,t-1}),$$
$$R_{2,t} = R_{2,t-1} + t^{-1}(z^2_{2,t-1} - R_{2,t-1}).$$

The $R_{j,t}$, $j = 1, 2$ are recursive estimates of the variances of the explanatory variables z_j.

Predictor proportions are updated according to the discrete choice probabilities. Mean profits of the two predictors $j = 1, 2$ are estimated by

$$\hat{E}\pi_{j,t} = \hat{E}\pi_{j,t-1} + \delta_t\left(\pi_{j,t} - \hat{E}\pi_{j,t-1}\right),$$

[14] For an overview of stability under RLS in dynamic macroeconomics see Evans and Honkapohja [11].

where $0 < \delta_t < 1$ and we focus on two cases that are standard in the literature: (i) $\delta_t = t^{-1}$, which delivers a recursive algorithm for computing the average, with equal weights, of past realized profits, and (ii) $\delta_t = \delta$, which weights recent profits more heavily, with the weights declining geometrically at rate $1 - \delta$.

The mean profits map into predictor proportions according to the law of motion

$$n_{j,t} = \frac{\exp\left[\alpha \hat{E}\pi_{j,t}\right]}{\sum_{j=1}^{2} \exp\left[\alpha \hat{E}\pi_{k,t}\right]}.$$

The dynamic version of the model exhibits real-time learning in the sense that agents adaptively update previous estimates of their belief parameters and the mean profits of each predictor. Agents now choose their model in each time period based on these recursive estimates. We are interested in whether the sequence of estimates b_t^1, b_t^2 and predictor proportions $n_{1,t}$ converge to the ME.[15] Our aim is to use numerical illustrations to show that the equilibrium can be stable under real-time learning and to investigate whether dynamic instability can also arise. It is beyond the scope of this paper to establish analytical convergence results.

We continue with the particular parameterization that generated Intrinsic Heterogeneity in the previous section. We set

$$A = \begin{bmatrix} .3 & .1 \\ .1 & .7 \end{bmatrix}, \quad \Sigma_\varepsilon = \begin{bmatrix} .7 & .2 \\ .2 & .6 \end{bmatrix},$$

$\gamma' = [.7, .5]$, and $\phi = 2$, and simulate the model for 50,000 time periods. We set the initial value of the VAR equal to a realization of its white noise shock, i.e., $z_0 = \varepsilon_0$. The initial value for $n_{1,0}$ is .82, a value that was chosen to lie away from the end points and the ME. Initial estimated mean profits are equal to the realized profits under the initial conditions. The initial belief parameters were set to $b_0^1 = 1, b_0^2 = 2$. The initial estimated variances $R_{1,0}, R_{2,0}$ are the identity matrices. Except where otherwise specified we choose $\alpha = 100$.

4.2. Sample average estimator for mean profits

When $\delta_t = t^{-1}$ we have

$$\hat{E}\pi_{j,t} = \hat{E}\pi_{j,t-1} + t^{-1}\left(\pi_{j,t} - \hat{E}\pi_{j,t-1}\right),$$

which is equivalent to computing sample averages, for each predictor, for $t = 1, 2, 3, \ldots$. This case leads to dynamic stability of the ME.[16]

Fig. 4 illustrates the results of a representative simulation for $\alpha = 100$. The top panel plots the simulated proportion $n_{1,t}$ against time. The middle and bottom panels plot the simulated

[15] Since we conduct the analysis numerically, we are being deliberately vague in what sense these sequences converge.

[16] In their "Final Remarks" Brock and Hommes [7] mention the case of equal weighting, and conjecture convergence to rational expectations. In our framework full rationality is excluded by construction and we find convergence to the ME.

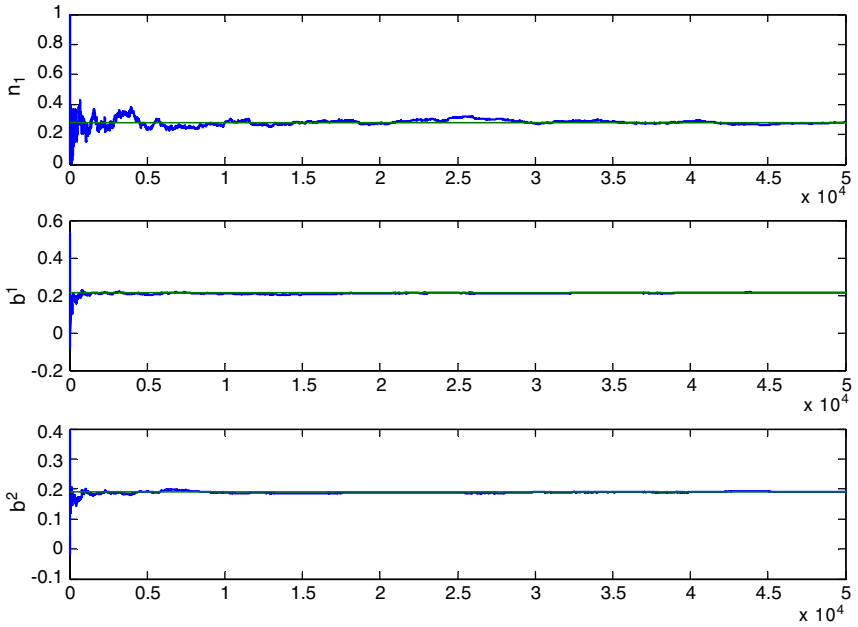

Fig. 4. Real time learning simulations with $\delta_t = t^{-1}$.

belief parameters b_t^1, b_t^2. In each plot the solid horizontal line represents the respective variables' values in the ME with Intrinsic Heterogeneity. As can be seen, there appears to be convergence to the ME. (The paths of law of motion parameters $\xi_{j,t}$, not shown, also indicate convergence). Initially there is considerable volatility in the proportion of agents who choose predictor 1. This volatility gradually dampens until the proportion approaches its equilibrium value. The dampening is much quicker in the belief parameters b_t^j, which approach their equilibrium values in a short period of time. Similar convergence results are obtained for other parameter settings, though speed of convergence is affected by α. For larger values of α it takes longer for the predictor proportions to settle down near the equilibrium values. However, the system appears to be stable for all $\alpha > 0$.

The intuition behind the stability is as follows. In our parameterization there is a unique ME with Intrinsic Heterogeneity. Heterogeneity arises because Condition P guarantees that under, say, z_1 homogeneity agents will have an incentive to mass on z_2, and vice-versa. For large α agents mass on the predictor that returns the highest mean profit. In our simulations the proportions of agents are initially well away from the ME. This implies that one predictor has a higher profit than the other. In the next period agents mass onto that predictor. Because Condition P holds, in the next period agents tend to mass onto the other predictor. As the rapid switching occurs agents update parameter estimates, which converge quickly, and accumulate data on relative mean forecast performance. As they learn about mean relative forecast performance, the volatility in predictor selection dampens and there is convergence towards the Misspecification Equilibrium. The system can therefore be said to generate "long-run intrinsic heterogeneity" in the sense that the heterogeneity of forecasts at each moment in time emerges as the long-run outcome of a dynamic process of adjustment.

In light of Brock and Hommes [7] our results may seem surprising. However, in Brock and Hommes [7] the model is deterministic, the predictor choice is between a costly stabilizing predictor and a costless destabilizing predictor, and predictor fitness is the most recent period's realized profits. The stability results in our model are the result of agents looking at the mean relative performance of the predictors using sample averages. Measuring fitness or expected profits using only the most recent period's profits is not appropriate in our stochastic framework. However, this is a special case of geometrically declining weighting, to which we now turn.

4.3. Geometric estimator for mean profits

We now consider the case $0 < \delta_t = \delta \leqslant 1$, i.e.

$$\hat{E}\pi_{j,t} = \hat{E}\pi_{j,t-1} + \delta\left(\pi_{j,t} - \hat{E}\pi_{j,t-1}\right).$$

Given initial values for $\hat{E}\pi_{j,0}$, for $t = 1, 2, 3, \ldots$. this weights $\pi_{j,t-i}$ by $\delta(1-\delta)^i$, for $i = 0, \ldots, t-1$ and $\hat{E}\pi_{j,0}$ by $(1-\delta)^t$. For $\delta \to 0$ this estimator of mean profits (or "fitness" measure) carries the limiting interpretation of a sample average in which all past profits are weighted equally. The $\delta = 1$ case, in which only the most recent period's profits matter, was emphasized in Brock and Hommes [7,8]. Brock and Hommes [7] do, however, set out a more general fitness measure and explicitly mention the geometric case.

Figs. 5 and 6 plot two real-time learning trajectories, with $\alpha = 100$, for polar cases $\delta =$ .0001 and .25. We maintain the same model parameter values, and the forecast parameter estimates b_t^j continue to be estimated by RLS. [17] In Fig. 5 the dynamics settle down very close to the ME. In contrast, in Fig. 6, the proportions oscillate frequently and erratically, in many periods taking values near $n_{1t} = 0$ or $n_{1t} = 1$.

The intuition for these results is as follows. For a sufficiently small weight on recent observations, the geometric average approximates the sample average used in the previous subsection. As the weight on recent observations increases, agents will "overreact" to the latest data and switch predictor functions. This will tend to cause price to move in the opposite direction making the other predictor return a higher profit in that period. After several periods this will be sufficient to induce agents to mass on the other predictor, with the exact timing depending on the realized sequence of random shocks. [18]

Other simulations (not shown) reveal the following qualitative features. As δ is increased from very small levels, the paths of n_{1t} and ξ_{jt} remain centered around ME values, but

[17] RLS could also be replaced by "constant gain" versions, which use $\lambda \in (0, 1)$ in place of t^{-1}. These have been advocated in models in which agents are concerned about possible regime change, e.g. Sargent [20]. See Evans and Ramey [14] for a setup in which unknown regime changes are present and agents aim to optimally chose λ. Because in our setup there is a unique ME and no structural change, we use the RLS algorithm here. We note, though, that for $\lambda \to 0$ the results will be qualitatively very close. In a companion paper we pursue constant gains in a setting where there are multiple Misspecification Equilibria.

[18] What would happen if the exogenous noise were removed? If the innovation variances Σ_ε and σ_v^2 are replaced by $\tau\Sigma_\varepsilon$ and $\tau\sigma_v^2$ for $0 < \tau < 1$ this would simply shrink the variation of p_t around the steady state, approaching the steady state path asymptotically as $\tau \to 0$. Within our framework one cannot easily look at the nonstochastic case $\tau = 0$ since in that case $\Omega = 0$ and the least-squares projections are undefined.

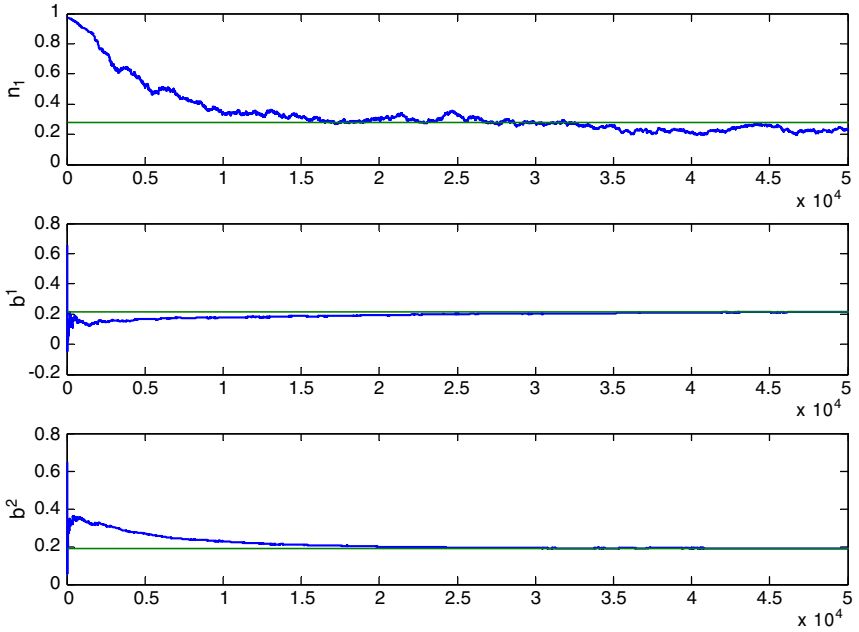

Fig. 5. Real time learning simulations with $\delta_t = \delta = .0001$.

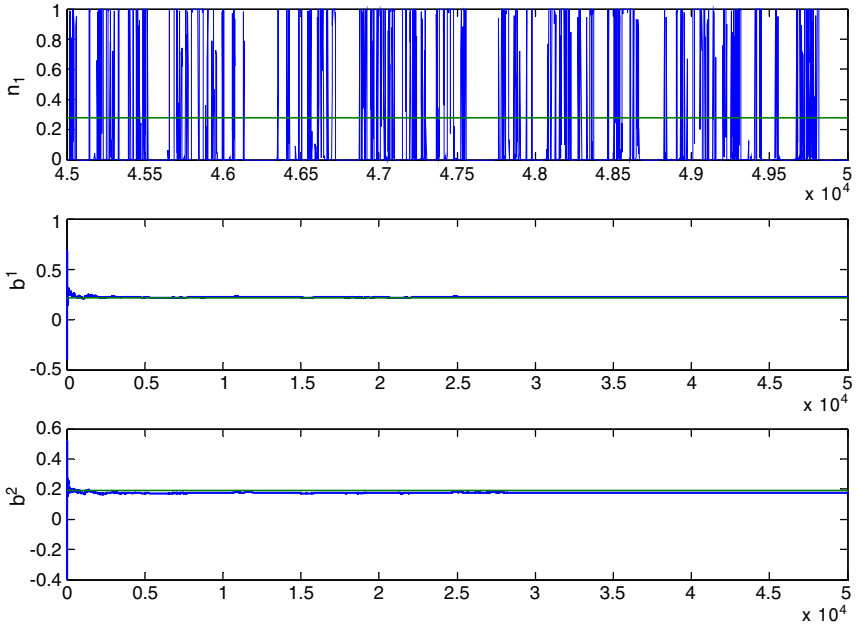

Fig. 6. Real time learning simulations with $\delta_t = \delta = .25$.

Table 1
Summary of numerical results

	$\alpha = 100$ $n_1^* = .2763$			$\alpha = 20$ $n_1^* = .3852$		
δ	$\bar{n}_1$	σ_{n_1}	ψ_{n_1}	$\bar{n}_1$	σ_{n_1}	ψ_{n_1}
0.0001	.2895	.0001	0	.4005	.0003	0
0.001	.2664	.0005	0	.3853	.0007	0
0.01	.2905	.0462	.1103	.3954	.0082	0
0.05	.3367	.1274	.4407	.3997	.0338	.0113
0.10	.3660	.1623	.5903	.4087	.0575	.0538
0.25	.4109	.1983	.7411	.4340	.0965	.1829

Note: n_1^* denotes the ME values. $\bar{n}_1$ and σ_{n_1} give the sample mean and standard deviation for the simulation. ψ_{n_1} gives the proportion of values for n_1 that were either above .95 or below .05 in the simulation. Each simulation was of length 60,000 with the first 10,000 data points discarded.

increase in volatility, as measured, for example, by their variances. As δ is increased further, the time-series variance in n_{1t} increases until eventually a high proportion of the time is spent with n_{1t} near its boundary values.

These results are illustrated in Table 1, computed from simulations of length 50,000 periods. The table also shows how the qualitative features depend on the value of α: for larger values of α we require smaller values of δ to keep the sample paths close to ME values.

We do not have a strong prior view as to the appropriate value of δ, and the results of Table 1 document the range of possibilities. In stochastic economies it seems clear that predictor performance should be assessed using an average of past profits, and the results of this section show how the degree of smoothing controls key qualitative features of cross-sectional and time-series heterogeneity.

4.4. Discussion

In this section we have numerically investigated dynamics incorporating real-time learning and dynamic predictor selection, focusing on the case in which the misspecification equilibrium exhibits Intrinsic Heterogeneity. For the sample average estimate of mean profits ($\delta_t = t^{-1}$), our numerical simulations show long-run intrinsic heterogeneity, i.e. convergence over time to a ME with intrinsic heterogeneity. Similar results are obtained using geometric estimators of mean profits with sufficiently small values of δ: predictor proportions now remain stochastic over time, but the sample paths are close to ME values. However, as δ increases, the variability of sample paths for n_{1t} increases and eventually exhibits sample paths with frequent rapid switches and in which most of the time is spent near boundary values of $n_{1t} = 0$ or 1. These latter results are clearly analogous to those of Brock and Hommes [7] for the case of large α and a positive cost to the rational expectations predictor, but arise in the current case in a stochastic model with two equally costly and equally sophisticated predictors.

One way to interpret the results of this section is that because of the negative feedback in the cobweb model from expectations to prices, when the conditions for intrinsic heterogeneity are met, heterogeneous expectations must arise in some form, even for large intensities of choice α. How heterogeneity arises is controlled by the sensitivity δ_t of the mean profit or "fitness" measure $\hat{E}\pi_{j,t}$ to the most recent period's profits. When $\delta_t = t^{-1}$ or takes a small constant value, heterogeneity takes the form of constant or near-constant positive proportions of agents using each of the forecast rules. This generates a cross-sectional dispersion of price forecasts at each point in time. When δ is sufficiently large, agents tend to mass onto a particular predictor that appears currently to yield higher profits, and the heterogeneity instead takes the form of frequent switching between estimators over time.

These results also suggest that the validity of the Friedman Hypothesis depends on the nature in which agents discount the past. Whether market forces are stabilizing and guide the economy to our ME depends on the sensitivity of firms to recent predictor profits versus more permanent expected profit measures of predictor success.

5. Conclusion

This paper demonstrates how to obtain heterogeneous expectations as an equilibrium outcome in a model with optimizing agents. Our setup is the standard cobweb model in which rational expectations was originally developed. We obtain our results with a discrete choice model for predictors, when agents are constrained to choose from a set of misspecified models. As in Brock and Hommes [7] the proportion of agents using the different predictors depends on their relative performance according to an intensity of choice parameter, but we focus on the case in which predictor performance is measured by expected profits. We extend this approach by considering predictors that are alternative econometric forecasting models. As the intensity of choice increases agents will select only the most successful predictors. In Brock and Hommes [7] heterogeneity of expectations is a reflection of finite intensities of choice and disappears in the neoclassical limit. One of the main contributions of our paper is to show that heterogeneity can remain for high intensities of choice as a result of the availability of multiple misspecified models.

Because of limits to cognition, knowledge of the economy, and degrees of freedom, we assume that agents must underparameterize by neglecting a variable or lag from their forecasting model. The importance of misspecification is widely recognized in applied econometrics and one that we believe should be reflected in realistic models of bounded rationality. Although we constrain agents to choose from a list of misspecified models, at the same time we require that the parameters of each chosen model are formed optimally in the sense that forecast errors are orthogonal to the explanatory variables of that model.

Our major theoretical contribution is to obtain existence results for a ME within this framework and to obtain a suitable condition under which heterogeneous expectations persist for high intensities of choice. When this condition is satisfied we say the model exhibits Intrinsic Heterogeneity.

Our central finding that misspecification can lead to heterogeneous expectations is not at all obvious. If the intensity of choice is large, a key requirement for this possibility is that the model be self-referential, i.e., that there be feedback from expectations to actual

outcomes. Heterogeneous expectations are not a necessary outcome when the intensity of choice is large, but do arise under a suitable joint condition on the model and the exogenous driving processes. We illustrate the results in a simple bivariate model. In particular, we show that, ceteris paribus, Intrinsic Heterogeneity arises when the parameter governing the self-referential extent of the model is sufficiently large. This surprising feature of self-referential models has not been noted in previous work.

Finally, we have shown numerically that a ME with intrinsic heterogeneity can emerge from a dynamic process of least-squares parameter learning and evolving predictor choices, provided agent's predictor choices are based on mean profits as estimated by sample averages. If instead agents place a high weight on more recent profits, as in Brock and Hommes [7,8], then the system can generate complicated dynamics around the ME.

In this paper we have focused on the cobweb model. In future work, we will examine the framework in a Lucas-type monetary model. The Lucas-type model shares a similar reduced-form as the cobweb model, but expectations have a positive feedback on price. Since the self-referential feature of these models is central, a model with positive feedback can be expected to yield distinct results.

Acknowledgments

We are greatly indebted to Garey Ramey for early discussions. We thank Jim Bullard, Cars Hommes, Didier Sornette, and participants at the 2002 CeNDEF Workshop on Economic Dynamics for helpful comments. We are particularly grateful to an Associate Editor and anonymous referee for detailed suggestions. This material is based upon work supported by the National Science Foundation under Grant No. 0136848.

Appendix

Proof of Proposition 1. Consider the matrix

$$[\Delta] = \left(I + \phi \sum_{j=1}^{K} n^j \Omega' u^{j'} \left(u^j \Omega u^{j'} \right)^{-1} u^j \right).$$

The absolute value of the indicated sum has a maximum value when considered as a function of $n \in S$. Hence there exists $\varepsilon > 0$ such that, for all $0 \leqslant \phi < \varepsilon$, $[\Delta]$ is strictly diagonally dominant (see Horn and Johnson [18, p. 302] for all $n \in S$. Strictly diagonally dominant matrices have nonzero determinants and hence are invertible. □

Proof of Theorem 3. Suppose to the contrary that the model does not exhibit intrinsic heterogeneity. From Theorem 2 we know that a ME exists for every α. Since the model does not have intrinsic heterogeneity, then for all $\bar{n} < 1$ there are infinitely many α such that $n_k^* > \bar{n}$ for some component $k = 1, \ldots, K$ where $n^* \in N_\alpha$. Hence there exists a sequence indexed by $\hat{s}$ such that $\alpha(\hat{s}) \to \infty$ with fixed points $n^*(\hat{s})$ satisfying $n^*_{k(\hat{s})}(\hat{s}) \to 1$. It follows that for some $i \in \{1, \ldots, K\}$ there exists a subsequence indexed by s such that $\alpha(s) \to \infty$

and $n_i^*(s) \to 1$. The expected profit functions $\tilde{F}_j(n)$ are continuous and hence for this sequence

$$E\pi^k(s) - E\pi^i(s) \to \tilde{F}_k(e_i) - \tilde{F}_i(e_i)$$

for all $k = 1, \ldots, K$, where e_i is the unit coordinate vector with component i equal to one. However, condition P implies that there exists $j \neq i$ such that $\tilde{F}_j(e_i) - \tilde{F}_i(e_i) > 0$. It follows from (12) that

$$n_i^*(s) = \frac{1}{1 + \sum_{k \neq i} \exp\{\alpha(s)(E\pi^k(s) - E\pi^i(s)\}}.$$

Thus $n_i^*(s) \to 0$ as $s \to \infty$. This contradicts $n_i^*(s) \to 1$ and hence the model must exhibit intrinsic heterogeneity. □

Proof of Proposition 4. We show, for every $\phi \geqslant 0$, that

$$\Delta(n_1) = (1 + n_1\phi)\left((1 + \phi) - \phi n_1\right) - \phi^2\rho\tilde{\rho}\left(n_1 - n_1^2\right) > 0$$

for all $0 \leqslant n_1 \leqslant 1$. Equivalently

$$\Delta(n_1) = \phi^2\left(\rho\tilde{\rho} - 1\right)n_1^2 + \phi^2\left(1 - \rho\tilde{\rho}\right)n_1 + (1 + \phi) > 0$$

for all $0 \leqslant n_1 \leqslant 1$. Since $|\rho\tilde{\rho}| \leqslant 1$ by the Schwarz inequality, $\Delta(n_1)$ is a concave quadratic function of n_1. Clearly $\Delta(0) > 0$ and $\Delta(1) > 0$. Since $\Delta(n_1)$ is concave and is positive at the end points of the domain $0 \leqslant n_1 \leqslant 1$, it follows that $\Delta(n_1) > 0$ for all $0 \leqslant n_1 \leqslant 1$. Hence Condition Δ is satisfied. □

Proof of Lemma 5. We can rewrite (14) as

$$S(n_1)\xi = A'\gamma,$$

where $\xi' = (\xi_1, \xi_2)$ and $S(n_1)$ is the indicated 2×2 matrix. Differentiating we obtain $(dS)\xi + S(d\xi) = 0$ and

$$\frac{d\xi}{dn_1} = -S^{-1}\frac{dS}{dn_1}\xi.$$

It is easily seen that

$$\frac{dS}{dn_1} = \phi\begin{pmatrix} 1 & \rho \\ -\tilde{\rho} & -1 \end{pmatrix}.$$

Somewhat abusing notation, it is now convenient to rewrite $F(n_1)$ as $F(\xi(n_1))$. To establish the result we compute $dF/dn_1 = (dF/d\xi)'(d\xi/dn_1)$. It can be verified that

$$\left(\frac{dF}{d\xi}\right)' = \phi D(1 - r^2)Ez_1^2\xi'\begin{pmatrix} 1 & 0 \\ 0 & -Q \end{pmatrix},$$

where $Q = Ez_2^2/Ez_1^2$.

Thus

$$dF/dn_1 = -\phi^2 D(1-r^2)Ez_1^2\xi' K(n_1)\xi, \text{ where}$$
$$K = \begin{pmatrix} 1 & 0 \\ 0 & -Q \end{pmatrix} S^{-1} \begin{pmatrix} 1 & \rho \\ -\tilde{\rho} & -1 \end{pmatrix}.$$

Here $r^2 = \rho\tilde{\rho}$ with $0 \leqslant r^2 < 1$. Computation of K yields

$$K(n_1) = \begin{pmatrix} \frac{1+\phi(1-n_1(1-r^2))}{1+\phi+\phi^2(1-r^2)n_1(1-n_1)} & \frac{\sqrt{Q}r(1+\phi)}{1+\phi+\phi^2(1-r^2)n_1(1-n_1)} \\ \frac{\sqrt{Q}r(1+\phi)}{1+\phi+\phi^2(1-r^2)n_1(1-n_1)} & \frac{Q(1+\phi r^2+\phi(1-r^2)n_1)}{1+\phi+\phi^2(1-r^2)n_1(1-n_1)} \end{pmatrix}.$$

$K(n_1)$ is symmetric with $K_{11}(n_1) > 0$ and

$$\det(K(n_1)) = \frac{Q(1-r^2)}{1+\phi+\phi^2(1-r^2)n_1(1-n_1)} > 0.$$

Thus $K(n_1)$ is positive definite and $\xi' K(n_1)\xi \geqslant 0$ for all ξ. The result follows since $dF/dn_1 \leqslant 0$ for all $0 \leqslant n_1 \leqslant 1$. □

Proof of Proposition 7. Take part (1), which states that Condition P implies Intrinsic Heterogeneity. We will establish that (i) for each α, $\exists n_1^*(\alpha) \in N_\alpha$ uniquely, (ii) $\exists \{\alpha(s)\}_s$ *s.t.* $\alpha(s) \to \infty \Rightarrow n_1^*(\alpha(s)) \to \hat{n}_1$ where $\hat{n}_1 \in N_\infty \equiv \{n_1 \in [0,1] :$ for $\alpha \to \infty$ $n_1 = T_\alpha(n_1)\}$ and (iii) $F(\hat{n}_1) = 0$.

Claim (i) that there exists a unique fixed point $n_1^*(\alpha)$ for each α comes directly from Theorem 6.

Claim (ii) is that there is a sequence $\alpha(s)$ indexed by s defined so that as $\alpha(s) \to \infty$ the corresponding sequence of fixed points from claim (i) $n_1^*(\alpha(s)) \to \hat{n}_1$. That there exists a sequence $\alpha(s) \to \infty$ and a similarly corresponding sequence $n_1^*(\alpha)$ follows from claim (i) and since $\alpha \in \mathbb{R}_+$ there are infinitely many such sequences. Theorem 6 used Brouwer's theorem and Lemma 5 to establish that there exists a unique fixed point for each α. Hence there exists a limit to the sequence of fixed points indexed by s and define it to be $n_1^*(\alpha(s)) \to \hat{n}_1$. By construction, $\hat{n}_1 \in N_\infty$.

Claim (iii) is that $F(\hat{n}_1) = 0$. Assume $\hat{n}_1 \in N_\infty$, Condition P, and $F(\hat{n}_1) \neq 0$. It follows that $F(\hat{n}_1) > 0$ or $F(\hat{n}_1) < 0$. Recall, $n_1(\alpha) = H_\alpha(F(n_1))$. By definition, as $\alpha \to \infty$

$$H_\alpha(x) \to \begin{cases} 1 & \text{if } x > 0, \\ 0 & \text{if } x < 0, \\ 1/2 & \text{if } x = 0. \end{cases}$$

So we have $n_1^*(\alpha) \to \hat{n}_1 \in \{0,1\}$. But, assuming Condition P implies $F(1) < 0$ and $F(0) > 0$. Hence, $\hat{n}_1$ is not an ME which contradicts our initial assumption. It must be the case that, with Condition P, $F(\hat{n}_1) = 0$.

Note now that Lemma 5 establishes $\hat{n}_1$ is the unique point where $F(\hat{n}_1) = 0$. Thus, we conclude that Condition P implies $n_1^*(\alpha) \to \hat{n}_1$ where $F(\hat{n}_1) = 0$.

A similar argument establishes parts (2) and (3) of the proposition. Note that Condition P1 implies $F(1) > 0$ and $F(0) > 0$ and Condition P0 has $F(1) < 0$ and $F(0) < 0$. The

monotonicity of F means that $\forall n_1, \alpha\ F(n_1(\alpha)) \neq 0$ and the result follows immediately from above. □

Further details for Section 3.4: Using the expressions for $E\pi_1$ and $E\pi_2$ obtained in Section 3.1, compute $F(1)$ and $F(0)$ by setting $n_1 = 1, 0$, respectively. Using $\rho = Ez_{1t}z_{2t}/Ez_{1t}^2$ and $\tilde{\rho} = Ez_{1t}z_{2t}/Ez_{2t}^2$ we get

$$\begin{aligned}\frac{F(1)}{Ez_{1t}^2} &= -\phi D\{(\xi_1^2(1)\tilde{\rho} - \xi_2^2(1)\rho)\rho + (1/2)(\xi_2^2(1) \\ &\quad -\tilde{\rho}^2\xi_1^2(1))Q - (1/2)(\xi_1^2(1) - \rho^2\xi_2^2(1))\} \\ \frac{F(0)}{Ez_{2t}^2} &= \phi D\{\tilde{\rho}[\xi_2^2(0)\rho - \xi_1^2(0)\tilde{\rho}] + (1/2)[(\xi_1^2(0) \\ &\quad -\xi_2^2(0)\rho^2)Q^{-1} - (\xi_2^2(0) - \xi_1^2(0)\tilde{\rho}^2)]\}.\end{aligned}$$

Thus, for example,

$$F(1) < 0 \quad \text{if} \quad \left[\xi_1^2(1) - \xi_2^2(1)\right]\left(Q\tilde{\rho}^2 - 1\right) > 0.$$

Using $Q\tilde{\rho}^2 = r^2 < 1$ it follows that

$$F(1) < 0 \quad \text{if} \quad \left[\xi_1^2(1) - \xi_2^2(1)\right] < 0. \qquad \square$$

References

[1] W.A. Branch, Local convergence properties of a Cobweb model with rationally heterogeneous expectations, J. Econom. Dynamics Control 27 (1) (2002) 63–85.
[2] W.A. Branch, B. McGough, Replicator Dynamics in a Cobweb Model with Rationally Heterogeneous Expectations, Mimeo, 2003.
[3] W.A. Branch, B. McGough, Consistent expectations and misspecification in stochastic non-linear economies, J. Econom. Dynamics Control, forthcoming.
[4] M. Bray, N. Savin, Rational expectations equilibria, learning and model specification, Econometrica 54 (1986) 1129–1160.
[5] W.A. Brock, S.N. Durlauf, Discrete choice with social interactions, Rev. Econom. Stud. 68 (2) (2001) 235–260.
[6] W.A. Brock, P. de Fountnouvelle, Expectational diversity in monetary economics, J. Econom. Dynamics Control 24 (2000) 725–759.
[7] W.A. Brock, C.H. Hommes, A rational route to randomness, Econometrica 65 (1997) 1059–1160.
[8] W.A. Brock, C.H. Hommes, Heterogeneous beliefs and routes to chaos in a simple asset pricing model, J. Econom. Dynamics Control 22 (1998) 1235–1274.
[9] W.A. Brock, C.H. Hommes, Rational animal spirits, in: P.J.J. Herings, A.J.J. Talman, G. Van Der Laan (Eds.), The Theory of Markets, North-Holland, Amsterdam, 2000, pp. 109–137.
[10] W.A. Brock, C.H. Hommes, F.O.O. Wagener, Evolutionary dynamics in financial markets with many trader types, J. Math. Econom., 2001, forthcoming.
[11] G.W. Evans, S. Honkapohja, Learning and Expectations in Macroeconomics, Princeton University Press, Princeton, NJ, 2001.
[12] G.W. Evans, S. Honkapohja, R. Marimon, Monetary inflation models with heterogeneous learning rules, Macroeconom. Dynamics 5 (2001) 1–31.
[13] G.W. Evans, G. Ramey, Expectation calculation and macroeconomic dynamics, Amer. Econom. Rev. 82 (1992) 207–224.

[14] G.W. Evans, G. Ramey, Adaptive Expectations, Underparameterization and the Lucas Critique, University of Oregon, Working Paper No. 2001-8, J. Monet. Econom., forthcoming.
[15] M. Friedman, Essays in Positive Economics, University of Chicago Press, Chicago, IL, 1953.
[16] J. Haltiwanger, M. Waldman, Limited rationality and strategic complementarity: the implications for macroeconomics, Quart. J. Econom. 104 (1989) 463–483.
[17] S. Honkapohja, K. Mitra, Learning Stability in Economies with Heterogeneous Agents, Mimeo.
[18] R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge University Press, Cambridge, 1985.
[19] C.F. Manski, D. McFadden, Structural Analysis of Discrete Data with Econometric Applications, MIT Press, Cambridge, MA, 1981.
[20] T.J. Sargent, The Conquest of American Inflation, Princeton University Press, Princeton, NJ, 1999.
[21] R. Sethi, R. Franke, Behavioural heterogeneity under evolutionary pressure: macroeconomic implications of costly optimisation, Econom. J. 105 (1995) 583–600.
[22] N.L. Stokey, R.E. Lucas, Jr., (with E.C. Prescott), Recursive Methods in Economic Dynamics, Harvard University Press, Cambridge, 1989.
[23] R. Townsend, Forecasting the forecast of others, J. Polit. Economy 91 (1983) 546–588.