Information Equilibria in Dynamic Economies*

Giacomo Rondina[†] Todd B. Walker[‡]

March 1, 2010

Abstract

We establish an information equilibrium concept that provides existence and uniqueness conditions for dynamic economies with incomplete information. Our equilibrium concept overturns non-existence results once thought to be pervasive in models with non-trivial informational dynamics, and establishes a connection between hierarchical and dispersed information structures. We show that the equilibria belonging to this class are characterized by a generalization of the celebrated Hansen-Sargent formula. A ubiquitous characteristic of this generalized Hansen-Sargent formula is a propagation effect triggered by perpetual learning about structural innovations from equilibrium variables. We provide analytic characterizations of equilibrium dynamics, which permit closed-form solutions of higher order belief dynamics. We also derive an equivalence between non-fundamental moving average representations and dynamic signal extraction problems. This equivalence allows for a novel rational expectations interpretation of moving average processes.

*We would like to thank Christian Hellwig, Ken Kasa, Eric Leeper, Kristoffer Nimark, Bart Taub, Pierre-Olivier Weill, Chuck Whiteman and seminar participants at UC–Berkeley, the Econometric Society summer meetings, and the Yale University Cowles Foundation Summer Conference on Information and Beliefs in Macroeconomics for useful comments.

[†]UCSD, grondina@ucsd.edu

[‡]Indiana University, walkertb@indiana.edu

1 Introduction

In market economies, agents use diverse sources of information to set demand and supply strategies. While some sources of information are *exogenous* to the specific market under consideration, other sources, such as prices and interest rates, are *endogenous* in that the information is generated as a by-product of the functioning of market forces. In this paper, we study rational expectations equilibria with competitive markets where agents have access to both exogenous and endogenous sources of information. We refer to these equilibria as "information equilibria."

Since the rational expectations revolution, solving for equilibria in dynamic macroeconomic models relies on imposing the restriction that equilibrium dynamics are a function of expectations of stochastic variables. Such expectations, in turn, determine a mapping from exogenous stochastic processes to endogenous variables. This mapping is often referred to as cross-equation restrictions, which are the "hallmark of rational expectations models," Sargent (1981). In this paper, we examine the cross-equation restrictions arising from a class of dynamic rational expectations models containing informational frictions, and derive conditions for which the informational friction persists in equilibrium. We show that the determination of the mapping between equilibrium variables and exogenous shocks, and the resulting cross-equation restrictions, are crucially affected by the interaction between exogenous *and* endogenous information.

The novelty of our results arises in part because the informational assumptions typically imposed on dynamic models do not nest incomplete information. Consider the following equilibrium equation for a standard speculative market

$$p_t = \beta \mathbb{E} \left(p_{t+1} | \Omega_t \right) + s_t, \tag{1.1}$$

where s_t is assumed to be exogenous, p_t endogenous, and $|\beta| < 1$. Suppose that there is a proportion of "informed" agents who are endowed with current and past structural innovations to s_t, say ε_t, so that $\Omega_t = \varepsilon^t \equiv \langle \varepsilon_t, \varepsilon_{t-1}, ... \rangle$, and a proportion of "uninformed" rational agents who only observe the history of equilibrium outcomes p^t, but not the history of structural innovations ε^t directly. Suppose also that s_t follows an autoregressive, moving average process of order one, ARMA$(1, 1)$

$$s_t = \rho s_{t-1} + \varepsilon_t + \theta \varepsilon_{t-1}. \tag{1.2}$$

The typical assumption is to set $\theta = 0$, and for s_t to follow a purely autoregressive process. An important result derived below is that agents observing endogenous information only, p^t generated by (1.1), will *always* be able to recover ε^t if $\theta = 0$. A straightforward and convenient implication of this property is that one can abstract from exogenous informational differences altogether. The downside, however, is that there exists a set of interesting "information equilibria" that are disregarded. This insight extends to more complex dynamic models where, even under the standard AR(1) assumption for the exogenous shocks, interesting information equilibria are usually overlooked.

In this paper, we characterize these equilibria and show that their propagation properties can be dramatically different from the fully-informed equilibrium. The paper makes the following contributions.

First, we provide a novel informational interpretation for moving average (MA) representations within the context of rational expectations models. Incomplete information models of several types are shown to produce endogenous variables with non-fundamental MA representations. As noted above, focusing solely on autoregressive components misses entirely this set of equilibria. We facilitate the informational interpretation of MA representations by providing an "informational equivalence" between non-fundamental MA representations and the more familiar dynamic signal extraction problem.

Second, we establish an information equilibrium concept that provides existence and uniqueness conditions for dynamic economies with incomplete information. The defining characteristic of information equilibria is a fixed point condition in information. Solving for the fixed point condition is tantamount to identifying which linear combination of structural shocks agents are able to infer from endogenous and exogenous variables. We demonstrate how rationality and common knowledge of rationality delivers an *additional* linear combination of innovations beyond that contained in endogenous and exogenous variables. Accounting for this additional restriction overturns non-existence results once thought to be pervasive in models with non-trivial informational dynamics [Futia (1981)].

Third, we *analytically* characterize the space of information equilibria in both the symmetric and hierarchical setup of Futia (1981) and a dispersed informational setup where each agent is equally uninformed. The assumption of rationality in dynamic models with incomplete information leads naturally to agents forming higher-order beliefs [Townsend (1983)] and signal extraction from endogenous variables [Sargent (1991)]. These characteristics make solving and characterizing information equilibria using traditional state space, recursive methods challenging, and has resulted in numerical approximation of equilibria. We take advantage of a version of the powerful Riesz-Fischer Theorem—which provides an alternative to a state space / Kalman filter approach—to solve for the information equilibria in closed form.[1] Our solution method not only easily handles the technical difficulties associated with incomplete information, as argued by Kasa (2000) and Kasa *et al.* (2008), but also permits generalized conditions for existence and uniqueness, as demonstrated by Whiteman (1983). Therefore we are able to examine the robustness of the information equilibrium to perturbations in parameter values and informational distributions, and for various stochastic processes.

Finally, we examine the dynamic properties of models with incomplete information. We show that the equilibria belonging to this class are characterized by a generalization of the celebrated Hansen and Sargent (1980) formula. The ubiquitous characteristic of this generalized Hansen-Sargent formula is a propagation effect triggered by *perpetual* learning about structural innovations from equilibrium variables. One interpretation of this effect lends itself directly to the "waves of optimism and pessimism" that Pigou (1929) argued is a key source of cyclical variation in economic activity.

Another important characteristic of asymmetric information models is the failure of the law of iterated expectations and the formation of higher-order beliefs. Given that we derive an analytical solution and establish conditions of existence and uniqueness, we are able to

[1] In a companion paper Rondina and Walker (2009), we demonstrate the connection between our approach and the recursive approach, and discuss the advantages and disadvantages of both. See Sargent (1987) for a discussion of the Riesz-Fischer Theorem.

sharply characterize these aspects of the equilibrium.

2 Connection to Literature

Models of incomplete information are becoming increasingly prominent in several literatures such as asset pricing, optimal policy communication, international finance, and business cycles.[2] The role of incomplete information in many of these settings was acknowledged very early on; Keynes (1936) believed higher-order expectations played a fundamental role in asset markets, while Pigou (1929) argued that business cycles may be subject to "waves of optimism and pessimism." The idea that incomplete information could induce a propagation mechanism and contribute substantially to business cycle fluctuations was first formalized in a rational expectations setting by Lucas (1975), Townsend (1983) and King (1982). The defining characteristic of these models was asymmetrically informed agents who observed endogenous variables that did not fully reveal the structural innovations hitting the economy. This incomplete information induced forecast errors that were correlated across agents, which resulted in business cycle fluctuations that exceeded the initial aggregate shock in both size and persistence. We too find that incomplete information induces propagation and amplification of innovations.

Solving for equilibria in dynamic models with incomplete information is challenging. Bacchetta and van Wincoop (2006) attribute the lack of research following the early work of Lucas (1972), Lucas (1975), King (1982) and Townsend (1983) to the technical challenges of solving for equilibrium, even though these models harbored much potential. The primary difficulties are rational agents forming higher-order beliefs [Townsend (1983)], which makes the typical recursive state space formulation approach problematic because the state may be infinite dimensional, and signal extraction from endogenous variables [Sargent (1991)], which leads to a delicate fixed point condition in information.

Following Townsend (1983), the customary way of solving for information equilibria in dynamic models with incomplete information is to assume that the innovations are perfectly observed at some arbitrary distant point in the past.[3] This allows one to put the system in state space recursive form, which permits the use of the Kalman filter to solve for the signal extraction problem, and implicitly solves the informational fixed point condition. Rondina and Walker (2009) show that truncating the state space in this manner runs the risk of revealing the entire history of innovations up to the current period, regardless of the point of truncation. Clearly, if the model generates a signal extraction problem of the type described here, this assumption has the undesirable implication of completely removing an important informational friction from the equilibrium outcome. Hence, the approximation error associated with truncation can be quite large.[4] Our solution procedure and equilibrium concept does not rely on truncating the state and therefore does not dampen the effects from incomplete information. We also provide general restrictions which guarantee the existence

[2] The literature is too voluminous to cite every worthy paper. Recent examples include: Morris and Shin (2002), Woodford (2003), Allen *et al.* (2006), Bacchetta and van Wincoop (2006), Gregoir and Weill (2007) Angeletos and Pavan (2007), Kasa *et al.* (2008), Lorenzoni (2009), Rondina (2009), Angeletos and La'O (2009).

[3] There have been other approaches to handle these technical issues. Most notably Nimark (2007) maintains the recursive structure and employs the Kalman filter, allowing for a large, yet finite, state space.

[4] Walker (2007) demonstrates this point in the model of Singleton (1987).

and uniqueness of information equilibria by deriving the informational fixed point condition *endogenously*.

Our work most closely relates to that of Futia (1981). Futia's key insight was that, in dynamic settings, moving average representations could be used to preserve asymmetric information in equilibrium. In a simple speculative market model, Futia examined both a symmetric and a hierarchical information structure, assuming that equilibrium prices convey information to price-taking investors. Using analytic function methods to solve for equilibrium linear pricing functions, Futia derived non-existence conditions for a symmetric information equilibrium, where the non-existence of equilibria was attributed to the endogeneity of information.[5] The non-existence result of Futia (1981) has since been regarded as a problematic feature of endogenous information equilibria in rational expectations models. We extend Futia in three directions.

i. We derive general existence conditions that are consistent with Futia's hierarchical information example, but diverge from Futia's symmetric case in the sense that we show that Futia's non-existence result disappears. We argue that this discrepancy can be attributed to rationality and common knowledge of rationality. We refer to this concept as "knowledge of the model" and argue that it plays a crucial role in characterizing the space of any information equilibria.

ii. We introduce a dispersed information structure and show how the equilibrium properties in this case are related to the hierarchical information equilibrium through a simple reinterpretation of a key informational parameter.

iii. We show that the equilibrium characterization can be interpreted as a "generalized" Hansen-Sargent formula, which facilitates the comparison to the typical representative agent rational expectations model. We also derive and examine important properties of the information equilibria (i.e. higher-order beliefs) in closed form.

Our work is also related to Kasa *et al.* (2008) (KWW). KWW examine Futia's speculative market model under a symmetric, heterogeneous information structure and derive conditions under which information remains heterogeneous in equilibrium. Once appropriately reinterpreted (taking into account the differences in information structure), the existence conditions for an heterogeneous information equilibrium in Futia (1981), KWW (2008) and the results derived herein can be shown to be consistent. KWW then study how the stochastic properties of such equilibrium can help in understanding the empirical properties of asset prices.

3 Information Equilibrium: Preliminaries

3.1 Equilibrium To fix notation and ideas, we define an information equilibrium within the model of Futia (1981). We work within this framework to juxtapose our definition of equilibrium to that of Futia's and to allow a broad range of interpretations. Futia assumed stochastic "market fundamentals" (s_t), which he interpreted as a speculative component of

[5] Taub (1989), Kasa (2000), Walker (2007), Kasa *et al.* (2008) and Rondina (2009) also use the space of analytic functions to characterize equilibrium in models with informational frictions. Seiler and Taub (2008), Bernhardt and Taub (2008), and Bernhardt *et al.* (2009) show how these methods can be used to accurately approximate asymmetric information equilibria in models with richer specifications of information.

supply. Agents are risk neutral and discount the future at rate β. For now, we assume a continuum of asymmetrically informed agents indexed by i. The model is given by

$$p_t = \beta \int_0^1 \mathbb{E}_t^i p_{t+1} di + s_t \tag{3.1}$$

where $\mathbb{E}_t^i$ is the conditional expectation of agent i.

The exogenous process (s_t) is driven by a Gaussian shock

$$s_t = A(L)\varepsilon_t, \qquad \varepsilon_t \overset{iid}{\sim} N(0, \sigma_\varepsilon^2) \tag{3.2}$$

where $A(L)$ is a square summable polynomial in the lag operator L.

Information is assumed to originate from two sources–exogenous and endogenous. Exogenous information, denoted by U_t^i, is that which is not affected by market forces. This dimension of information must be endowed by the modeler. Endogenous information is generated through market interactions. When agents are asymmetrically informed, endogenous variables may convey additional information not contained in the exogenous information set. We separate endogenous information into two components–$\mathbb{V}_t(p)$ and $\mathbb{M}_t(p)$. The notation $\mathbb{V}_t(p)$ denotes the smallest closed subspace that is spanned by current and past p_t and $\mathbb{M}_t(p)$ embeds the assumption that agents know the equilibrium process p_t evolves according to (3.1). This distinction is important and elaborated on below. The time t information of trader i is then $\Omega_t^i = U_t^i \vee \mathbb{V}_t(p) \vee \mathbb{M}_t(p)$, where the operator $\vee$ denotes the span (i.e., the smallest closed subspace which contains the subspaces) of the U_t^i, $\mathbb{V}_t(p)$ and $\mathbb{M}_t(p)$ spaces. If the exogenous and endogenous information are disjoint, then the linear span becomes a direct sum. We use similar notation as Futia (1981) in that $\mathbb{V}_t(x) = \mathbb{V}_t(y)$ means the space spanned by $\{x_{t-j}\}_{j=0}^{\infty}$ is equivalent, in mean square, to the space spanned by $\{y_{t-j}\}_{j=0}^{\infty}$.

Uncertainty is assumed to be driven entirely by the Gaussian stochastic process ε_t, which rules out sunspots and implies the equilibrium lies in a well-known Hilbert space (the space spanned by square-summable linear combinations of ε_t). Normality implies that optimal projection formulas are equivalent to conditional expectations,

$$\mathbb{E}_t^i(p_{t+1}) = \Pi[p_{t+1}|\Omega_t^i] = \Pi[p_{t+1}|U_t^i \vee \mathbb{V}_t(p) \vee \mathbb{M}_t(p)]. \tag{3.3}$$

where Π denotes linear projection. We now define an information equilibrium.

Definition IE. *An Information Equilibrium (IE) is a stochastic process for $\{p_t\}$ and a stochastic process for the information sets $\{\Omega_t^i, i \in [0,1]\}$ such that: (i) each agent i, given the price and the information set, follows an optimal strategy and forms expectations according to (3.3); (ii) p_t satisfies the equilibrium condition (3.1).*

An IE consists of two objects: a *price* and a *distribution of information*. The two objects are both endogenously and simultaneously determined in equilibrium. An IE can be summarized by two statements: (a) given a distribution of information sets, there exists a market clearing price determined by each agent i's optimal prediction conditional on the information sets; (b) given a price process, there exists a distribution of information sets generated by the price process that provides the basis for optimal prediction. Both statements (a) and (b) must be satisfied by the same price and the same distribution of information simultaneously in order to satisfy the requirements of an IE.

3.2 Information In dynamic settings, agents continually collect new observations, but what is crucial is how much the current and past observations reveal about the structural innovations—ε_{t-j}. One of the key contributions of this paper with respect to the existing literature on information asymmetries is the emphasis on a class of signal extraction that pertains only to *dynamic* settings. Even without exogenous and superimposed noise, a general dynamic structure exists such that the effects of a contemporaneous shock ε_t cannot be parsed out from past realizations ε_{t-1}, ε_{t-2},....

As an example, consider the problem of extracting information about ε_t from

$$x_t = \varepsilon_t + \theta\varepsilon_{t-1}. \tag{3.4}$$

If we assume $|\theta| \leq 1$, then there exists a linear combination of current and past x_t's that allows the exact recovery of ε_t. This is

$$\mathbb{E}_{|\theta|\leq 1}\left(\varepsilon_t|x^t\right) = \left(x_t - \theta x_{t-1} + \theta^2 x_{t-2} - \theta^3 x_{t-3} + ...\right) = \varepsilon_t. \tag{3.5}$$

Note that the infinite sum converges as θ^n goes to zero for n "big enough." In this case the x_t process spans the same informational space as the ε_t process, $[\mathbb{V}_t(x) \equiv \mathbb{V}_t(\varepsilon)]$.

When $|\theta| > 1$ information is partially lost. Obviously, (3.5) is no longer well defined as the coefficients for the past realizations of x_t grow without bound. Nevertheless, there is still a linear combination of x_t that minimizes the forecast error for ε_t; this is given by

$$\mathbb{E}_{|\theta|>1}\left(\varepsilon_t|x^t\right) = \theta\left(x_t - \theta^{-1}x_{t-1} + \theta^{-2}x_{t-2} - \theta^{-3}x_{t-3} + ...\right) = \tilde{\varepsilon}_t. \tag{3.6}$$

In this case, the representation for x_t spans $\tilde{\varepsilon}_t$, $[\mathbb{V}_t(x) \equiv \mathbb{V}_t(\tilde{\varepsilon})]$. $\tilde{\varepsilon}_t$ contains strictly less information than ε_t, in the sense that the mean squared forecast error conditional on $\tilde{\varepsilon}_t$ is bigger than ε_t (which is identically zero). More specifically,

$$\mathbb{E}\left[(\varepsilon_t - \tilde{\varepsilon}_t)^2\right] = \left(1 - \frac{1}{\theta^2}\right)\sigma_\varepsilon^2 > 0.$$

The forecast error approaches zero as $|\theta| \to 1$; that is, the information in $\tilde{\varepsilon}_t$ approaches the full information ε_t.

A dynamic interpretation of this informational imperfection is to imagine an original state of the world that is imperfectly observed at $t = 0$ (e.g. ε_{-1}). If $|\theta| \leq 1$, the forecast error associated with this "original ignorance" approaches zero over time. The dynamics of the signal do not prevent the agents from learning the correct state of the world as the sample size increases. Conversely, when $|\theta| > 1$, the same original ignorance never diminishes and agents never perfectly recover ε_t, at any arbitrary time in the future.

It is useful at this point to establish a connection between the information contained in $\tilde{\varepsilon}_t$ when $|\theta| > 1$ and a dynamic signal extraction problem cast in a more familiar setting. Suppose that agents observe an infinite history of the signal

$$z_t = \varepsilon_t + \eta_t, \tag{3.7}$$

where $\eta_t \overset{iid}{\sim} N\left(0, \sigma_\eta^2\right)$. The optimal prediction is well known and given by $\mathbb{E}(\varepsilon_t|z^t) = \frac{\tau}{1+\tau}z_t$, where τ is the standard signal-to-noise ratio parameter $\tau = \sigma_\varepsilon^2/\sigma_\eta^2$.

Rondina and Walker (2009) show the equivalence (in terms of the variance of the prediction errors) between the signal extraction problem (3.4) when $|\theta| > 1$ and (3.7) is equivalent when

$$\theta^2 = 1 + \frac{1}{\tau}. \tag{3.8}$$

where equivalence is defined as equality of variance of the forecast error conditioned on the infinite history of the observed signal, i.e.

$$\mathbb{E}\left[\left(\varepsilon_t - \mathbb{E}_{|\theta|>1}\left(\varepsilon_t|x^t\right)\right)^2\right] = \mathbb{E}\left[\left(\varepsilon_t - \mathbb{E}\left(\varepsilon_t|z^t\right)\right)^2\right],$$

This suggests that agents, concerned with minimizing the variance of forecast errors, should be indifferent between receiving the signal x_t or z_t, assuming that (3.8) is satisfied. However from a positive point of view, there are important differences. For example, the impulse responses of x_t and z_t to an innovation in ε_t are far from similar. For z_t, the prediction formula reacts at impact according to the magnitude $\frac{\tau}{1+\tau}$, but is zero thereafter. Conversely, the response of x_t extends for many periods beyond impact. To see this more clearly, rewrite (3.6) as

$$\begin{aligned}\tilde{\varepsilon}_t &= \underbrace{\theta^{-1}} \quad \underbrace{\varepsilon_t} + \underbrace{(1-\theta^{-2})}\underbrace{[\varepsilon_{t-1} - \theta^{-1}\varepsilon_{t-2} + \theta^{-2}\varepsilon_{t-3} - \cdots]}. \\ &= \text{weight} \times \text{signal} + \quad \text{weight} \quad \times \quad \text{noise}\end{aligned} \tag{3.9}$$

This equation demonstrates how θ^{-1} controls the information that $\tilde{\varepsilon}_t$ contains about ε_t through two channels–a signal with weight θ^{-1}, and a noise component with weight $(1-\theta^{-2})$. As θ increases there are three effects. First, the weight on the signal decreases and x_t contains less information about ε_t. Second, the weight on the noise increases, but this is offset (somewhat) by the third effect, which is a reduction in the noise associated with innovations dated $t-2$ and earlier $(\varepsilon_{t-2}, \varepsilon_{t-3}, \ldots)$. In the limit as $\theta \to \infty$, the distribution of $\tilde{\varepsilon}_t$ is degenerate at ε_{t-1} so that the best prediction for ε_t is last period's realization ε_{t-1}.[6] As $|\theta|$ approaches 1, these three effects are reversed and past realizations contribute substantially to the noise while the weight on the noise decreases.

Figure 1 plots the impulse response functions for $\mathbb{E}(\varepsilon_t|z^t)$ and $\mathbb{E}_{|\theta|>1}(\varepsilon_t|x^t)$. For both values of τ, the reaction at impact of the impulse responses are the same for x_t and z_t. However, the impulse response for prediction conditional on z_t dies out immediately after impact, while the impulse response for prediction based on x_t exhibits the interesting dynamics characterized by (3.9). For the high signal case ($\tau = 1$, blue line), x_t declines after impact, mimicking the actual behavior of ε_t (the black line); then subsequently overshoots zero and continues to oscillate for many periods. For the low signal case ($\tau = .1$, red line), x_t initially increases at impact, producing a hump-shaped response of the same order of magnitude as the initial innovation. The impulse then overshoots and oscillates for only a few periods. These dynamics continually resurface in the models studied below.

[6] Note that the mean squared forecast error in this limiting case is equal to σ_ε^2, which is what one would get with an unconditional forecast. Agents prefer to use (3.9) rather than the unconditional forecast, $\mathbb{E}(\varepsilon_t) = 0$, because the resulting mean squared forecast error is smaller than σ_ε^2 for finite values of θ.

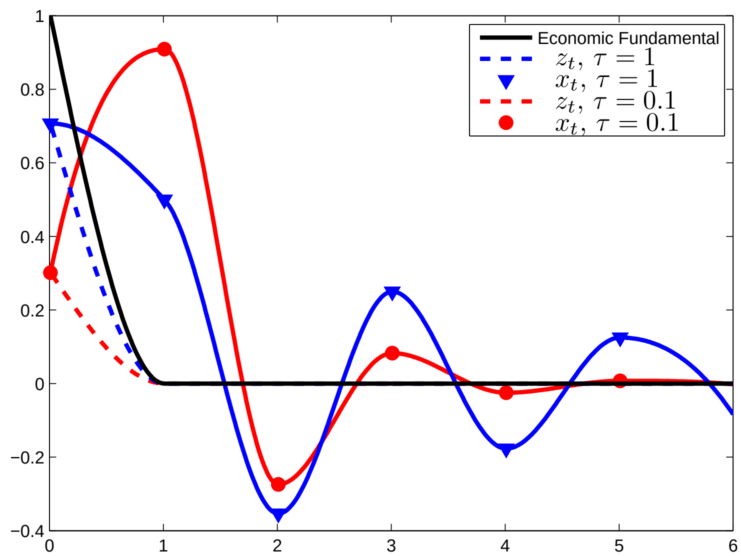

Figure 1: Impulse Responses of x_t and z_t to a one unit change in ε_t for signal-to-noise ratios of $\tau = 1$, $\theta = \sqrt{2}$ (dashed, triangle markers) and $\tau = 0.1$, $\theta = \sqrt{11}$ (dashed-dot, circle markers). The solid line is the innovation itself, ε_t, normalized to 1.

3.3 Solution Procedure As stated in Definition 3.1, an Information Equilibrium (IE) must satisfy two conditions. Given a distribution of information sets, there exists a market clearing price determined by each agent's optimal prediction. Given a price process, there exists a distribution of information generated by the price that provides the basis for optimal prediction. Our solution procedure described here uses a recursion in the space of analytic functions to solve for the fixed point conditions, initiated with a candidate solution that is "minimal" from an informational point of view.

In solving for the fixed point conditions of Definition 3.1, we use the analytic function approach advocated by Futia (1981), Whiteman (1983), Taub (1989) and Kasa (2000). The analytic function approach is particularly useful when solving for an informational fixed point because the information encoded into endogenous equilibrium variables can be easily detected by the behavior of the analytic functional representation of these variables in the complex plane. As shown above, the property of invertibility of an analytic function inside the complex unit circle informs about the existence of a linear combination of observed information that reveals the state of the system. If perfect recovery is not possible, the invertible representation delivers the precise mapping into the smaller space that is revealed.

The solution procedure involves two steps: [i] guess a candidate solution that is minimal with respect to information and impose equilibrium conditions [ii] check the invertibility of the endogenous variables to ensure the informational fixed point condition holds. Through market interactions, the information conveyed by the candidate solution may be larger than the initial information set of step [i]. If this is the case, the new enlarged information set is used to generate a new candidate solution, and the process is repeated until convergence. Since the expansion of the information set is bounded above by the full information benchmark, the iteration is sure to converge.

A critical component of the solution procedure is initializing the recursion in information.

In principle, for any specification of the exogenous information structure $\{U_t^i\}$ there might exist a unique convergence point in the recursion described above. It is useful to think of IE as a set of equilibria where each equilibrium is indexed by the assumed exogenous information structure. It will never be an empty set as one can always assume exogenous information such that the equilibrium is fully revealing.[7]

In initializing the recursion in information, we follow the spirit of Radner (1979), who advocated forming an "exogenous information equilibrium" as an initial guess for the IE. The exogenous information equilibrium assumes agents are only able to condition on exogenous information, which places a lower-bound restriction on the initial condition for information. Radner argued that such an equilibrium would persist only if every agent remained unsophisticated and ignored the information coming from the model. A dynamic interpretation of Radner is to say that a "sophisticated" agent acting rationally will not generate forecast errors that are serially correlated *with respect to their own information sets.* As we see below, this does not preclude the possibility that agent i's forecast errors will be serially correlated with respect to agent j's information set.

The machinery of the previous section can now be used to generate a initial guess for the equilibrium that is consistent with Radner's motivation. For the models described below, this guess is given by

$$p_t = Q(L)\prod_{i=1}^{n}(L-\lambda_i)\varepsilon_t \tag{3.10}$$

where $|\lambda_i| < 1$ for all i is assumed, and $Q(L)$ is assumed to contain no zeros inside the unit circle.[8] As noted in Section 3.2, if $|\lambda_i| < 1$ for all i, conditioning on the price process implies agents will not be able to infer $\{\varepsilon_{t-j}\}_{j=0}^{\infty}$ perfectly. This guess for the price process spans the space of $\tilde{\varepsilon}_t$, where now

$$\tilde{\varepsilon}_t = \mathcal{B}_{\lambda_1}(L)\mathcal{B}_{\lambda_2}(L)\cdots\mathcal{B}_{\lambda_n}(L)\varepsilon_t \tag{3.11}$$

and $\mathcal{B}_{\lambda_i}(L) = |\lambda_i|(L-\lambda_i)/\lambda_i(1-\lambda_i L)$, i.e. a product of Blaschke factors must be used to derive the information set of the agents [see Hansen and Sargent (1991), Lippi and Reichlin (1994)]. Using the tools of previous section, it is easy to show that for every additional $|\lambda_i|$ inside the unit circle, the conditioning set $\mathbb{V}_t(p)$ contains strictly *less* information than before. As our solution procedure permits the number of zeros inside the unit circle (n) to be arbitrarily large, (3.10) represents an "informational lower bound." As we show for the specific informational assumptions that follow, our solution procedure endogenizes λ_i and $Q(L)$ and therefore is very flexible in that it does not impose $|\lambda_i| < 1$ for any i in equilibrium. Thus the candidate equilibrium is minimal from an informational standpoint but allows for a potentially larger information set endogenously.

[7] In Rondina and Walker (2009) we show that the choice of casting the solution to an IE in a recursive fashion (i.e. specifying a state representation for the problem) is subject to the risk of implicitly initializing the agents' information set in a way that excludes entire classes of information equilibria.

[8] In what follows, we also impose the restriction that $\lambda_i \neq \beta$ for all i. This restriction, along with $|\beta| < 1$, ensures existence and uniqueness of the equilibrium, for a given exogenous information structure.

4 Symmetric Information

Consider model (3.1) where the heterogenous beliefs collapse to a common knowledge, symmetric information structure,

$$p_t = \beta \mathbb{E}_t(p_{t+1}) + s_t. \tag{4.1}$$

Symmetric information implies the law of iterated expectations holds and the above difference equation may be written as the contemporaneous expectation of the discounted sum of future s_t's,

$$p_t = \sum_{j=0}^{\infty} \beta^j \mathbb{E}_t(s_{t+j}). \tag{4.2}$$

4.1 Existence of Information Equilibria It is useful to establish a benchmark equilibrium solution to the above equation. Assume that the exogenous information provided to the agents is the full knowledge of the innovations up to time t, i.e.

$$U_t^i = \mathbb{V}_t(\varepsilon), \forall i \in [0,1]\,. \tag{4.3}$$

Here, and in the following analysis, we assume that agents observe the endogenous information $\mathbb{V}_t(p) \vee \mathbb{M}_t(p)$. In lieu of characterizing each term in the summation, we posit that the solution to (4.2) has the functional form $p_t^\varepsilon = P^\varepsilon(L)\varepsilon_t$, where $P^\varepsilon(L)$ satisfies square summability. Expectations are given by the Wiener-Kolmogorov optimal prediction formula, $\mathbb{E}[p_{t+1}^\varepsilon|\mathbb{V}_t(\varepsilon)] = L^{-1}[P^\varepsilon(L) - P_0^\varepsilon]\varepsilon_t$ which follows from our assumption that agents have knowledge of current and past innovations, $\{\varepsilon_{t-j}\}_{j=0}^{\infty}$. Substituting the expectation into the equilibrium yields a functional equation for p_t. As noted above, we solve for the functional fixed point problem in the space of analytic functions. The z-transform of the p_t process may be written as

$$P^\varepsilon(z) = \frac{zA(z) - \beta P_0^\varepsilon}{z - \beta}. \tag{4.4}$$

The z-transform must be analytic in the frequency domain, which is tantamount to square summability in the time domain. If $|\beta| \geq 1$, then (4.4) is analytic and the free parameter P_0^ε can be set arbitrarily. Uniqueness, then, requires $|\beta| < 1$, in which case the free parameter P_0^ε is set to ensure the function is analytic for $|z| < 1$.[9] The equilibrium is then characterized by

$$p_t = \left(\frac{LA(L) - \beta A(\beta)}{L - \beta}\right)\varepsilon_t \tag{4.5}$$

which is the celebrated Hansen-Sargent formula [Hansen and Sargent (1991)]. This characterization of a rational expectations equilibrium is not controversial and can be obtained through many different solution procedures [*e.g.*, Blanchard and Kahn (1980), Sims (2002)].

[9]See Futia (1981) and Whiteman (1983) for more on solving rational expectations models using z-transform techniques.

As noted by Hansen and Sargent, this equation clearly captures the cross-equation restrictions, which are the "hallmark" of rational expectations models. It is the *unique* solution to (4.2) when information is specified as (4.3). However, the definition of an IE does not rule out the existence of other equilibria when the exogenous information is specified differently. The relevant questions are then: under what exogenous informational assumption does (4.5) represent an IE? What are the characteristics of an IE when the exogenous information does not support (4.5) as an equilibrium?

In contrast to (4.3), assume exogenous information is given by the empty set $U_t^i = \{0\}$ $\forall i$. Hence, all information is coming from current and past observations of the endogenous variable $\mathbb{V}_t(p)$ and knowledge that this endogenous variable is generated by (4.1), $\mathbb{M}_t(p)$. We focus on this exogenous information for two reasons. First, signal extraction from endogenous variables in a dynamic asymmetric information setting is nontrivial, and this section lays the groundwork for that case. Second, this is the same informational setup as Futia (1981); however, we overturn the non-existence pathologies derived therein.

The solution procedure outlined above advocates forming an initial guess of the endogenous variable that is minimal with respect to information. This "informational lower bound" may be achieved by assuming (3.10) as the guess for the exogenous information equilibrium. The conditional expectation may be derived by applying the Wiener-Kolmogorov optimal prediction formula to (3.10) conditional on observing $\{\tilde{\varepsilon}_{t-j}\}_{j=0}^{\infty}$ given by (3.11). This conditional expectation is[10]

$$\mathbb{E}[p_{t+1}|\mathbb{V}_t(\tilde{\varepsilon})] = L^{-1}[Q(L)\prod_{i=1}^{n}(1-\lambda_i L) - Q_0]\tilde{\varepsilon}_t. \tag{4.6}$$

Substituting this into (4.1) and solving the model in the space of analytic functions yields the following theorem.

Theorem 1. *Under the exogenous information assumption $U_t^i = \{0\}$ $\forall i$, a unique Information Equilibrium for (4.1) with $|\beta| < 1$ always exists and is determined as follows: let $\{|\lambda_i| < 1\}_{i=1}^n$ be a collection of real numbers such that*

$$A(\lambda_i) = 0, \tag{4.7}$$

then the information equilibrium price process is

$$p_t = Q(L)\prod_{i=1}^{n}(L-\lambda_i)\varepsilon_t = \frac{1}{L-\beta}\left\{LA(L) - \beta A(\beta)\frac{\prod_{i=1}^{n}\mathcal{B}_{\lambda_i}(L)}{\prod_{i=1}^{n}\mathcal{B}_{\lambda_i}(\beta)}\right\}\varepsilon_t \tag{4.8}$$

where

$$\mathcal{B}_{\lambda_i}(L) = \frac{L-\lambda_i}{1-\lambda_i L}.$$

If condition (4.7) does not hold for any $|\lambda_i| < 1$, the IE is given by (4.5).

[10]Kasa *et al.* (2008) emphasize the conditioning down onto the smaller subspace $\tilde{\varepsilon}_t$ in the conditional expectation. We show that this conditioning down also applies to the equilibrium characterization and takes the form of (4.8).

Proof. See Appendix A. □

As will be emphasized throughout, an IE consists of both an information set and a price process. The statement of Theorem 1 highlights this duality by requiring an IE to satisfy two conditions–(4.7) and (4.8). Restriction (4.7) states that the initial guess of the price (3.10), will only be an IE if $A(\lambda_i) = 0$ holds for every $i = 1, ...n$. Restriction (4.7), then, determines the exact number of $|\lambda_i| < 1$ for $i = 1, ...n$ that exist in equilibrium, and hence determines the endogenous information available to the agents. This restriction stipulates that the exogenous process, s_t, must vanish when evaluated at each of the λ's given by (4.8). Given that the s_t and p_t process share common zeros (λ), the information content of s_t must be equal to that of p_t.

The intuition behind this result is best understood by distinguishing between information generated by observing the price sequence or "time series information" of p_t ($\mathbb{V}_t(p)$), and information generated by the model or "equilibrium information" of p_t ($\mathbb{M}_t(p)$). Knowledge of the time series properties of p_t is straightforward, while information generated by the model is a more subtle concept. In the symmetric information case, knowledge of the model implies that agents know that all other agents are similarly informed. This common knowledge suggests that, whatever the time series properties of the equilibrium process are, the mere fact that equilibrium holds implies

$$p_t - \beta \mathbb{E}_t(p_{t+1}) = s_t.$$

Therefore, the minimal information agents receive in equilibrium is that which is generated by the exogenous process, s_t (i.e., $\mathbb{M}_t(p) = \mathbb{V}_t(s)$). What is important is that this relationship holds no matter the process for p_t, so long as a unique equilibrium exists. It is in this sense that the model–and common knowledge of the model–places significant structure on the information sets of agents, and provides agents with an *additional* linear combination of structural shocks beyond that given by the price sequence alone. The initial guess assumed n zeros inside the unit circle without specifying the specific value of those zeros. Unless the supply process contained the same number of zeros in exactly the same location, knowledge of the model would reveal additional information to the agents.

This distinction is important because if one fails to recognize the information generated by the model when imposing common knowledge, one could end up wrongly concluding that information equilibria may not exist. Non-existence pathologies of this type emerge in Futia (1981). Corollary 3.16 of Futia argues that a necessary and sufficient condition for the existence of an IE when agents are symmetrically informed is $\mathbb{V}_t(p) = \mathbb{V}_t(s)$. Futia provides an example where $s_t = (1 + \theta L)\, \varepsilon_t$ with $\theta = 5/8$ and $\beta \approx 1$. This parameter setting implies that the p_t spans a strictly smaller space than ε_t, while s_t spans the space of ε_t ($\mathbb{V}_t(p) \subset \mathbb{V}_t(s) = \mathbb{V}_t(\varepsilon)$). Futia argues that no rational expectations equilibrium exists for this parameter setting. This notion of an information equilibrium, however, ignores the information being generated by the model. In fact, observing the equilibrium process p_t and *knowing* that it is generated by (4.1) implies knowledge of $\mathbb{V}_t(s)$ by construction–$\mathbb{V}_t(p) \vee \mathbb{M}_t(p) = \mathbb{V}_t(s)$. This means that where Futia thought an IE did not exist, an IE does exist and is equal to (4.5), the fully revealing equilibrium.

More generally, our results show that an IE will *always* exist for (4.1) given $|\beta| < 1$, provided that one looks for it in the appropriate space. Representation (4.5) is the *unique* equilibrium that resides in $\mathbb{V}_t(\varepsilon)$, while (4.8) is the *unique* equilibrium residing in $\mathbb{V}_t(\tilde{\varepsilon})$. The exogenous informational assumption $\{U_t^i\}$ delivers uniqueness, and hence there are no issues with multiplicity. However, without a precise definition of an IE, it would difficult to distinguish between the two.

4.2 Characterization of Information Equilibria Given that existence and uniqueness has been established, we now examine the properties of the IE of Theorem 1. Notice that while the IE given by (4.5) is the "typical" Hansen-Sargent formula, (4.8) is a modified version of the same formula.[11] There is an informational interpretation to the Hansen-Sargent formula which applies to both equations. The first component of (4.5) and (4.8) is the perfect foresight equilibrium,

$$p_t^{pf} = \sum_{j=0}^{\infty} \beta^j s_{t+j} = \frac{LA(L)}{L-\beta}\varepsilon_t \tag{4.9}$$

This is the IE that would emerge if agents knew current, past *and future* values of ε_t.

The second component represents what must be subtracted off from the perfect foresight equilibrium because future values of ε_t are not known at t. In other words, the second component isolates the conditioning down that corresponds with the agents' information set. When agents observe current and past ε_t, this conditioning down amounts to subtracting off a particular linear combination of *future* values of ε_t. Appendix A of Hansen and Sargent (1980) shows that this component is given by the principal part of the Laurent series expansion of $A(z)$ around β, specifically

$$\beta A(\beta) \sum_{j=1}^{\infty} \beta^j \varepsilon_{t+j} \tag{4.10}$$

In the modified Hansen-Sargent formula (4.8), the conditioning down amounts to subtracting off the usual component (4.10) plus a specific linear combination of *past* values of ε_t determined by λ_i. Assuming $n = 1$ and using partial fractions yields the combination of future and past ε_t's that must be subtracted from the perfect foresight price,

$$\beta A(\beta)\left(\sum_{j=1}^{\infty} \beta^j \varepsilon_{t+j} + \frac{1-\lambda^2}{\beta-\lambda}\sum_{j=0}^{\infty} \lambda^j \varepsilon_{t-j}\right) \tag{4.11}$$

The second component on the RHS is the exact linear combination of past ε_t's that the agents do not observe. The denominator of this term, $(\beta-\lambda)$, cancels given restriction (4.7) and we get the noise term of (3.9). Hence the intuition laid out in Section 3.2 applies here.

[11] The IE given by (4.8) also nests the sticky information setup of Mankiw and Reis (2002) ($\lambda_i = 0$). However, the structural interpretation of our setup is quite different. "Inattentiveness" relies on an assumption that agents do not fully incorporate widely-available macroeconomic data into, say, price setting decisions. Our approach allows for a reinterpretation of this behavioral assumption in that our agents are acting rationally but are unable to infer the true innovations hitting the economy. Our approach also allows for more flexibility in the degree of uncertainty.

Characterizing the equilibrium in $\tilde{\varepsilon}_t$ space reveals the more familiar Hansen-Sargent representation. Assuming $n = 1$, (4.7)–(4.8) may be written in $\tilde{\varepsilon}_t$ space as

$$\tilde{s}_t = (1 - \lambda L)\hat{A}(L)\tilde{\varepsilon}_t \tag{4.12}$$

$$p_t = \left(\frac{L(1 - \lambda L)\hat{A}(L) - \beta(1 - \lambda\beta)\hat{A}(\beta)}{L - \beta}\right)\tilde{\varepsilon}_t \tag{4.13}$$

where $\hat{A}(L)$ is assumed to have no zeros inside the unit circle. This representation of equilibrium bears the more familiar features of the Hansen-Sargent formula. Because the Hansen-Sargent formula is derived from optimality conditions, a version of this formula must hold in the space spanned by the agents' conditioning set. Theorem 1 establishes conditions for which the space of existence is $\tilde{\varepsilon}_t$, and (4.12)–(4.13) demonstrates the existence of the Hansen-Sargent formula in this space. This representation of equilibrium delivers exactly the uninformed agents' belief of market fundamentals: $\tilde{s}_t = (1 - \lambda L)\hat{A}(L)\tilde{\varepsilon}_t$.

We conclude this section with a specific example that puts more structure on the general results already derived. Assume that the moving-average representation of s_t is given by

$$s_t = \rho s_{t-1} + \varepsilon_t + \theta\varepsilon_{t-1}, \qquad |\rho| < 1. \tag{4.14}$$

According to Theorem 1 the type of IE encountered hinges upon the whether s_t spans the space of ε_t, which is determined solely by θ. If $|\theta| < 1$, then the s_t process spans ε_t. In this case, the information equilibrium is obtained by plugging (4.14) into (4.5), which yields

$$p_t - \rho p_{t-1} = \left(\frac{1 + \theta\beta}{1 - \rho\beta}\right)\varepsilon_t + \theta\varepsilon_{t-1}. \tag{4.15}$$

If $|\theta| > 1$, then the specification of the exogenous information given to the agents is crucial. If we assume $U_t^i = \mathbb{V}_t(\varepsilon), \forall i$, then the IE would be equal to (4.15). However, if we maintain the assumption that $U_t^i = 0, \forall i$, the IE is found by plugging (4.14) into (4.8),

$$\tilde{p}_t - \rho\tilde{p}_{t-1} = \left(\frac{\theta + \beta}{1 - \rho\beta}\right)\tilde{\varepsilon}_t + \tilde{\varepsilon}_{t-1} = \left(\frac{1 + \theta L}{L + \theta}\right)\left[\left(\frac{\theta + \beta}{1 - \rho\beta}\right)\varepsilon_t + \varepsilon_{t-1}\right] \tag{4.16}$$

where the second equality shows the mapping into ε-space.

Notice also that the way the agents discount the news or innovations is different across the two equilibria. Equation (4.14) shows that last period's innovation, ε_{t-1}, receives a smaller discount than the contemporaneous innovation, ε_t, when $|\theta| > 1$. The opposite is true for the equilibrium that lies in $\tilde{\varepsilon}_t$ space. Here the agents believe the exogenous process is given by

$$\tilde{s}_t = \rho\tilde{s}_{t-1} + \theta\tilde{\varepsilon}_t + \tilde{\varepsilon}_{t-1} \tag{4.17}$$

where the contemporaneous innovation receives the smaller discount. This "flipping" of discount factors for s_t has an obvious impact on the conditional forecast error in predicting p_{t+1}. The effect on equilibrium dynamics can be seen by examining impulse response functions.

Figure 2 plots the impulse response functions for p_t and $\tilde{p}_t$ for $\theta = \sqrt{11}$ (which, according to (3.8), corresponds to a signal-to-noise ratio of 1/10) in the left panel, and $\theta = \sqrt{2}$ (which

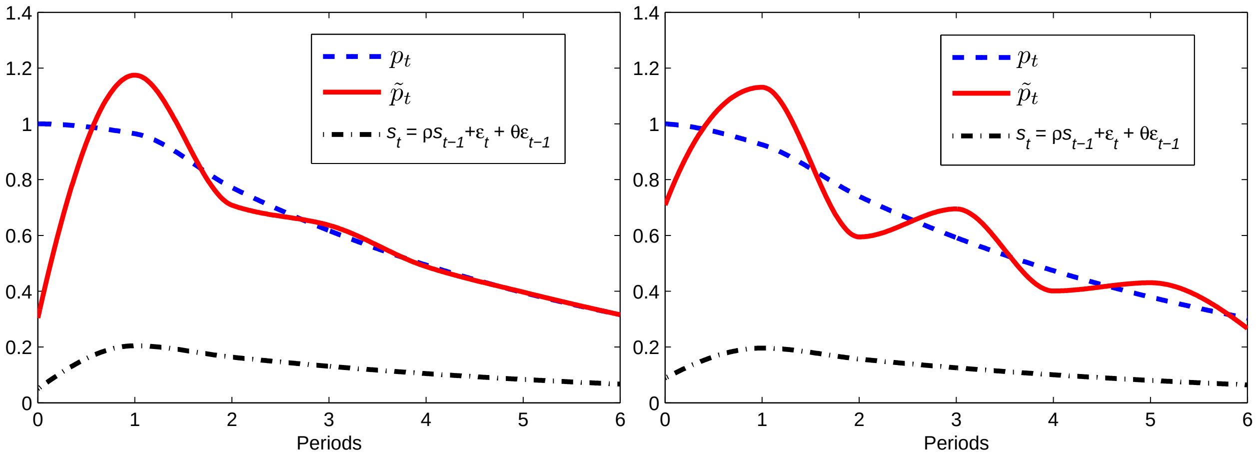

Figure 2a: $\theta = \sqrt{11}$ High Noise

Figure 2b: $\theta = \sqrt{2}$ Low Noise

Figure 2: Responses of p_t (4.15) and $\tilde{p}_t$ (4.16) to innovation in ε_t.

corresponds to a signal-to-noise ratio of 1) for a one-unit shock to ε at time t. The impulse responses are normalized with respect to the impulse response at impact for the price under complete information p_t. The additional parameters values are: $\rho = 0.8$, $\beta = 0.985$, $\sigma_\varepsilon = 1$.

For $\tilde{p}_t$, a one-unit shock to the structural innovation ε_t at time t has an interesting propagation effect. At impact $\tilde{p}_t$ underreacts with respect to the full information price p_t, while it overreacts one period after impact. The pattern then settles into waves of under– and overreaction over the subsequent periods as in Section 3.2. Comparing the impulse responses across the two panels reveals that in the presence of low noise, the initial underreaction at impact is smaller compared to the high noise case. In contrast, the subsequent "mood swings" are of greater magnitude and more persistent in the low noise case, while they tend to decay fairly quickly in the high noise case. Interestingly, if one were to measure the efficiency loss in terms of the relative discrepancy from the full information benchmark, it is not immediately clear whether one would prefer the low noise case to the high noise case. We leave the analysis of this issue to future work.

There are two aspects of the equilibrium dynamics that should be emphasized. First, despite the fact that the model itself is very simplistic–a univariate present value model–the propagation effects of $\tilde{p}_t$ can be quite rich. Here we have added a single parameter (θ) to a simple model and through an interesting informational angle are able to deliver significant propagation. Second, the difference in the dynamics between p_t and $\tilde{p}_t$ is quite dramatic given that the only distinguishing characteristic is information. It is a well known result that fundamental and non-fundamental moving average representations have the same covariance generating functions. Thus, the covariance generating functions of (4.14) and (4.17) are identical. The only difference here is that in one scenario, agents form expectations by conditioning on current and past ε_t; in the other, agents are assumed to only observe current and past p_t and s_t, which does not fully reveal ε_t.

5 Asymmetric Information

Having established the benchmark Information Equilibria for the symmetric information case, we now introduce asymmetric information structures. We examine two forms: the clustered informational setup of Futia (1981), where agents are clustered into groups of "informed" and "uninformed", and a dispersed informational setup, where each agent receives an idiosyncratic signal about market fundamentals. We derive an explicit connection between the two informational structures.

5.1 Clustered / Hierarchical Information Structure

There are two types of agents, *informed* and *uninformed*. The proportion of the informed agents is denoted by $\mu \in [0,1]$ and they are assumed to observe the entire history of the structural shock ε up to time t. The remaining $1-\mu$ agents are uninformed in the sense that they observe only equilibrium outcomes $(\mathbb{V}_t(p), \mathbb{M}_t(p))$. Using our notation for exogenous information

$$\begin{aligned} U_t^i &= \mathbb{V}_t(\varepsilon) \quad \text{for} \quad i \in \mu \\ U_t^i &= \{0\} \quad \text{for} \quad i \in 1-\mu. \end{aligned}$$

For $\mu = 1$ this setup is equivalent to the full information equilibrium (4.5); while for $\mu = 0$ it corresponds to the incomplete information equilibrium (4.8). Both agents are assumed to be rational and have common knowledge of rationality. The equilibrium is given by

$$p_t = \beta\big[\mu\mathbb{E}\big(p_{t+1}|V_t(\varepsilon) \vee \mathbb{M}_t(p)\big) + (1-\mu)\,\mathbb{E}\big(p_{t+1}|V_t(p) \vee \mathbb{M}_t(p)\big)\big] + s_t. \tag{5.1}$$

5.1.1 Existence of Information Equilibrium

Without loss of generality, we assume our initial guess, (3.10), contains exactly one zero inside the unit circle. The following theorem delivers existence and uniqueness conditions.

Theorem 2. *Under the exogenous information assumption $U_t^i = \mathbb{V}_t(\varepsilon)$ for $i \in \mu$ and $U_t^i = \{0\}$ for $i \in 1-\mu$, a unique Information Equilibrium for (5.1) with $|\beta| < 1$ always exists and is determined as follows: If there exists a $|\lambda| < 1$ such that*

$$A(\lambda) - \frac{\mu\beta A(\beta)}{h(\beta)} = 0 \tag{5.2}$$

then the IE of (5.1) is given by

$$p_t = (L-\lambda)Q(L)\varepsilon_t = \frac{1}{L-\beta}\left\{LA(L) - \beta A(\beta)\frac{h(L)}{h(\beta)}\right\}\varepsilon_t \tag{5.3}$$

with

$$h(L) \equiv \mu\lambda - (1-\mu)\mathcal{B}_\lambda(L), \qquad \mathcal{B}_\lambda(L) \equiv \frac{L-\lambda}{1-\lambda L}$$

If restriction (5.2) does not hold for $|\lambda| < 1$, the IE converges to (4.5).

Proof. See Appendix A. ∎

The intuition behind Theorem 2 is similar to that of Theorem 1 with the important difference that now restriction (5.2) must be satisfied in order for asymmetric information to persist in equilibrium. The initial exogenous informational guess of $p_t = (L-\lambda)Q(L)\varepsilon_t$ with $|\lambda| < 1$ implies uninformed agents, through knowledge of the price process alone ($\mathbb{V}_t(p)$), will be able to infer the linear combination of current and past $\tilde{\varepsilon}_t = \mathcal{B}_\lambda(L)\varepsilon_t$. In order for this informational assumption to survive in equilibrium, it must be the case that knowledge of the model does not provide any *additional* information. More precisely, through knowledge of the model $\mathbb{M}_t(p)$, uniformed agents are able to subtract off their expectation ($\mathbb{E}^{\mathcal{U}}$) from the equilibrium price. What remains is the expectation of the informed ($\mathbb{E}^{\mathcal{I}}$) and the exogenous process, s_t. That is,

$$\begin{aligned} p_t - \beta(1-\mu)\mathbb{E}^{\mathcal{U}}(p_{t+1}|\mathbb{M}_t(p) \vee \mathbb{V}_t(p)) &= \beta\mu\mathbb{E}^{\mathcal{I}}(p_{t+1}|\mathbb{V}_t(\varepsilon)) + s_t \\ &= \beta\mu L^{-1}\left[(L-\lambda)Q(L) - \frac{\lambda A(\beta)}{h(\beta)}\right]\varepsilon_t + A(L)\varepsilon_t \end{aligned} \tag{5.4}$$

where the last equality follows from the proof of Theorem 2 in Appendix A. (5.4) provides the exact linear combination of structural shocks that the uninformed agents are able to glean from having knowledge of the model. Therefore, the information provided by (5.4) must be equivalent to $\tilde{\varepsilon}_t$ in order for the exogenous informational guess to be consistent with an IE. This will be true if and only if (5.4) vanishes at $L = \lambda$. Condition (5.2) ensures that this is the case.[12]

Theorem 2 provides general restrictions for *any* stochastic process $A(\cdot)$. We now turn to a specific example to provide a sharper characterization of existence. Assuming s_t follows the ARMA(1,1) process (4.14) yields the following corollary,

Corollary 1. *The model described by* (5.1) *and* (4.14) *with* $\beta, \rho \in (0,1)$ *and* $\theta > 0$ *defines a space of existence for unique asymmetric IE of the form* (5.3)*. The space is characterized as follows.*

(1.a) If $\theta \leq 1$ *an asymmetric information equilibrium does not exist.*
(1.b) If $\theta > 1$ *an asymmetric equilibrium exists for any* $\mu > 0$ *and* $\rho \geq 0$ *if*

$$\theta \geq \left(\frac{1}{1-\beta(1+\rho)}\right) \tag{1.b}$$

(1.c) If $\theta > 1$*, and (1.b) is not satisfied, an asymmetric IE exists for* μ *if and only if* $\mu \in (0, \mu^*)$ *with*

$$\mu^* = \frac{(\theta-1)(1-\rho\beta)}{\beta(1+\rho)(1+\theta\beta)}$$

Proof. See Appendix A. □

Figure 3 characterizes the IE for the ARMA(1,1) process in (β, θ) space. Three points are noteworthy. First, as is evident from the figure and condition (1.a), if $\theta \leq 1$ an asymmetric

[12] Restriction (5.2) corresponds to Equation (6.18) in Futia (1981) and Assumption 3.7 in Kasa *et al.* (2008) once the exogenous process s_t and information structure are appropriately defined.

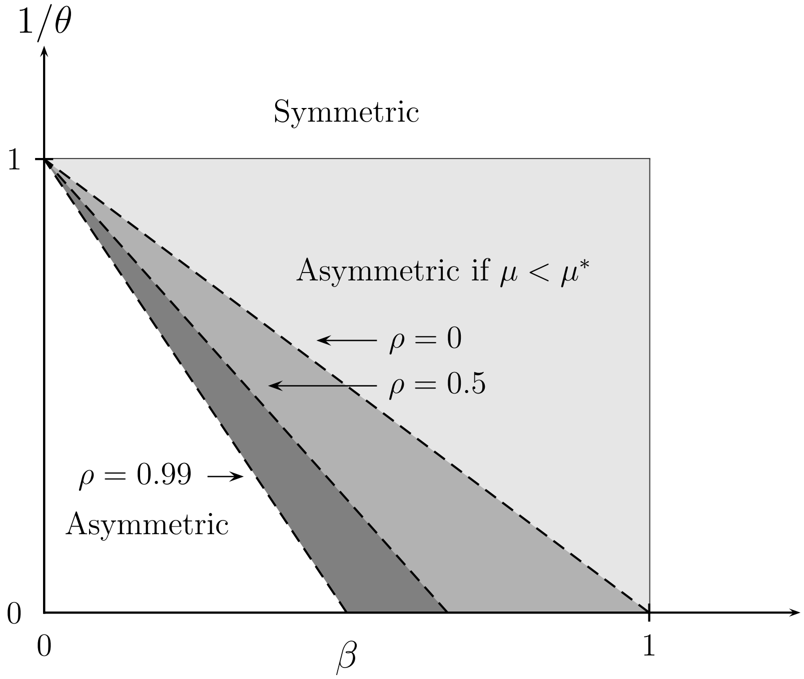

Figure 3: Existence of Symmetric and Asymmetric Information Equilibria following Corollary 1 for $s_t = \rho s_{t-1} + \varepsilon_t + \theta\varepsilon_{t-1}$.

IE does not exist regardless of the other parameters in the model. As emphasized in the introduction, it is the MA component which acts as a noise that prevents the uniformed from learning the true innovations, and the typical assumption of an AR(1) cannot deliver asymmetric information in equilibrium. A pure autoregressive representation will always reveal the information of the informed agents to the uninformed. Second, from condition (1.c) and figure 3, for a certain region of the parameter space (to the right of the dashed lines in figure 3) an asymmetric IE exists only if the distribution of informed traders is sufficiently small. The dashed lines represent the IE that prevails as $\mu \to 1$, plotted for various serial correlation parameters. To the left of the dashed line, asymmetric information will always be preserved in equilibrium regardless of the ratio of informed to uninformed. The derivations of section 3.2 demonstrate that an increase in θ may be interpreted as an increase in the noise associated with the signal extraction problem. The informational disparity between the informed and uninformed may become so large that no matter how many informed agents participate in the market, the uninformed will not fully learn the structural innovations in equilibrium. How the discount factor β alters the space of existence is similar to that of the serial correlation parameter ρ, which is the final point to be made. As the serial correlation in the s_t process increases and β increases, it is more difficult to preserve asymmetric information, *ceteris paribus* (the dashed line shifts to the left as ρ increases from 0 to 0.99). An increase in β and ρ leads to a longer lasting effect of current information. This results in a higher $|\lambda|$ and a decrease in the informational discrepancy between the informed and uninformed.

5.1.2 Characterization of Information Equilibrium The equilibrium representation (5.3) of Theorem 2 is algebraically the cleanest because it makes clear $\mu \to 0$ implies convergence to the equilibrium of Theorem 1, and as $\mu \to 1$, the equilibrium approaches the fully revealing (4.5). However there are equivalent representations which have a more natural economic interpretation. We state this as a corollary to Theorem 2.

Corollary 2. *If $|\lambda| < 1$, the equilibrium described in Theorem 2 has an equivalent representation in $\tilde{\varepsilon}$ space given by*

$$p_t = \frac{1}{L-\beta}\left\{(1-\lambda L)M^{\mathcal{U}}(L) - (1-\lambda\beta)M^{\mathcal{U}}(\beta)\right\}\tilde{\varepsilon}_t, \tag{5.5}$$

and a representation in ε space given by

$$p_t = \frac{1}{L-\beta}\left\{M^{\mathcal{I}}(L) - M^{\mathcal{I}}(\beta)\right\}\varepsilon_t \tag{5.6}$$

where $M^{\mathcal{U}}(L) = \frac{(LA(L)+\mu\beta\lambda Q_0)}{(L-\lambda)}$ and $M^{\mathcal{I}}(L) = LA(L) - (1-\mu)\beta Q_0 \frac{L-\lambda}{1-\lambda L}$

Proof. Follows directly from Theorem 2. □

These representations highlight what the uninformed and informed agents believe to be "market fundamentals." For the informed (uninformed) agents, the market fundamental is a combination of the exogenous process, s_t, *and* the forecast error of the uninformed (informed) agents. The modification of the Hansen-Sargent formula is due to the speculative dynamics associated with IE. This is true even though (as we show below) uninformed agents are not forming higher-order expectations.

Compared to the symmetric IE of Section 4, the partial fractions expansion of the IE in Theorem 2, given by

$$\begin{aligned} p_t &= p_t^{pf} - \beta A(\beta)\left(\sum_{j=1}^{\infty}\beta^j\varepsilon_{t+j} + \frac{(1-\mu)(1-\lambda^2)}{\beta-\lambda-\mu\beta(1-\lambda^2)}\sum_{j=0}^{\infty}\lambda^j\varepsilon_{t-j}\right) \\ &= p_t^{pf} - \beta A(\beta)\sum_{j=1}^{\infty}\beta^j\varepsilon_{t+j} - M^{\mathcal{U}}(\beta)(1-\lambda^2)\sum_{j=0}^{\infty}\lambda^j\varepsilon_{t-j}, \end{aligned} \tag{5.7}$$

shows that now the linear combination of past ε_t's that must be subtracted from the perfect foresight price is weighted by the proportion of informed agents in the economy. Moreover, changes in μ will alter λ and in turn, the information content of the price.

Recent papers have emphasized the role of higher-order belief (HOB) dynamics and the subsequent breakdown in the law of iterated expectations with respect to the average expectations operator in models with asymmetric information, but resort to numerical analysis or truncation in demonstrating the dynamic case [Allen *et al.* (2006), Bacchetta and van Wincoop (2006), Nimark (2005), Bacchetta and van Wincoop (2004)]. We are able to characterize these objects in closed form and show precisely *why* HOB exist and *why* and *when* HOB imply a failure of the law of iterated expectations. This is of first order given the findings of, for example, Pearlman and Sargent (2005), who document that HOB do not exist in the model of Townsend (1983), as previously believed. The following proposition shows why HOB are formed and why HOB lead to the break down in the law of iterated expectations for the average expectations operator.

Proposition 1. *If the information equilibrium given by Theorem 2 holds for $|\lambda| < 1$, then*
i. the informed agents form higher order beliefs, while the uninformed do not;
ii. the average expectations operator does not satisfy the law of iterated expectations.

Proof. The proof of the proposition is perhaps more instructive than the proposition itself and hence selected parts of the proof follow, while the proof in its entirety can be found in Appendix A. □

In a model with asymmetrically informed agents there exists an incentive to form higher-order beliefs. The *average* expectation of the price determines equilibrium according to (5.1). So if agent j could observe agent k's forecast of tomorrow's price, her forecast error would be smaller.[13] From the definition of equilibrium, we may write the informed and uninformed agents' expectation of tomorrow's price as

$$\mathbb{E}_t^{\mathcal{I}}(p_{t+1}) = \beta \mathbb{E}_t^{\mathcal{I}} \bar{\mathbb{E}}_{t+1} p_{t+2} + \mathbb{E}_t^{\mathcal{I}} s_{t+1}, \qquad \mathbb{E}_t^{\mathcal{U}}(p_{t+1}) = \beta \mathbb{E}_t^{\mathcal{U}} \bar{\mathbb{E}}_{t+1} p_{t+2} + \mathbb{E}_t^{\mathcal{U}} s_{t+1} \tag{5.8}$$

where $\bar{\mathbb{E}}_t$ denotes the time-t average forecast. Therefore, the optimal conditional expectation of each agent type is a discounted expectation of next period's average expectation. Writing the price as $p_t = (L - \lambda) Q(L) \varepsilon_t$ where $|\lambda| < 1$, the appendix shows the time $t+1$ average expectation of the price at $t+2$ is

$$\mu \mathbb{E}_{t+1}^{\mathcal{I}} p_{t+2} + (1-\mu) \mathbb{E}_{t+1}^{\mathcal{U}} p_{t+2} = p_{t+2} - Q_0[(1-\mu)\mathcal{B}_\lambda(L) - \mu\lambda]\varepsilon_{t+2} \tag{5.9}$$

The second term in the RHS of (5.9) represents the market's average forecast error. If we take the informed agent's time t expectation of this average

$$\begin{aligned} \mathbb{E}_t^{\mathcal{I}} \bar{\mathbb{E}}_{t+1} p_{t+2} &= \mathbb{E}_t^{\mathcal{I}} p_{t+2} + \mu\lambda Q_0 \mathbb{E}_t^{\mathcal{I}} \varepsilon_{t+2} - Q_0(1-\mu)\mathbb{E}_t^{\mathcal{I}} \mathcal{B}(L)\varepsilon_{t+2} \\ &= \mathbb{E}_t^{\mathcal{I}} p_{t+2} + \qquad 0 \qquad - Q_0(1-\mu)(1-\lambda^2)\left(\frac{\mu\lambda}{1-\lambda L}\right)\varepsilon_t \end{aligned} \tag{5.10}$$

we see that the *uninformed* agents' forecast error, given by the Blaschke factor $\mathcal{B}(L)\varepsilon_{t+2}$, is serially correlated with respect to the *informed* agents' information set. Re-arranging (5.10),

$$\mathbb{E}_t^{\mathcal{I}}[p_{t+2} - \bar{\mathbb{E}}_{t+1} p_{t+2}] = Q_0(1-\mu)(1-\lambda^2)\left(\frac{\mu\lambda}{1-\lambda L}\right)\varepsilon_t$$

gives the interpretation of the informed agents' expectation of the average forecast error in forecasting p_{t+2} [Bacchetta and van Wincoop (2006)]. Conditional on the informed's information set, the uninformed's forecast errors are serially correlated. Hence, the informed agents will always do better, if they correct their expectation of the average price according to the forecast errors of the uninformed.

Conversely, the uninformed do not form HOB because the forecast errors of the informed are not forecastable conditional on the uninformed's information set at time t

$$\begin{aligned} \mathbb{E}_t^{\mathcal{U}} \bar{\mathbb{E}}_{t+1} p_{t+2} &= \mathbb{E}_t^{\mathcal{U}} p_{t+2} + Q_0 \mu\lambda \mathbb{E}_t^{\mathcal{U}} \varepsilon_{t+2} - Q_0(1-\mu)\mathbb{E}_t^{\mathcal{U}} B(L)\varepsilon_{t+2} \\ &= \mathbb{E}_t^{\mathcal{U}} p_{t+2} + 0 - 0 \end{aligned} \tag{5.11}$$

[13] Note that we are abstracting from a Grossman-Stiglitz type market for information. While this type of market would have interesting features, we leave this for future research.

Agents form HOB if the *average* forecast error is serially correlated (along some dimension) with respect to their own information set. Notice that as information becomes symmetric (that is, as $\lambda \to 1$ and $\mu \to 1$) this term disappears. This result sheds light on the finding of Pearlman and Sargent (2005) who show that when agents' information sets are symmetric, no HOB exist.

Moreover, this analysis makes clear why the law of iterated expectations fails with respect to the average expectations operator. Combining (5.10) and (5.11) gives

$$\begin{aligned}\bar{\mathbb{E}}_t\bar{\mathbb{E}}_{t+1}p_{t+2} &= \mu\mathbb{E}_t^{\mathcal{I}}\bar{\mathbb{E}}_{t+1}p_{t+2} + (1-\mu)\mathbb{E}_t^{\mathcal{U}}\bar{\mathbb{E}}_{t+1}p_{t+2}\\ &= \bar{\mathbb{E}}_t p_{t+2} - (1-\mu)(1-\lambda^2)\left(\frac{Q_0\mu\lambda}{1-\lambda L}\right)\varepsilon_t\end{aligned}$$

Iterating on this equation, as shown in Appendix A, yields

$$\overline{\mathbb{E}}_t\overline{\mathbb{E}}_{t+1}\cdots\overline{\mathbb{E}}_{t+j}p_{t+j+1} = \overline{\mathbb{E}}_t p_{t+j+1} - (1-\mu)(1-\lambda^2)\left(\frac{\sum_{i=1}^{j}(\mu\lambda)^i Q_{j-i}}{1-\lambda L}\right)\varepsilon_t \tag{5.12}$$

When either $\mu = 0$ or $\mu = 1$, the law of iterated expectations holds as the average expectation collapses to the expectation of the informed or uninformed, respectively. This is because the law of iterated expectations certainly holds with respect to individual traders' information sets. Thus, it is the formation of HOB that leads directly to the failure of the law of iterated expectations. The degree to which the law of iterated expectations fails is determined by the distribution of informed agents, μ, and degree of asymmetric information, as indexed by λ.

Given that we have an analytic solution for the HOBs component, we are able to isolate the contribution of HOBs to equilibrium dynamics. The following proposition is the boundedly rational equilibrium that would emerge if all HOBs were removed.

Proposition 2. *Under the exogenous information assumption $U_t^i = \mathbb{V}_t(\varepsilon)$ for $i \in \mu$ and $U_t^i = \{0\}$ for $i \in 1-\mu$ and assuming that the informed agents do not form HOBs, a unique boundedly rational equilibrium always exists and is determined as follows: If there exists a $|\lambda| < 1$ such that*

$$A(\lambda) - \frac{\beta\mu A(\beta)}{(1-\lambda\beta)k(\beta)} = 0 \tag{5.13}$$

then the boundedly-rational IE is given by

$$p_t = \frac{1}{L-\beta}\left(LA(L) - \beta A(\beta)\frac{k(L)}{k(\beta)}\right)\varepsilon_t \tag{5.14}$$

where $k(L) = \frac{\mu\lambda}{1-\lambda\beta} - (1-\mu)\mathcal{B}_\lambda(L)$. If (5.13) does not hold for $|\lambda| < 1$, the IE converges to (4.5).

Proof. See Appendix A. ∎

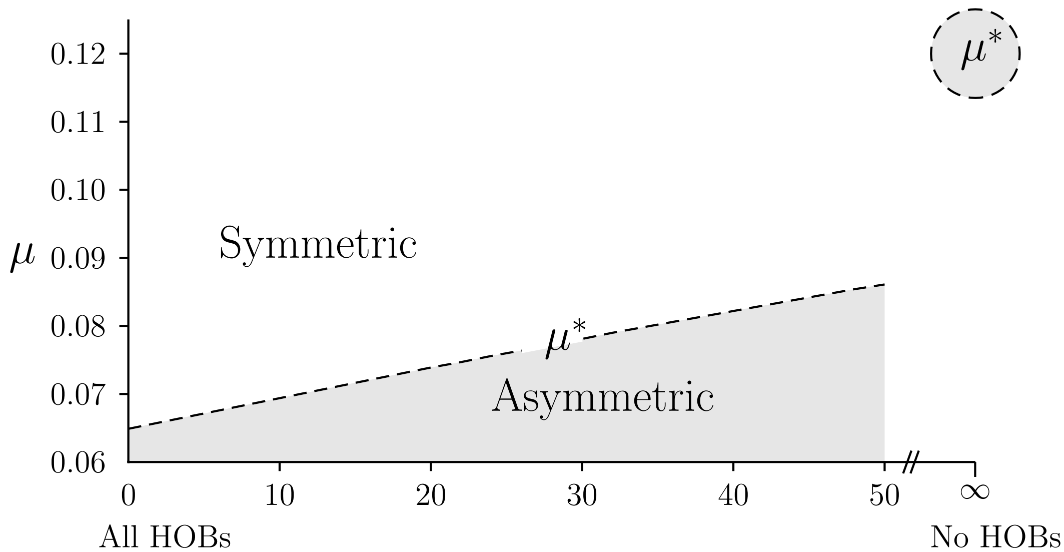

Figure 4: Existence space for the boundedly rational equilibrium as higher-order beliefs are removed from the expectation of informed agents: $s_t = 0.8s_{t-1} + \varepsilon_t + \sqrt{11}\varepsilon_{t-1}$, $\beta = 0.985$.

There are two elements of the boundedly rational equilibrium that are noteworthy. First, higher-order belief dynamics play a crucial role in disseminating information. As discussed above, the informed agents are correcting for the bias in the uniformed agents' expectations. But there is an important feedback mechanism at work. The uninformed agents are able to extract information about their own forecast errors by observing the endogenous variables due to the formation of HOBs. One consequence of this informational feedback effect is highlighted in figure 4. This figure shows the existence space of the boundedly rational information equilibrium as higher-order belief dynamics are removed from the expectation of the informed agents. As HOBs are removed, the asymmetric information equilibrium can support more informed agents. This is because the information that the uninformed are extracting from the endogenous variable is declining as fewer HOBs are being formulated. Second, the functional form of the equilibrium when HOBs are removed is identical to the functional form of Theorem 2. However, the coefficients characterizing the equilibrium will be different. Figure 5 plots the impulse responses to an innovation in ε_t for $\mu = 1$, $\mu = 0$, and $\mu = 0.06$ with and without HOBs. The figure shows the the boundedly rational equilibrium fluctuates relatively more wildly. These fluctuations can be attributed to the decline in information content of the endogenous variable.

5.2 Dispersed Information Structure In this section we assume that all agents are identical in terms of the imperfect quality of information they possess. In particular, we assume each agent observes its own particular "window" of the world, as in Phelps (1969). Agents observe a noisy signal of the innovation ε, which is idiosyncratic across agents. Information is dispersed in the sense that, although complete knowledge of the

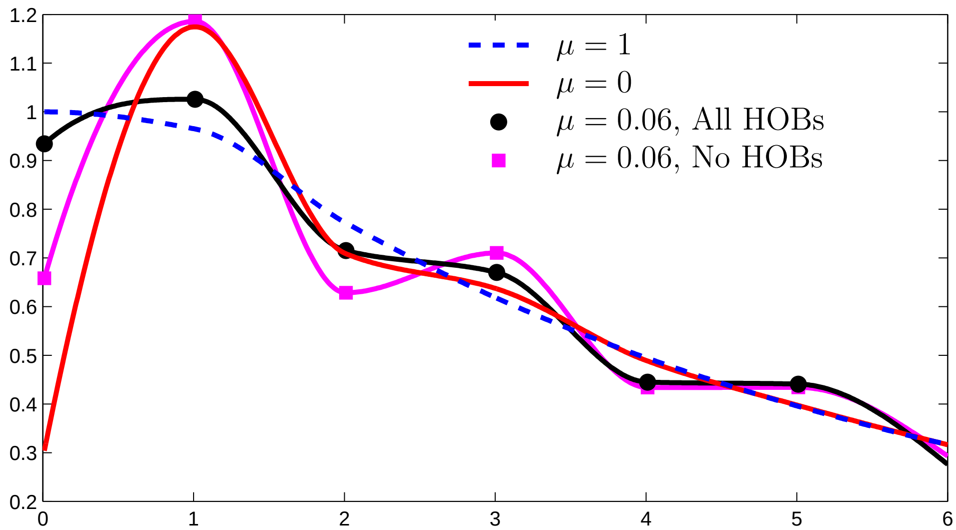

Figure 5: Responses of p_t to an innovation in ε_t with $s_t = 0.8 s_{t-1} + \varepsilon_t + \sqrt{11}\varepsilon_{t-1}$, $\beta = 0.985$ for $\mu = 1$, $\mu = 0$, $\mu = 0.6$ with and without HOBs.

fundamentals is not given to anyone agent, by pooling the noisy signal of all agents, it is possible to recover the full information about the state of the economy ε_t. The noisy signal is specified as

$$\varepsilon_{it} = \varepsilon_t + v_{it} \quad \text{with} \quad v_{it} \sim N\left(0, \sigma_v^2\right). \tag{5.15}$$

The exogenous dispersed information assumption corresponds to

$$U_t^i = \mathbb{V}_t\left(\varepsilon_i\right) \qquad \text{for} \quad i \in [0,1]\,. \tag{5.16}$$

Notice that when the noise is driven to zero, $\sigma_v^2 \to 0$, this setup is equivalent to the full information symmetric equilibrium (4.5), while an infinite noise, $\sigma_v^2 \to \infty$, yields the symmetric equilibrium (4.8).

As we have seen in Section 3.2, the information conveyed by the noisy signal ε_i can be measured by the signal-to-noise ratio, $\tau = \sigma_\varepsilon^2/\sigma_v^2$. Each agent then forms

$$\mathbb{E}_{it}\left(p_{t+1}\right) = \mathbb{E}\left(p_{t+1} | \mathbb{V}_t(\varepsilon_i) \vee \mathbb{V}_t(p) \vee \mathbb{M}_t(p)\right) \tag{5.17}$$

and the equilibrium is now given by

$$p_t = \beta \int_0^1 \mathbb{E}_{it}\left(p_{t+1}\right) di + s_t. \tag{5.18}$$

5.2.1 Existence of Information Equilibria What is unique about this setup is that each agent formulates a forecast by extracting optimally the information from a vector of two signals (p_t, ε_{it}). The basic idea of deriving a fundamental representation developed in Section 3.2 extends naturally to a multivariate setting. The mapping between the signal and innovations is now a matrix, and the invertibility of that matrix determines the information content of the signals. We maintain the assumption that (3.10) contains exactly one zero inside the unit circle; again, this is without loss of generality. The mapping between innovations and signals is given by

$$\begin{pmatrix} \varepsilon_{it} \\ p_t \end{pmatrix} = \begin{bmatrix} 1 & 1 \\ (L-\lambda)\,Q\,(L) & 0 \end{bmatrix} \begin{pmatrix} \varepsilon_t \\ v_{it} \end{pmatrix}. \tag{5.19}$$

Given the candidate price function, this matrix is of rank 1 at $L = \lambda$ and so it cannot be inverted. As shown in Appendix A and Rondina (2009), the invertible representation is derived through use of a Blaschke factor and factorization of the signal ε_{it}. The expectation (5.17) will always be given by the sum of two terms: a linear combination of current and past innovations ε_t and a linear combination of current and past idiosyncratic noise v_{it}. Taking the average of the expectations across agents, the second term would be zero yielding

$$\begin{aligned} \bar{\mathbb{E}}_t\left(p_{t+1}\right) &= \left(\ (L-\lambda)\,Q\,(L) + \lambda Q_0\ \right) L^{-1} \tfrac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2+\sigma_v^2}\varepsilon_t \\ &\quad + \left(\ (1-\lambda L)\,Q\,(L) - Q_0\ \right) L^{-1} \tfrac{\sigma_v^2}{\sigma_\varepsilon^2+\sigma_v^2} \left(\frac{L-\lambda}{1-\lambda L}\right)\varepsilon_t \\ &= \tfrac{\tau}{1+\tau}\mathbb{E}_t^{\mathcal{I}}(p_{t+1}) + \tfrac{1}{1+\tau}\mathbb{E}_t^{\mathcal{U}}(p_{t+1}) \end{aligned} \tag{5.20}$$

where the last line follows from the results in the previous section. Theorem 3 follows immediately.

Theorem 3. *Let $\tau \equiv \sigma_\varepsilon^2/\sigma_v^2$ be the signal-to-noise ratio of the noisy signal (5.15). Under the exogenous dispersed information assumption $U_t^i = \mathbb{V}_t(\varepsilon_i)$, a unique Information Equilibrium for (5.18) with $|\beta| < 1$ always exists and is equivalent to the equilibrium characterized in Theorem 2 where μ is now defined as*

$$\mu = \frac{\tau}{1+\tau}.$$

Under the ARMA(1,1) assumption for the process s_t, the existence space of an IE under dispersed information is identical to that provided by Corollary 1.

Corollary 3. *The model described by (5.18) and* (4.14) *with $\beta, \rho \in (0,1)$ and $\theta > 0$ defines a space of existence for unique asymmetric IE of the form* (5.3) *as described in Corollary 1, where μ is now specified as in Theorem 3.*

Proof. See Appendix A. □

Theorem 3 and Corollary 3 show that, from an aggregate point of view, the dynamics of the IE under dispersed information display a remarkable connection to the clustered/hierachical setup. In this setup, agents use the exogenous signal (5.15) to mitigate the dynamic noise associated with the non-fundamental MA representation. As the signal-to-noise ratio approaches 0, the average conditional expectation given by (5.20), and therefore the equilibrium, converges to the fully uninformed equilibrium of Theorem 1 and (4.8). Conversely as the signal-to-noise ratio approaches infinity, the IE converges to the fully informed equilibrium of (4.5). Hence the equilibrium may be interpreted as a linear combination of informed and uninformed agents, where the proportion of informed to uninformed is given by the signal-to-noise ratio.

As noted by Theorem 3, the restriction for asymmetric information to persist in equilibrium is given by (5.2), with μ appropriately defined. As with previous definitions of an information equilibrium, knowledge of the model plays a crucial role. In this setup knowledge of the model for agent i ($\mathbb{M}_t(p, \varepsilon_i)$) is given by: $p_t - \beta \mathbb{E}(p_{t+1}|\varepsilon_i^t, p^t) = \beta[\bar{\mathbb{E}}(p_{t+1}) - \mathbb{E}(p_{t+1}|\varepsilon_i^t, p^t)] + s_t$. Appendix A shows that adding this additional piece of information to the vector of observables (5.19) gives

$$\begin{pmatrix} \varepsilon_{it} \\ p_t \\ \mathbb{M}_t(p, \varepsilon_i) \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ (L-\lambda)\, Q\,(L) & 0 \\ A\,(L) & \mu \left(\frac{1-\lambda^2}{1-\lambda L}\right) \beta Q_0 \end{pmatrix} \begin{pmatrix} \varepsilon_t \\ v_{it} \end{pmatrix} \tag{5.21}$$

An IE stipulates that this enlarged information set cannot reveal any additional information than the price function. Therefore, the matrix mapping innovations to signals must also be of rank 1 at $L = \lambda$. It is straightforward to show that 2 of the 3 minors of this matrix have rank 1 at $L = \lambda$. For the third minor the condition for rank 1 is

$$\mu \left(\frac{1-\lambda^2}{1-\lambda L} \right) \beta Q_0 - A\,(L) = 0 \quad \text{at} \quad L = \lambda.$$

which is identical to (5.2).

The intuition for the existence of a dispersed information equilibrium as μ changes lies in the information discrepancy and discounting mechanisms outlined in Section 5.1. As the precision of the private signal ε_{it} increases, agent i will rely more on the signal to forecast the innovation in s_t. In so doing, all agents will put more weight on the current innovation ε_t, which reduces the discounting on current information. This is analogous to the direct effect triggered by an increase in μ upon the information conveyed by the model to the uninformed agents in the hierarchical case. The direct effect triggers an equilibrium effect in this case as well. Because all the agents rely more on their private signal, on average expectations will discount current information less and thus the equilibrium price, being a function of the average expectations, will carry more information about the current innovation. As a consequence, the equilibrium price will become more informative and the dispersion of information in equilibrium reduced, up to the point of disappearance for μ large enough.

5.2.2 Characterization of Information Equilibria

While a reinterpretation of μ allows for a connection to the IE of the previous section, there are noteworthy differences between the two setups. For example, the cross sectional distribution of beliefs in the hierarchical setup was degenerate; whereas in the dispersed information, a well defined cross section emerges with interesting properties. First, individual expectations are persistently different from the average expectation. Second, the cross sectional variation is perpetual in the sense that the unconditional cross sectional variance is positive. These two results are stated in the following proposition.

Proposition 3. *The difference between the average market expectation and individual expectations in the IE of Theorem 3 is given by the AR(1) process*

$$\mathbb{E}_t^i(p_{t+1}) - \bar{\mathbb{E}}_t\,(p_{t+1}) = -\mu \frac{A\,(\beta)}{h\,(\beta)} \frac{1-\lambda^2}{1-\lambda L} v_{it}. \tag{5.22}$$

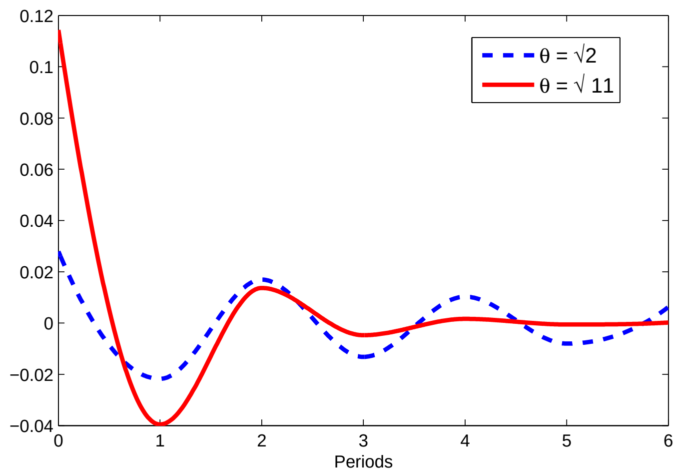

Figure 6: Impulse responses of the deviation of agent i's expectation from the market average (5.22) for $\theta = \sqrt{11}$ (red line) and $\theta = \sqrt{2}$ (blue line)

The cross-sectional unconditional variance of the difference in beliefs is

$$\mu^2 \left(1 - \lambda^2\right) \left(\frac{A(\beta)}{h(\beta)}\right)^2 \sigma_v^2. \tag{5.23}$$

Proof. See Appendix A. □

This proposition points out a remarkable feature of the IE of Theorem 3. Even though agents observe a common source of information (the equilibrium price), the presence of exogenous dispersed information prevents the agents from learning perfectly the aggregate innovation. The pooling of information through the market interaction fails to result in a sufficient statistic for the state of the world. Rational agents have dispersed beliefs, in equilibrium, that are persistently far away from the average market beliefs and may be so for many periods. The extent of the divergence of opinion depends on the parameters of the model. Agents' beliefs tend to converge when they all become very uninformed or when they all become very informed. As $\mu \to 0$, or as $\mu \to 1$ which implies $|\lambda| \to 1$, the unconditional variance (5.23) converges to zero. Figure 6 reports the impulse response function of (5.22) to a unit variance positive innovation in the noise process v_{it}, assuming s_t is given by (1.2). A familiar pattern emerges from the figure: the disagreement of the individual agent with respect to the market average goes through waves of under and overreaction with respect to the market expectations. Such waves are specific to each agent as they are the result of the individual innovation component v_{it}.

An additional implication of diverse beliefs in equilibrium is that *all* the agents will form higher order beliefs, whereas only the informed agents did so in the equilibrium of Theorem 2. Here HOB do not imply that agent i is forecasting the forecasts of agent j, which would not make sense as each agent is atomistic. Instead agent i uses her exogenous signal to forecast the forecasts of the *market*. Hence, the aggregate HOB take the same form as those in the hierarchical case. We summarize the description of the HOB for the dispersed information case in the following proposition

Proposition 4. *If the information equilibrium given by Theorem 3 holds for $|\lambda| < 1$, then*
i. all agents form higher order beliefs
ii. the average expectations operator does not satisfy the law of iterated expectations
iii. higher-order belief dynamics follow an AR(1) *process given by*

$$\bar{\mathbb{E}}_t \bar{\mathbb{E}}_{t+1} \cdots \bar{\mathbb{E}}_{t+j} p_{t+j+1} = \bar{\mathbb{E}}_t p_{t+j+1} - (1-\mu)(1-\lambda^2)\left(\frac{\sum_{i=1}^{j} (\mu\lambda)^i Q_{j-i}}{1-\lambda L}\right)\varepsilon_t$$

where $\mu \equiv \frac{\tau}{1+\tau}$.

The reason all agents form HOB is due to the presence of imperfectly informative private *and* public signals. Morris and Shin (2002) show that noisy public information is a key factor that causes agents to rationally engage in the guessing game about the average beliefs of the average beliefs of the average beliefs, and so on. When an agent is faced with the problem of forming an opinion about the average market expectation, she will take into account the fact that all the other agents observe a common signal, in this case the price. Hence the price plays an informative role as it is an important predictor of the average opinion of the market.

For this model, agent i's expectation of the market expectation is given by

$$\mathbb{E}_{it} \bar{\mathbb{E}}_{t+1} \left(p_{t+2}\right) = \mathbb{E}_{it}(p_{t+2}) - Q_0 \left(1-\mu\right) \mathbb{E}_{it} \left(\frac{L-\lambda}{1-\lambda L} \varepsilon_{t+2}\right).$$

which is obtained by taking (5.20) one period forward and noticing that $\mathbb{E}_{it}\left(\varepsilon_{t+j}\right) = 0$ for $j > 0$ since any information set at time t contains no information about future ε's under our assumptions. The non-fundamental MA term $(L-\lambda)/(1-\lambda L)\varepsilon_{t+2}$ represents the noise generated by the public signal. Following the intuition of Morris and Shin (2002), if this term was not present or if $|\lambda| > 1$, the agents would not form HOB, as the expectation of the market forecast would coincide exactly with their forecast. The higher the relative precision of the public signal $(1-\mu)$, the more important that signal will be in forming expectations about market beliefs. In the Appendix we show that this term is not zero and that agent i's beliefs about the market expectations are given by

$$\mathbb{E}_{it} \bar{\mathbb{E}}_{t+1} \left(p_{t+2}\right) = \mathbb{E}_{it}\left(p_{t+2}\right) - Q_0 \left(1-\mu\right) \mu\lambda \frac{(1-\lambda^2)}{1-\lambda L} \varepsilon_{it}. \tag{5.24}$$

The second term on the RHS of (5.24) shows that the exogenous signal received by each agent (ε_{it}) is correlated with the endogenous noise. In other words, the exogenous signal has predictive power and agents will use it at each date in order to adjust their forecast of the market average. While individual agents have uncorrelated forecast errors, the forecast error of the market is a function of the noise implicit in the public signal. Rational agents will recognize this and will smooth the forecast error of the market by conditioning on their own private information.

Taking the average across agents gives

$$\bar{\mathbb{E}}_t \bar{\mathbb{E}}_{t+1} \left(p_{t+2}\right) = \bar{\mathbb{E}}_t \left(p_{t+2}\right) - Q_0 \left(1-\mu\right) \mu\lambda \frac{(1-\lambda^2)}{1-\lambda L} \varepsilon_t, \tag{5.25}$$

which is exactly equal to (5.10). This shows that the law of iterated expectations fails to hold in the dispersed information case, even though no agent is superiorly informed - as was necessary in the hierarchical information setup. The reason for the convergence to (5.10) is because, on average, the exogenous signal, $\varepsilon_{it} = \varepsilon_t + v_{it}$, will reveal ε_t perfectly. Individual agents will adjust their expectation of the market average according to (5.24), but integrating over all agents implies that, on average, the market behaves as the informed agents of the previous section.

One crucial difference between our setup and that of Morris and Shin (2002) is that the public signal in our case is an equilibrium variable and the noise is itself a feature that emerges endogenously in equilibrium. Our methods present a framework that can easily accommodate the informational assumptions of Morris and Shin (2002) but where the non-linear interaction between private and public information is pervasive, in the sense that it can take place in any market just because rational agents extract information from the commonly and perfectly observed equilibrium price. In other words, our results suggest that any competitive speculative dynamic market, because of its functioning through a commonly observed signal, the price, by definition contains the seeds of the informational inefficiency formalized by Morris and Shin (2002) and extensively analyzed by Angeletos and Pavan (2007) and Angeletos and Pavan (2009). Our methods then suggest interesting applications where part of the public noise can be endemic to the dynamics of the equilibrium and can interact in interesting ways with other sources of noise or with economic policies.

6 Concluding Comments

Models with incomplete information offer a rich set of results unobtainable in representative agent, rational expectations economies and have implications for business cycle modeling, asset pricing and optimal policy, to name a few applications. There are two important characteristics of these models emphasized in this paper. First, the dynamic signal extraction of the type studied here offers an endogenous propagation mechanism. A robust finding in the empirical macroeconomic literature is that data prefer DSGE models with internal propagation mechanisms such as habit formation, investment adjustment costs, nominal rigidities, etc. [Cogley and Nason (1995)]. Our paper suggests that in lieu of these mechanisms, modeling uncertainty in a more nuanced manner might provide the needed propagation. Second, the law of iterated expectations does not hold with respect to the average expectations operator in dynamic models of asymmetric information. A robust finding in the empirical asset pricing literature is a rejection of the martingale hypothesis. Therefore, the breakdown in the law of iterated expectations due to speculative dynamics may play a pivotal role in understanding this empirical finding.

More broadly, the results of this paper suggest that models with dynamic incomplete information show great promise for many applications. This has been known (or at least believed) since Lucas (1972). However, solving and characterizing equilibrium has proven to be a significant challenge, impeding the progress of these models. In this paper, we derived existence and uniqueness conditions, along with a solution methodology that yields analytic solutions to dynamic models with incomplete information. While there is much more work to be done, this solution methodology is a step towards making these models usable for analysis.

7 Appendix A: Proofs

Theorem 1

Substituting the conditional expectation (4.6) into the equilibrium (4.1) yields the z-transform in ε_t-space

$$\begin{aligned} Q(z)\prod_{i=1}^{n}(z-\lambda_i) &= \beta z^{-1}[Q(z)\prod_{i=1}^{n}(1-\lambda_i z) - Q_0]\prod_{i=1}^{n}\mathcal{B}_{\lambda_i}(z) + A(z) \\ &= \beta z^{-1}[Q(z)\prod_{i=1}^{n}(z-\lambda_i) - Q_0\prod_{i=1}^{n}\mathcal{B}_{\lambda_i}(z)] + A(z) \end{aligned}$$

A bit of algebra yields

$$Q(z)(z-\beta)\prod_{i=1}^{n}(z-\lambda_i) = zA(z) - Q_0\prod_{i=1}^{n}\mathcal{B}_{\lambda_i}(z) \tag{7.1}$$

For $|\beta| < 1$, uniqueness requires the $Q(\cdot)$ process to be analytic inside the unit circle, which will not be the case unless the process vanishes at the poles $z = \{\lambda_i, \beta\}$ for every i. For simplicity, we assume $\lambda_i \neq \lambda_j$ for any $i \neq j$, however this restriction can be relaxed [see, Whiteman (1983)]. If $n = 1$, we also rule out $\lambda = \beta$, because the zero in the p_t process (λ_i) would cancel the pole in the denominator (β) and the rational expectations solution would not be unique (i.e., Q_0 could be set arbitrarily). Evaluating at $z = \lambda_i$ gives the restriction on the $A(\cdot)$ process, $A(\lambda_i) = 0$ for all i, which corresponds with (4.7). Evaluating at $z = \beta$ gives

$$Q_0 = \frac{\beta A(\beta)}{\prod_{i=1}^{n}\mathcal{B}_{\lambda_i}(\beta)} \tag{7.2}$$

Substituting this into (7.1) yields (4.8).

Theorem 2

Given the price process follows (3.10) for $n = 1$, the conditional expectations for the informed and uninformed are given by

$$\begin{aligned} E_t^I(p_{t+1}) &= L^{-1}[(L-\lambda)Q(L) + \lambda Q_0]\varepsilon_t \\ E_t^U(p_{t+1}) &= L^{-1}[(L-\lambda)Q(L) - Q_0\mathcal{B}_\lambda(L)]\varepsilon_t \end{aligned}$$

Substituting the expectations into the equilibrium gives the z-transform in ε_t space as

$$(z-\lambda)Q(z) = \beta\mu z^{-1}[(z-\lambda)Q(z) + \lambda Q_0] + \beta(1-\mu)z^{-1}[(z-\lambda)Q(z) - Q_0\mathcal{B}_\lambda(z)] + A(z) \tag{7.3}$$

and re-arranging yields the following functional equation

$$(z-\lambda)(z-\beta)Q(z) = zA(z) + \beta Q_0[\mu\lambda - (1-\mu)\mathcal{B}_\lambda(z)]$$

The $Q(\cdot)$ process will not be analytic unless the process vanishes at the poles $z = \{\lambda, \beta\}$. Evaluating at $z = \lambda$ gives the restriction on $A(\cdot)$, $A(\lambda) = -\beta\mu Q_0$. Rearranging terms

$$\begin{aligned} (z-\beta)Q(z) &= \frac{1}{z-\lambda}\big\{zA(z) + \beta Q_0[\mu\lambda - (1-\mu)\mathcal{B}_\lambda(z)]\big\} \\ &= \frac{1}{z-\lambda}\big\{zA(z) + \beta Q_0 h(z)\big\} \end{aligned} \tag{7.4}$$

where $h(z) \equiv [\mu\lambda - (1-\mu)\mathcal{B}_\lambda(z)]$. Evaluating at $z = \beta$ gives Q_0 as $Q_0 = -\frac{A(\beta)}{h(\beta)}$. This implies the restriction on $A(\cdot)$ is

$$A(\lambda) = \frac{\beta\mu A(\beta)}{h(\beta)}$$

which is (5.2). Substituting this into (7.4) delivers (5.3).

Corollary 1

The proof follows immediately from the restriction (5.2). Condition (1.a) is derived by taking the limit of (5.2) as $\mu \to 0$. This is the IE that would exist if no informed agents populated the model. Intuitively, if no asymmetric IE exists in this case, then none would exist if informed agents had positive measure. This restriction is given by $A(\lambda) = 0$ for $|\lambda| < 1$, which for the process $A(\lambda) = (1+\theta\lambda)/(1-\rho\lambda)$, implies $\theta \in (0, 1)$. Notice that because $\theta > 0$, $\lambda \to -1$ from above. Substituting $\lambda = -1$ into (5.2) and solving for μ gives condition (1.c). When $\lambda = -1$, the IE equilibrium converges to the symmetric information case. Setting μ^* equal to unity and solving for θ gives condition (1.b).

Corollary 2

Solving the asymmetric IE in $\tilde{\varepsilon}$ space. The guess for the price is $p_t = (1 - \lambda L)Q(L)\tilde{\varepsilon}_t$. The expectations are given by

$$E_t^U(p_{t+1}) = L^{-1}[(1 - \lambda L)Q(L) - Q_0]\tilde{\varepsilon}_t$$

$$E_t^I(p_{t+1}) = L^{-1}[(L - \lambda)Q(L) + \lambda Q_0]\varepsilon_t = L^{-1}[(1 - \lambda L)Q(L) + \frac{(1 - \lambda L)\lambda Q_0}{L - \lambda}]\tilde{\varepsilon}_t$$

where the second equality follows from multiplying and dividing by the Blaschke factor. Substituting in the expectations and a bit of algebra gives the equilibrium in $\tilde{\varepsilon}$ space as

$$\begin{aligned} p_t &= \frac{1}{L-\beta}\left\{(1 - \lambda L)\frac{(LA(L) + \mu\beta Q_0\lambda)}{(L-\lambda)} - \beta Q_0(1-\mu)\right\}\tilde{\varepsilon}_t \\ &= \frac{1}{L-\beta}\left\{(1 - \lambda L)M(L) - (1 - \lambda\beta)M(\beta)\right\}\tilde{\varepsilon}_t \end{aligned} \tag{7.5}$$

where $M(L) = \frac{(LA(L)+\mu\beta Q_0\lambda)}{(L-\lambda)}$. There is a pole at $z = \lambda$ on the right-hand side of (7.5) in two places and a pole at $z = \beta$. Unless these poles are removed the equilibrium is not "informationally stable" in $\tilde{\varepsilon}$ space. Notice that if $\mu = 0$, the assumption that $A(z)$ has a zero at λ will suffice to ensure that (7.5) is informationally stable (as shown in the symmetric case). However, here there is an additional pole (still at λ) in the informed agent's forecast error. This is the additional information that the uninformed see in this model vis-a-vis the model where $\mu = 0$. So now the restriction on $A(\cdot)$ is

$$\begin{aligned} [\beta Q_0\mu\lambda + zA(z)]\Big|_{\lambda} &= 0 \\ A(\lambda) + \beta\mu Q_0 &= 0 \end{aligned} \tag{7.6}$$

which is, of course, identical to the restriction in ε space, (5.2). Q_0 is then set to remove the pole at $z = \beta$,

$$Q_0 = -\frac{A(\beta)\mathcal{B}_\lambda(\beta)^{-1}}{\mu\lambda\mathcal{B}_\lambda(\beta)^{-1} - (1-\mu)} = -\frac{A(\beta)}{h(\beta)}$$

Proposition 1

Write the equilibrium price as $p_t = (L - \lambda)Q(L)\varepsilon_t$ where $|\lambda| < 1$ and $Q(L)$ satisfies (5.3). For $j = 1$, the time $t + 1$ average expectation of the price at $t + 2$ is given by

$$\begin{aligned} \overline{\mathbb{E}}_{t+1}p_{t+2} &= \mu\mathbb{E}_{t+1}^I p_{t+2} + (1-\mu)\mathbb{E}_{t+1}^U p_{t+2} \\ &= L^{-1}(L-\lambda)Q(L)\varepsilon_{t+1} + L^{-1}Q_0[\mu\lambda - (1-\mu)\mathcal{B}_\lambda(L)]\varepsilon_{t+1} \\ &= p_{t+2} + L^{-1}Q_0[\mu\lambda - (1-\mu)\mathcal{B}_\lambda(L)]\varepsilon_{t+1} \end{aligned} \tag{7.7}$$

The informed agent's time t expectation of the average expectation at $t+1$ is

$$\begin{aligned}
&\mathbb{E}_t^I\overline{\mathbb{E}}_{t+1}p_{t+2}\mathbb{E}_t^I p_{t+2} + \mu\lambda Q_0\mathbb{E}_t^I\varepsilon_{t+2} - Q_0(1-\mu)\mathbb{E}_t^I\mathcal{B}_\lambda(L)\varepsilon_{t+2} \\
&= \mathbb{E}_t^I p_{t+2} + 0 - Q_0(1-\mu)L^{-2}\{\mathcal{B}_\lambda(L) - \mathcal{B}_\lambda(0) - \mathcal{B}_\lambda(1)L\}\varepsilon_t
\end{aligned} \tag{7.8}$$

Note

$$\begin{gathered}
\mathcal{B}_\lambda(L) = \frac{L-\lambda}{1-\lambda L} = (L-\lambda)(1+\lambda L+\lambda^2L^2+\lambda^3L^3+\cdots) \\
\mathcal{B}_\lambda(0) = -\lambda, \qquad \mathcal{B}_\lambda(1) = (1-\lambda)(1+\lambda) = (1-\lambda^2) \\
\therefore\ \mathbb{E}_t^I\overline{\mathbb{E}}_{t+1}p_{t+2} = \mathbb{E}_t^I p_{t+2} - (1-\mu)Q_0L^{-2}\{\mathcal{B}_\lambda(L)+\lambda-(1-\lambda^2)L\}\varepsilon_t
\end{gathered}$$

where

$$\frac{L-\lambda}{1-\lambda L} + \lambda - (1-\lambda^2)L = \frac{\lambda(1-\lambda^2)L^2}{1-\lambda L}$$

Therefore, the informed agent's expectation of the average expectation is

$$\mathbb{E}_t^I\overline{\mathbb{E}}_{t+1}p_{t+2} = \mathbb{E}_t^I p_{t+2} - (1-\lambda^2)(1-\mu)\left(\frac{Q_0\lambda}{1-\lambda L}\right)\varepsilon_t \tag{7.9}$$

For the uninformed,

$$\begin{aligned}
\mathbb{E}_t^U\overline{\mathbb{E}}_{t+1}p_{t+2} &= \mathbb{E}_t^U p_{t+2} + Q_0\mu\lambda\mathbb{E}_t^U\varepsilon_{t+2} - Q_0(1-\mu)\mathbb{E}_t^U\mathcal{B}_\lambda(L)\varepsilon_{t+2} \\
&= \mathbb{E}_t^U p_{t+2} + 0 - Q_0(1-\mu)\mathbb{E}_t^U\mathcal{B}_\lambda(L)\varepsilon_{t+2} \\
&= \mathbb{E}_t^U p_{t+2} + 0 - Q_0(1-\mu)\mathbb{E}_t^U\tilde{\varepsilon}_{t+2} \\
&= \mathbb{E}_t^U p_{t+2} + 0 - 0
\end{aligned}$$

Thus the uninformed are *not* forming higher-order expectations.

Therefore, we have that

$$\begin{aligned}
\overline{\mathbb{E}}_t\overline{\mathbb{E}}_{t+1}p_{t+2} &= \mu\mathbb{E}_t^I\overline{\mathbb{E}}_{t+1}p_{t+2} + (1-\mu)\mathbb{E}_t^U\overline{\mathbb{E}}_{t+1}p_{t+2} \\
&= \overline{\mathbb{E}}_t p_{t+2} - (1-\mu)(1-\lambda^2)\left(\frac{Q_0\mu\lambda}{1-\lambda L}\right)\varepsilon_t \\
&\neq \overline{\mathbb{E}}_t p_{t+2}
\end{aligned} \tag{7.10}$$

For $j=2$, we need to calculate $\overline{\mathbb{E}}_t\overline{\mathbb{E}}_{t+1}\overline{\mathbb{E}}_{t+2}p_{t+3}$. From (7.7)

$$\overline{\mathbb{E}}_{t+2}p_{t+3} = p_{t+3} + Q_0[\mu\lambda - (1-\mu)\mathcal{B}_\lambda(L)]\varepsilon_{t+3}$$

We now need the uninformed and informed's time $t+1$ expectations of $\overline{\mathbb{E}}_{t+2}p_{t+3}$. The uninformed

$$\begin{aligned}
\mathbb{E}_{t+1}^U[\overline{\mathbb{E}}_{t+2}p_{t+3}] &= \mathbb{E}_{t+1}^U p_{t+3} \\
&= \left[\frac{(1-\lambda L)Q(L)}{L^2}\right]_+\mathcal{B}_\lambda(L)\varepsilon_{t+1} \\
&= L^{-2}[(L-\lambda)Q(L) - \{Q_0+(Q_1-\lambda Q_0)L\}\mathcal{B}_\lambda(L)]\varepsilon_{t+1}
\end{aligned} \tag{7.11}$$

The informed

$$\begin{aligned}
\mathbb{E}_{t+1}^I[\overline{\mathbb{E}}_{t+2}p_{t+3}] &= \mathbb{E}_{t+1}^I p_{t+3} + Q_0\mu\lambda\mathbb{E}_{t+1}^I\varepsilon_{t+3} - Q_0(1-\mu)\mathbb{E}_{t+1}^I\mathcal{B}_\lambda(L)\varepsilon_{t+3} \\
&= L^{-2}[(L-\lambda)Q(L) + \lambda Q_0 - (Q_0-\lambda Q_1)L]\varepsilon_{t+1} - \left[\frac{Q_0(1-\mu)(1-\lambda^2)\lambda}{1-\lambda L}\right]\varepsilon_{t+1}
\end{aligned} \tag{7.12}$$

Combining (7.11) and (7.12) gives

$$\overline{\mathbb{E}}_{t+1}\overline{\mathbb{E}}_{t+2}p_{t+3} = \mu\{L^{-2}[(L-\lambda)Q(L) + \lambda Q_0 - (Q_0 - \lambda Q_1)L]\varepsilon_{t+1} - \left[\frac{Q_0(1-\mu)(1-\lambda^2)\lambda}{1-\lambda L}\right]\varepsilon_{t+1}\}$$
$$+(1-\mu)L^{-2}[(L-\lambda)Q(L) - \{Q_0 + (Q_1 - \lambda Q_0)L\}\mathcal{B}_\lambda(L)]\varepsilon_{t+1}$$
$$\overline{\mathbb{E}}_{t+1}\overline{\mathbb{E}}_{t+2}p_{t+3} = p_{t+3} + \mu\{\lambda Q_0 - (Q_0 - \lambda Q_1)L\}\varepsilon_{t+3} - \mu\left[\frac{Q_0(1-\mu)(1-\lambda^2)\lambda}{1-\lambda L}\right]\varepsilon_{t+1}$$
$$-(1-\mu)[\{Q_0 + (Q_1 - \lambda Q_0)L\}\mathcal{B}_\lambda(L)]\varepsilon_{t+3} \tag{7.13}$$

It is obvious again that the uninformed's expectations of (7.13) are just

$$\mathbb{E}_t^U[\overline{\mathbb{E}}_{t+1}\overline{\mathbb{E}}_{t+2}p_{t+3}] = \mathbb{E}_t^U p_{t+3} \tag{7.14}$$

This is because the uninformed cannot forecast the forecast errors of the informed and

$$\mathbb{E}_t^U\left[\frac{\kappa}{1-\lambda L}\right]\varepsilon_{t+1} = \mathbb{E}_t^U\left[\frac{\kappa}{L-\lambda}\right]e_{t+1} = \kappa\mathbb{E}_t^U\sum_{j=0}^{\infty}\lambda^j e_{t+2+j} = 0$$

where $\kappa = \mu Q_0(1-\mu)(1-\lambda^2)\lambda$.

For the informed

$$\mathbb{E}_t^I[\overline{\mathbb{E}}_{t+1}\overline{\mathbb{E}}_{t+2}p_{t+3}] = \mathbb{E}_t^I p_{t+3} - \kappa\mathbb{E}_t^I(1-\lambda L)^{-1}\varepsilon_{t+1} - (1-\mu)\mathbb{E}_t^I\Gamma(L)\varepsilon_{t+3}$$
$$\mathbb{E}_t^I[\overline{\mathbb{E}}_{t+1}\overline{\mathbb{E}}_{t+2}p_{t+3}] = \mathbb{E}_t^I p_{t+3} - \left[\frac{Q_0\mu(1-\mu)(1-\lambda^2)\lambda^2}{1-\lambda L}\right]\varepsilon_t - \left[\frac{Q_1(1-\mu)(1-\lambda^2)\lambda}{1-\lambda L}\right]\varepsilon_t$$
$$\mathbb{E}_t^I[\overline{\mathbb{E}}_{t+1}\overline{\mathbb{E}}_{t+2}p_{t+3}] = \mathbb{E}_t^I p_{t+3} - (1-\mu)(1-\lambda^2)\left[\frac{Q_0\mu\lambda^2 + Q_1\lambda}{1-\lambda L}\right]\varepsilon_t$$

Therefore the average expectation is

$$\overline{\mathbb{E}}_t\overline{\mathbb{E}}_{t+1}\overline{\mathbb{E}}_{t+2}p_{t+3} = \overline{\mathbb{E}}_t p_{t+3} - (1-\mu)(1-\lambda^2)\left[\frac{Q_0\mu^2\lambda^2 + Q_1\mu\lambda}{1-\lambda L}\right]\varepsilon_t \tag{7.15}$$

compare to

$$\overline{\mathbb{E}}_t\overline{\mathbb{E}}_{t+1}p_{t+2} = \overline{\mathbb{E}}_t(p_{t+2}) - (1-\mu)(1-\lambda^2)\left(\frac{Q_0\mu\lambda}{1-\lambda L}\right)\varepsilon_t$$

By induction, we are converging to

$$\overline{\mathbb{E}}_t\overline{\mathbb{E}}_{t+1}\cdots\overline{\mathbb{E}}_{t+j}p_{t+j+1} = \overline{\mathbb{E}}_t p_{t+j+1} - (1-\mu)(1-\lambda^2)\left(\frac{\sum_{i=1}^{j}(\mu\lambda)^i Q_{j-i}}{1-\lambda L}\right)\varepsilon_t$$

Proposition 2

To be added.

Theorem 3

The first step in the proof is to obtain a representation for the signal vector (ε_{it}, p_t) that can be used to formulate the expectation at the agent's level. The representation in terms of the innovation ε_t and the noise v_{it} is

$$\begin{pmatrix} \varepsilon_{it} \\ p_t \end{pmatrix} = \begin{pmatrix} \sigma_\varepsilon & \sigma_v \\ (L-\lambda)\,p\,(L) & 0 \end{pmatrix}\begin{pmatrix} \hat{\varepsilon}_t \\ \hat{v}_{it} \end{pmatrix} = \Gamma(L)\begin{pmatrix} \hat{\varepsilon}_t \\ \hat{v}_{it} \end{pmatrix}. \tag{7.16}$$

where we have re-scaled the mapping so that the innovations $\hat{\varepsilon}_t$ and the noise $\hat{v}_{it}$ have unit variance and we have implicitly defined $p(L) = Q(L)\sigma_\varepsilon$. Let the fundamental representation be denoted by

$$\begin{pmatrix} \varepsilon_{it} \\ p_t \end{pmatrix} = \Gamma^*(L) \begin{pmatrix} w_{it}^1 \\ w_{it}^2 \end{pmatrix}. \tag{7.17}$$

The lag polynomial matrix $\Gamma^*(L)$ is given by (see Rondina (2009))

$$\Gamma^*(L) = \Gamma(L) W_\lambda B_\lambda(L)$$

where

$$W_\lambda = \frac{1}{\sqrt{\sigma_\varepsilon^2 + \sigma_v^2}} \begin{pmatrix} \sigma_\varepsilon & -\sigma_v \\ \sigma_v & \sigma_\varepsilon \end{pmatrix} \quad \text{and} \quad B_\lambda(L) = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1-\lambda L}{L-\lambda} \end{pmatrix}.$$

The vector of fundamental innovations is then given by

$$\begin{pmatrix} w_{it}^1 \\ w_{it}^2 \end{pmatrix} = B_\lambda(L^{-1}) W_\lambda^T \begin{pmatrix} \hat{\varepsilon}_t \\ \hat{v}_{it} \end{pmatrix}.$$

The expectation term for agent i is provided by the second row of the Wiener-Kolmogorov prediction formula applied to the fundamental representation (7.17), which is

$$\mathbb{E}(p_{t+1}|\varepsilon_i^t, p^t) = \left[\Gamma_{21}^*(L) - \Gamma_{21}^*(0)\right] L^{-1} w_{it}^1 + \left[\Gamma_{22}^*(L) - \Gamma_{22}^*(0)\right] L^{-1} w_{it}^2. \tag{7.18}$$

It is straightforward to show that

$$\begin{aligned} \Gamma_{21}^*(L) &= (L-\lambda)\, p(L) \tfrac{\sigma_\varepsilon}{\sqrt{\sigma_\varepsilon^2+\sigma_v^2}}, \quad \Gamma_{21}^*(0) = -\lambda p_0 \tfrac{\sigma_\varepsilon}{\sqrt{\sigma_\varepsilon^2+\sigma_v^2}} \\ \Gamma_{22}^*(L) &= -(1-\lambda L)\, p(L) \tfrac{\sigma_v}{\sqrt{\sigma_\varepsilon^2+\sigma_v^2}}, \quad \Gamma_{22}^*(0) = -p_0 \tfrac{\sigma_v}{\sqrt{\sigma_\varepsilon^2+\sigma_v^2}}. \end{aligned}$$

Solving for the equilibrium price requires averaging across all the agents. In taking those averages, the idiosyncratic components of the innovation (the noise) will be zero and one would just have two terms that are function only of the aggregate innovation, namely

$$\int_0^1 w_{it}^1 di = w_t^1 = \tfrac{\sigma_\varepsilon}{\sqrt{\sigma_\varepsilon^2+\sigma_v^2}} \hat{\varepsilon}_t \quad \text{and} \quad \int_0^1 w_{it}^2 di = w_t^2 = -\tfrac{\sigma_v}{\sqrt{\sigma_\varepsilon^2+\sigma_v^2}} \frac{L-\lambda}{1-\lambda L} \hat{\varepsilon}_t.$$

The average market expectation is then

$$\bar{\mathbb{E}}(p_{t+1}) = \left[(L-\lambda)p(L) + \lambda p_0\right] L^{-1} \tfrac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2+\sigma_v^2} \hat{\varepsilon}_t + \left[(1-\lambda L)p(L) - p_0\right] L^{-1} \tfrac{\sigma_v^2}{\sigma_\varepsilon^2+\sigma_v^2} \frac{L-\lambda}{1-\lambda L} \hat{\varepsilon}_t. \tag{7.19}$$

Now, if we let

$$\mu \equiv \tfrac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2+\sigma_v^2},$$

and we substitute the functional form of the average expectations into the equilibrium equation for p_t we would get

$$(L-\lambda)p(L) = \beta\mu L^{-1}[(L-\lambda)p(L) + \lambda p_0] + \beta(1-\mu)L^{-1}[(L-\lambda)p(L) - p_0 \frac{L-\lambda}{1-\lambda L}] + A(L)\sigma_\varepsilon$$

which is equivalent to (7.3) since $p(L) = Q(L)\sigma_\varepsilon$. The rest of the proof follows the same lines of Theorem 2. For the sake of completeness, we need to show that, for the dispersed information case, the information conveyed by the knowledge of the model is consistent with the information used in the expectational equation for agent i presented above. Such knowledge can be represented by the variable

$$m_{it} \equiv p_t - \beta E\left(p_{t+1}|\varepsilon_i^t, p^t\right) = \beta \left(\ \bar{E}\left(p_{t+1}\right) - E\left(p_{t+1}|\varepsilon_i^t, p^t\right) \ \right) + s_t.$$

we then need to show that the fundamental representation of the signal vector $(\varepsilon_{it}, p_t, m_{it})$ is the same as the one we derived above. Essentially, we need to show that the mapping between this enlarged vector of signal and the vector of structural innovation is still of rank 1 at $L=\lambda$. Using the result in Corollary 3 to write down the explicit form of the difference between the individual expectations and the average market expectations, the mapping of interest is

$$\begin{pmatrix} \varepsilon_{it} \\ p_t \\ m_{it} \end{pmatrix} = \begin{pmatrix} \sigma_\varepsilon & \sigma_v \\ (L-\lambda)\, p(L) & 0 \\ A(L)\,\sigma_\varepsilon & \frac{\sigma_\varepsilon \sigma_v}{\sigma_\varepsilon^2+\sigma_v^2}\left(\frac{1-\lambda^2}{1-\lambda z}\right)\beta p_0 \end{pmatrix} \begin{pmatrix} \hat{\varepsilon}_t \\ \hat{v}_{it} \end{pmatrix}. \tag{7.20}$$

It is straightforward to show that 2 of the 3 minors of this matrix have rank 1 at $L=\lambda$. For the third minor the condition for rank 1 is

$$\frac{\sigma_\varepsilon \sigma_v}{\sigma_\varepsilon^2+\sigma_v^2}\left(\frac{1-\lambda^2}{1-\lambda z}\right)\sigma_\varepsilon \beta p_0 - A(L)\,\sigma_\varepsilon \sigma_v = 0 \quad \text{at} \quad L=\lambda.$$

Using the fact that $p_0 = Q_0 \sigma_\varepsilon$ one can immediately see that this condition is equivalent to (5.2), hence, in a dispersed information equilibrium, it is always true that the enlarged information matrix (7.20) carries the same information as the information matrix (7.16). This completes the proof of Theorem 3.

Proposition 3

Substituting $\Gamma_{21}^*(L)$ and $\Gamma_{22}^*(L)$ into (7.24) and collecting the terms that constitute (7.19), one gets

$$\begin{aligned} \mathbb{E}(p_{t+1}|\varepsilon_i^t, p^t) &= \bar{\mathbb{E}}(p_{t+1}) + \frac{\sigma_\varepsilon}{\sigma_\varepsilon^2+\sigma_v^2} L^{-1}[(L-\lambda)p(L) + \lambda p_0 - (L-\lambda)p(L) + p_0 \frac{L-\lambda}{1-\lambda L}]\sigma_v \hat{v}_{it} \\ &= \bar{\mathbb{E}}(p_{t+1}) + \frac{\sigma_\varepsilon}{\sigma_\varepsilon^2+\sigma_v^2} L^{-1}[\lambda p_0 + p_0 \frac{L-\lambda}{1-\lambda L}]\sigma_v \hat{v}_{it} \\ &= \bar{\mathbb{E}}(p_{t+1}) + \mu Q_0 \frac{1-\lambda^2}{1-\lambda L} v_{it}, \end{aligned} \tag{7.21}$$

which completes the proof.

Proposition 4

The notation of the proof is that of Theorem 3 unless otherwise specified. The crucial step in the proof is to show that

$$\mathbb{E}\left(\frac{L-\lambda}{1-\lambda L}\varepsilon_{t+2}|\varepsilon_i^t, p^t\right) = \mu\lambda \frac{(1-\lambda^2)}{1-\lambda L}\varepsilon_{it}. \tag{7.22}$$

where $\mu \equiv \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2+\sigma_v^2}$. Let $B(L) = \frac{L-\lambda}{1-\lambda L}$ and define

$$y_t = B(L)\,\varepsilon_t. \tag{7.23}$$

then we look for $\mathbb{E}(y_{t+2}|\varepsilon_i^t, p^t) = \pi_1(L)\,\varepsilon_{it} + \pi_2(L)\,p_t$. Following Theorem 1 in Rondina (2009) we know that

$$\begin{bmatrix} \pi_1(L) & \pi_2(L) \end{bmatrix} = \left[L^{-2} g_{y,(\varepsilon,p)}(L)\left(\Gamma^*(L^{-1})^T\right)^{-1}\right]_+ \Gamma^*(L)^{-1}$$

where $\Gamma^*(L)$ and $\left(w_{it}^1, w_{it}^2\right)$ are defined in (7.17) and $g_{y,(\varepsilon,p)}(L)$ is the variance-covariance generating function between the variable to be predicted and the variables in the information set. In our case we have that

$$g_{y,(\varepsilon,p)}(L) = \begin{bmatrix} B(L)\,\sigma_\varepsilon^2 & B(L)\left(L^{-1}-\lambda\right) p\left(L^{-1}\right)\sigma_\varepsilon \end{bmatrix}.$$

Plugging in the explicit forms and solving out the algebra

$$L^{-2} g_{y,(\varepsilon,p)}(L)\left(\Gamma^*(L^{-1})^T\right)^{-1} = \frac{1}{\sqrt{\sigma_\varepsilon^2+\sigma_v^2}} \begin{bmatrix} L^{-2}\frac{L-\lambda}{1-\lambda L}\sigma_\varepsilon^2 + L^{-2}\left(L^{-1}-\lambda\right) p\left(L^{-1}\right)\frac{\sigma_\varepsilon^2}{\sigma_v} & -L^{-2}\frac{\sigma_\varepsilon^2+\sigma_v^2}{\sigma_v}\sigma_\varepsilon \end{bmatrix}.$$

Applying the annihilator operator to the RHS we see that the second term of the vector goes to zero. For the first term, the assumption that $p(L)$ is analytic inside the unit circle ensures that $L^{-2}\left(L^{-1}-\lambda\right)p\left(L^{-1}\right)$ does not contain any term in positive power of L. We are then left with

$$\left[L^{-2}\frac{L-\lambda}{1-\lambda L}\right]_{+}=\frac{\lambda\left(1-\lambda^{2}\right)}{1-\lambda L}, \tag{7.24}$$

as shown in section 4.1. Summarizing we have shown that

$$\left[\begin{array}{cc}\pi_1(L) & \pi_2(L)\end{array}\right]=\frac{1}{\sqrt{\sigma_\varepsilon^2+\sigma_v^2}}\left[\begin{array}{cc}\frac{\lambda(1-\lambda^2)}{1-\lambda L}\sigma_\varepsilon^2 & 0\end{array}\right]\Gamma^*(L)^{-1}.$$

Notice that

$$\Gamma^*(L)^{-1}\left[\begin{array}{c}\varepsilon_{it}\\ p_t\end{array}\right]=\left[\begin{array}{c}w_{it}^1\\ w_{it}^2\end{array}\right]$$

so that

$$\mathbb{E}\left(y_{t+2}|\varepsilon_i^t,p^t\right)=\left[\begin{array}{cc}\pi_1(L) & \pi_2(L)\end{array}\right]\left[\begin{array}{c}\varepsilon_{it}\\ p_t\end{array}\right]=\frac{1}{\sqrt{\sigma_\varepsilon^2+\sigma_v^2}}\frac{\lambda\left(1-\lambda^2\right)}{1-\lambda L}\sigma_\varepsilon^2 w_{it}^1.$$

From the proof of Theorem 3 we know that $w_{it}^1=\frac{1}{\sqrt{\sigma_\varepsilon^2+\sigma_v^2}}\left(\varepsilon_t+v_{it}\right)$, which, once substituted in the above expression completes the proof.

References

Allen, F., Morris, S. and Shin, H. (2006). Beauty contests and iterated expectations in asset markets. *Review of Financial Studies*, **19** (3), 719–752.

Angeletos, G. and La'O, J. (2009). Noisy business cycles, mIT Working Paper.

— and Pavan, A. (2007). Efficient use of information and social value of information. *Econometrica*, **75** (4), 1103–1142.

— and — (2009). Policy with Dispersed Information: Policy with Dispersed Information. *Journal of the European Economic Association*, **7** (1), 11–60.

Bacchetta, P. and van Wincoop, E. (2004). Higher order expectations in asset pricing, working Paper, Studienzentrum Gerzensee.

— and — (2006). Can information heterogeneity explain the exchange rate puzzle? *American Economic Review*, **96** (3), 552–576.

Bernhardt, D., Seiler, P. and Taub, B. (2009). Speculative dynamics, forthcoming, *Economic Theory*.

— and Taub, B. (2008). Cross-Asset Speculation in Stock Markets. *Journal of Finance*, **63** (5), 2385–2427.

Blanchard, O. J. and Kahn, C. M. (1980). The solution of linear difference models under rational expectations. *Econometrica*, **48** (5), 1305–1312.

Cogley, T. and Nason, J. (1995). Output Dynamics in Real-Business-Cycle Models. *The American Economic Review*, **85** (3), 492–511.

Futia, C. A. (1981). Rational expectations in stationary linear models. *Econometrica*, **49** (1), 171–192.

Gregoir, S. and Weill, P. (2007). Restricted perception equilibria and rational expectation equilibrium. *Journal of Economic Dynamics and Control*, **31** (1), 81–109.

Hansen, L. P. and Sargent, T. J. (1980). Formulating and estimating dynamic linear rational expectations models. *Journal of Economic Dynamics and Control*, **2**, 7–46.

— and — (1991). Two difficulties in interpreting vector autoregressions. In L. P. Hansen and T. J. Sargent (eds.), *Rational Expectations Econometrics*, Westview Press.

Kasa, K. (2000). Forecasting the forecasts of others in the frequency domain. *Review of Economic Dynamics*, **3**, 726–756.

—, Walker, T. B. and Whiteman, C. H. (2008). Asset prices in a time series model with perpetually disparately informed, competitive traders, working paper.

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

King, R. (1982). Monetary Policy and the Information Content of Prices. *Journal of Political Economy*, **90** (2), 247–279.

Lippi, M. and Reichlin, L. (1994). Var analysis, nonfundamental representations, blaschke matrices. *Journal of Econometrics*, **63**, 307–325.

Lorenzoni, G. (2009). A theory of demand shocks, forthcoming, *American Economic Review*.

Lucas, R. E., Jr. (1972). Expectations and the neutrality of money. *Journal of Economic Theory*, **4**, 103–124.

— (1975). An equilibrium model of the business cycle. *Journal of Political Economy*, **83**, 1113–1144.

Mankiw, N. and Reis, R. (2002). Sticky Information Versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve. *Quarterly Journal of Economics*, **117** (4), 1295–1328.

Morris, S. and Shin, H. S. (2002). The social value of public information. *American Economic Review*, **92**, 1521–1534.

Nimark, K. (2007). Dynamic higher order expectations, working paper.

Nimark, K. P. (2005). Dynamic pricing, imperfect common knowledge and inflation inertia, eCB Working Paper No. 474.

Pearlman, J. G. and Sargent, T. J. (2005). Knowing the forecasts of others. *Review of Economic Dynamics*, **8** (2), 480–497.

Phelps, E. (1969). The New Microeconomics in Inflation and Employment Theory. *American Economic Review*, **59** (2), 147–160.

Pigou, A. C. (1929). *Industrial Fluctuations*. London: Macmillan, 2nd edn.

Radner, R. (1979). Rational Expectations Equilibrium: Generic Existence and the Information Revealed by Prices. *Econometrica*, **47** (3), 655–678.

Rondina, G. (2009). Incomplete information and informative pricing, working Paper. UCSD.

— and Walker, T. B. (2009). Solving for information equilibria, working Paper.

Sargent, T. (1981). Interpreting economic time series. *The Journal of Political Economy*, pp. 213–248.

Sargent, T. J. (1987). *Macroeconomic Theory*. Academic Press, 2nd edn.

— (1991). Equilibrium with signal extraction from endogenous variables. *Journal of Economic Dynamics and Control*, **15**, 245–273.

Seiler, P. and Taub, B. (2008). The Dynamics of Strategic Information Flows in Stock Markets. *Finance and Stochastics*, **12** (1), 43–82.

Sims, C. A. (2002). Solving Linear Rational Expectations Models. *Computational Economics*, **20** (1), 1–20.

Singleton, J., Kenneth (1987). Asset prices in a time series model with disparately informed, competitive traders. In W. Barnett and K. Singleton (eds.), *New Approaches to Monetary Economics*, Cambridge: Cambridge University Press.

Taub, B. (1989). Aggregate Fluctuations as an Information Transmission Mechanism. *Journal of Economic Dynamics and Control*, **13** (1), 113–150.

Townsend, R. M. (1983). Forecasting the forecasts of others. *Journal of Political Economy*, **91**, 546–588.

Walker, T. B. (2007). How equilibrium prices reveal information in time series models with disparately informed, competitive traders. *Journal of Economic Theory*, **137** (1), 512–537.

Whiteman, C. (1983). *Linear Rational Expectations Models: A User's Guide*. Minneapolis: University of Minnesota Press.

Woodford, M. (2003). Imperfect common knowledge and the effects of monetary policy. In P. Aghion, R. Frydman, J. Stiglitz and M. Woodford (eds.), *Knowledge, Information, and Expectations in Modern Macroeconomics*, Princeton, N.J.: Princeton University Press.