Information Equilibria in Dynamic Economies with Dispersed Information*

Giacomo Rondina[†] Todd B. Walker[‡]

First version: March 2009; This version: December 2011

Abstract

We study Rational Expectations equilibria in dynamic models with dispersed information and signal extraction from endogenous variables. Existence and uniqueness conditions for a new class of rational expectations equilibria in economies with dispersed information are established. The novelty of this class lies in the presence of confounding dynamics in the equilibrium process that can permanently sustain the information dispersion across agents, even when the equilibrium process is perfectly observed. A feature of the equilibria belonging to this class is a dynamic response of endogenous variables to economic shocks that display waves of optimism and pessimism that are not present in the full information counterpart. We derive an analytical characterization of the equilibria that generalizes the celebrated Hansen-Sargent optimal prediction formula, and also allows us to study higher-order beliefs representations. We show that the higher-order belief dynamics, contrary to what is normally believed, can generate a positive effect on information diffusion: if dispersedly informed agents were not engaging in formulating expectations about expectations about expectations and so on, information transmission through equilibrium prices would be reduced.

Keywords: Dispersed Information, Incomplete Information, Rational Expectations Equilibrium, Higher-Order Beliefs

*We would like to thank Manuel Amador, Marios Angeletos, Tim Cogley, Jesús Fernández-Villaverde, Ken Kasa, Jennifer La'O, Eric Leeper, Kristoffer Nimark, Gary Ramey, Tom Sargent, Martin Schneider, Bart Taub, Venky Venkateswaran, Pierre-Olivier Weill, Mirko Wiederholt, Chuck Whiteman and seminar participants at UC–Berkeley, UC–Santa Cruz, New York University, University of Pennsylvania, University of Texas–Austin, Stanford University, St. Louis Federal Reserve, the Econometric Society 2009 summer meetings, the 2009 Yale University Cowles Foundation Summer Conference on Information and Beliefs in Macroeconomics, the 2011 LAEF Spring Conference, and Information, the 2011 CREI conference on Beliefs and Expectations in Macroeconomics for useful comments. We also acknowledge financial support from the National Science Foundation under grant number SES–0962221.

[†]UCSD, grondina@ucsd.edu
[‡]Indiana University, walkertb@indiana.edu

1 Introduction

Dynamic models with dispersed information are becoming increasingly prominent in several literatures such as asset pricing, optimal policy communication, international finance, and business cycles.[1] The role of incomplete information in many of these settings was acknowledged very early on; Keynes (1936) argued that higher-order expectations played a fundamental role in asset markets, while Pigou (1929) advanced the idea that business cycles may be the consequence of "waves of optimism and pessimism" that originate in markets where agents, by observing common signals, generate correlated forecast errors. 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).

From this early literature it was immediately clear that solving for equilibria in dynamic models with incomplete information would be challenging. Sargent (1991) and 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 associated with solving for equilibrium, even though these models harbored much potential. The primary difficulty is that when endogenous variables transmit information, the equilibrium fixed point problem typical of the rational expectations paradigm involves a mapping from endogenous variables to the agents' information sets: given the equilibrium obtained under the expectations specified for a given information set, the information revealed in equilibrium should be consistent with the information used to solve for the equilibrium. In dynamic settings with incomplete information, this fixed point condition is nontrivial and a crucial aspect of the equilibrium.

We develop an equilibrium concept, which we refer to as an "Information Equilibria" (IE), that explicitly accounts for this fixed point condition, and yields existence and uniqueness conditions for rational expectations models with dispersed information. In particular, we focus on a new class of rational expectations equilibria in economies with dispersed information. The novelty of this class lies in the presence of confounding dynamics in the equilibrium process that can permanently sustain the information dispersion across agents, even when the equilibrium process is perfectly observed. We derive an analytical characterization of the equilibria that generalizes the celebrated Hansen-Sargent optimal prediction formula, and also allows us to study higher-order beliefs representations.

We develop our key existence, uniqueness and characterization results for models with dispersed information in several steps. We do this for two reasons: first, each step has value on its own in terms of possible applications, and second, decomposing the key result into steps allows us to obtain some crucial insights on the workings of information interactions when information is dispersed. The key steps are as follows. First, we begin by

[1] The literature is too voluminous to cite every worthy paper. Recent examples include: Morris and Shin (2002), Woodford (2003), Pearlman and Sargent (2005), Allen, Morris, and Shin (2006), Bacchetta and van Wincoop (2006), Hassan and Mertens (2011), Hellwig (2006), Gregoir and Weill (2007) Angeletos and Pavan (2007), Kasa, Walker, and Whiteman (2008), Lorenzoni (2009), Rondina (2009), Angeletos and La'O (2009b), Hellwig and Venkateswaran (2009).

deriving the existence and characterization conditions for the dispersed informational setup (Theorem 1). Next, we introduce an arbitrary fraction of agents that are perfectly informed about the current and past state of the fundamentals. We then show that the information equilibrium characterized under the assumption of some agents being perfectly informed is equivalent to the aggregate representation of the "dispersed information" case in which every agent receives a privately observed noisy signal about the state of the market fundamentals, together with the equilibrium price (Theorem 2). The equivalence holds once the parameter measuring the proportion of agents perfectly informed in Theorem 2 is reinterpreted as the signal-to-noise ratio of the privately observed signal of Theorem 1. This equivalence result stems from the optimal signal extraction of dispersedly informed agents that consists of a mixing strategy in interpreting the information available to them. With some probability agents will act as if their signal is exactly correct, mimicking thus the behavior of the perfectly informed agents. With the complementary probability they will act as if their signal contains no information about the state and so they will take into account only the information from the equilibrium price, thus mimicking the other fraction of agents.

Equipped with the analytical characterization of the market equilibria under dispersed information, we are able to characterize the higher-order belief (HOB) representation of such equilibria and study the role of higher order thinking in shaping the market price dynamics. Recent papers have emphasized the role of HOB dynamics and the subsequent breakdown in the law of iterated expectations with respect to the average expectations operator in models with asymmetric information [e.g., Allen, Morris, and Shin (2006), Bacchetta and van Wincoop (2006), Nimark (2008), Pearlman and Sargent (2005), Angeletos and La'O (2009a)]. Many resort to numerical analysis or truncation of the state space in demonstrating the dynamic case, making it difficult to isolate the specific role played by HOBs. With an analytical solution in hand, we are able to characterize these objects in closed form and show precisely *why* HOBs exist, and *why* and *when* HOBs imply a failure of the law of iterated expectations. In addition, it is possible to relate the formation of HOBs to the transmission of information in equilibrium by showing that the formation of HOBs increases the information impounded into endogenous variables. This, in turn, leads to a decrease in the variance of prediction errors. In other words, forming HOBs gives rise to a positive effect on information diffusion. This conclusion goes against the existing conjecture that HOBs are responsible for the slow reaction of endogenous variables to structural shocks. This idea stems from the observation that agents forming HOBs forecast the forecast errors of uninformed agents, thereby injecting additional persistence through the higher-order expectations. However, we find that this observation is incomplete as it does not take into account the effect of higher order thinking upon informational transmission. Once the both effects are considered, the latter one always dominates in our setting, and thus HOB formation *always* improves information in equilibrium, which in turn actually *reduces* the persistence in equilibrium.

A remarkable feature of the equilibria belonging to this class of models is that the market price can display continuously oscillating overpricing and underpricing compared to the market price that would emerge under complete information. This property pertains to a rational expectations equilibrium and is not the result of bounded rationality or ad-hoc learning. We show that this propagation stems from the dynamic signal extraction undertaken by market participants. To the best of our knowledge, this result is new to the rational expectations literature. We argue that this feature of the equilibrium makes models with incomplete information empirically more relevant than their complete information counterpart.

In order to make the derivation of our results as transparent as possible, we focus our attention on a simple forward-looking asset pricing framework. Such a framework is, nonetheless, flexible enough to encompass the key dynamic equations of many standard macroeconomic settings. Our results are therefore generally applicable to any dynamic model of higher economic complexity.[2]

2 Information Equilibrium: Preliminaries

This section establishes notation and lays important groundwork for interpreting the equilibrium characterizations that follow.

2.1 Equilibrium Model

To fix notation and ideas, we define an information equilibrium within a generic linear rational expectations framework. The forward-looking nature of the key equilibrium relationship is quite flexible in that it allows for a broad range of interpretation, so that our results apply to any setting where current variables depend on the expectations of future variables.

2.1.1 Market

In order to keep things grounded in a specific economic example, we interpret our equations as arising from the perfectly competitive equilibrium of an asset market in which investors take positions on a risky asset to maximize the expected utility of next period wealth.[3] The asset market works as follows: investors submit their demand schedules—a mapping that associates the asset price to net demand—to a Walrasian auctioneer. The auctioneer collects the demand schedules and then calls the price that equates demand to supply. To allow for trading in equilibrium, the net supply of the asset in a given period t, s_t, is assumed to be exogenous.[4] The net demand in the asset market is provided by a continuum of potentially diversely informed agents indexed by i. The market clearing price chosen by the Walrasian auctioneer is given by

$$p_t = \beta \int_0^1 \mathbb{E}_t^i p_{t+1} \phi(i) di + s_t \tag{2.1}$$

[2]For example, see Rondina and Walker (2011) for an application of our methods to a standard real business cycle model.

[3]In Appendix B we present a simple asset demand model that delivers the equilibrium equation that we use throughout the paper.

[4]In what follows we will let the supply of the asset be measured by $-s_t$. Therefore, an increase (decrease) in s_t will correspond to a decrease (increase) in the exogenous supply of the asset.

where $\beta \in (0,1)$, $\mathbb{E}_t^i$ is the conditional expectation of agent i, $\phi(\cdot)$ is the density of agents and the exogenous process (s_t) is driven by a Gaussian shock

$$s_t = A(L)\varepsilon_t, \qquad \varepsilon_t \overset{iid}{\sim} N(0, \sigma_\varepsilon^2) \tag{2.2}$$

and where $A(L)$ is assumed to be a square-summable polynomial in non-negative powers of the lag operator L.

2.1.2 Information 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 and is endowed by the modeler to the agents. Thus, the exogenous information profile $\{U_t^i, i \in [0,1]\}$ is a primitive of the model. Endogenous information is generated through market interactions. When agents are diversely informed, endogenous variables may convey additional information not already 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 linear closed subspace that is spanned by current and past p_t, we refer to it as "time-series information" of p_t. $\mathbb{M}_t(p)$, on the other hand, results from the assumption that agents know the equilibrium process p_t evolves according to (2.1); we refer to it as "*information from the model*."

To clarify what information is captured in $\mathbb{M}_t(p)$, it is useful to think about how the the knowledge of the model (2.1) affects the Walrasian market structure described above. When rational investors formulate the demand schedule to submit to the Walrasian auctioneer, they know that the auctioneer will pick a price that clears the market, i.e. that satisfies (2.1). Investors can use this information to reduce their forecast errors. To see this, suppose that all the investors have the same information and thus the individual demand schedule is given by $\beta\mathbb{E}_t p_{t+1} - p_t$, for some arbitrary information set. Given a candidate price p_t chosen by the auctioneer, investors know that at that price the market will clear, which means $\beta\mathbb{E}_t p_{t+1} - p_t + s_t = 0$. If this is the case, then investors will treat s_t as part of the information that they should use to derive $\mathbb{E}_t p_{t+1}$ for any arbitrary p_t. As investors submit their demand schedules they do not know what is the true value of s_t but they can formulate expectations that are consistent with the true value that will be revealed once the Walrasian auctioneer picks the market clearing price. If investors ignored this information, they would incur consistently higher forecast errors, which would violate rational expectations and imply their submitted demand schedules were not optimal. That subjective beliefs must be model consistent is a standard definition of a rational expectations equilibrium.[5] In rational expectations models with complete information and representative agents, information from the model is a trivial equilibrium condition. We show below that in models with incomplete information and heterogeneously informed agents, information from the model plays a crucial role in determining equilibrium.

[5]From a mere statistical point of view, the knowledge of the model is equivalent to the knowledge of the covariance generating function between the process s_t and the equilibrium price p_t. In other words, in equilibrium there is a true relationship between prices and supply that is summarized by the variance-covariance generating matrix $\begin{pmatrix} g_{pp}(z) & g_{ps}(z) \\ g_{ps}(z) & g_{ss}(z) \end{pmatrix}$. Knowledge of the model corresponds to knowing $g_{ps}(z)$ and using it to obtain s_t from p_t.

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.[6] Uncertainty is assumed to be driven entirely by the Gaussian stochastic process ε_t, which implies that optimal projection formulas are equivalent to conditional expectations,

$$\mathbb{E}_t^i(p_{t+1}) = \Pi[p_{t+1}|\Omega_t^i] = \Pi[p_{t+1}|U_t^i \vee \mathbb{V}_t(p) \vee \mathbb{M}_t(p)]. \tag{2.3}$$

where Π denotes linear projection. The normality assumption also 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.

2.1.3 Equilibrium Definition

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 $\left\{\Omega_t^i, i \in [0,1]\right\}$ such that: (i) each agent i, given the price and the information set, optimally forms expectations according to (2.3); (ii) p_t satisfies the equilibrium condition (2.1).*

An IE consists of two objects, a *price* and a *distribution of information*, and 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.

2.2 Confounding Dynamics and Signal Extraction

Central to the existence of the class of rational expectations equilibria examined in this paper is the idea that dynamics can conceal information. In this section we lay some groundwork on the relationship between the dynamics of a stochastic process and the information conveyed by that process. We isolate a signal extraction mechanism that operates at the heart of the new class of equilibria established in Section 3; this will allow us to gain insights in the interpretation of the equilibrium dynamics.

In dynamic settings, the information set of agents is continuously expanding as they collect new observations with each period t. A crucial question in such settings is whether an expanding information set over time corresponds to an ever increasing precision of information about the current and past structural innovations, $\{\varepsilon_{t-j}\}_{j=0}^{\infty}$.

The answer to this question depends upon the characteristics of the dynamics of the observed variables. Using the terminology of Rozanov (1967), if the structural innovations are *fundamental* for the observable variables,

[6] 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 $\{y_{t-j}\}_{j=0}^{\infty}$.

then agents would eventually learn the true underlying dynamics. Intuitively, if a dynamic stochastic process is invertible in current and past observables, then it is fundamental and the observed history would allow one to back out the exact history of the underlying fundamental innovations. On the other hand, if the process is non-fundamental, then the observed history will contain only imperfect information about the fundamentals. In this case we say that the observed variable displays *confounding dynamics*. In linear dynamic settings, confounding dynamics can be formalized by non-fundamental moving averages (MA) representations.

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

$$x_t = -\lambda\varepsilon_t + \varepsilon_{t-1}. \tag{2.4}$$

If $|\lambda| \geq 1$, the stochastic process x_t is invertible in current and past x_t, which means that there exists a linear combination of current and past x_t's that allows the exact recovery of ε_t; formally

$$\mathbb{E}\left(\varepsilon_t|x^t\right) = -1/\lambda\left(x_t + \lambda^{-1}x_{t-1} + \lambda^{-2}x_{t-2} + \lambda^{-3}x_{t-3} + ...\right) = \varepsilon_t. \tag{2.5}$$

Note that the infinite sum converges as λ^{-j} goes to zero for j "big enough".

When $|\lambda| < 1$ the process is no longer invertible in current and past x_t. Equation (2.5) 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

$$\mathbb{E}\left(\varepsilon_t|x^t\right) = -\frac{\lambda}{|\lambda|}\left(x_t + \lambda x_{t-1} + \lambda^2 x_{t-2} + \lambda^3 x_{t-3} + ...\right) = \tilde{\varepsilon}_t. \tag{2.6}$$

Non-invertibility implies that $\tilde{\varepsilon}_t$ 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, the mean square forecast error is

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

The mean squared forecast error approaches zero as the dynamics goes from non-invertible to invertible, i.e. as $|\lambda| \to 1$ from below.

The imperfect information described by (2.4) when $|\lambda| < 1$ corresponds to an ignorance about the initial state of the world at time $t = 0$ that *never* unravels because of the confounding dynamics of x_t. To see this, imagine that agents initially observe $x_1 = -\lambda\varepsilon_1 + \varepsilon_0$ and thus cannot distinguish between ε_1 and ε_0. If they knew ε_0 they could easily back out ε_1 from x_1 and then, as information about x_t accumulates, all the values of ε_t for $t > 1$ would be known. However, if all agents observe is x_0, then the best they can do is to get as close as possible to ε_t using (2.6). Whereas in standard signal extraction problems the informational friction is assumed in the form

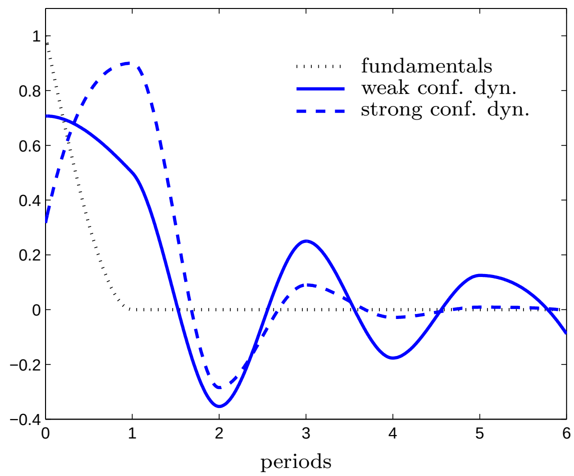

Figure 1: Impulse response of the optimal prediction formula for fundamentals ε_t in presence of confounding dynamics (Equation (2.7)) to a one time innovation $\varepsilon_0 = 1$. The dotted line is the process for fundamentals; the solid line is the response under "weak" confounding dynamics ($|\lambda| = 1/\sqrt{2}$); the dashed line is the response under "strong" confounding dynamics ($|\lambda| = 1/\sqrt{11}$).

of a superimposed signal-to-noise ratio, in (2.4) the noise is a result of the dynamic unfolding parameterized by λ that keeps the ignorance about the initial state ε_0 informationally relevant at any point in time.[7]

An additional important implication of confounding dynamics is that the optimal learning effort of the agents creates a persistent effect of past innovations. To see this let $\lambda < 0$ and rewrite (2.6) as

$$\tilde{\varepsilon}_t = \underbrace{-\lambda\varepsilon_t}_{= \text{ information}} + \underbrace{(1-\lambda^2)[\varepsilon_{t-1} + \lambda\varepsilon_{t-2} + \lambda^2\varepsilon_{t-3} + \cdots]}_{\text{noise from confounding dynamics}}. \tag{2.7}$$

This equation clarifies how λ controls the information that the history of x_t contains about ε_t through two channels: an informative signal with weight λ (the first term on the RHS), and a noise component with weight $(1-\lambda^2)$. Notice that the noise term is a linear combination of past innovations, which is the source of the persistent effect of past innovations. As the confounding dynamics become more pronounced, i.e. when λ decreases, there are three effects. First, the weight on the informative signal decreases as x_t contains less information about ε_t. Second, the weight $(1 - \lambda^2)$ on the noise increases; however, this increase is in part offset by the third effect, which is a reduction in the persistence of innovations dated $t - 2$ and earlier.

To visualize these effects, we report the impulse response function for the prediction equation (2.7) to a one time, one unit increase in ε_t in Figure 1 for both a low and a high value of λ with $\lambda < 0$.[8] First notice that for

[7] As long as $|\lambda| < 1$, whether λ is positive or negative does not matter for the informational content. In Appendix B we show that the signal extraction problem under confounding dynamics is equivalent, in forecast mean square error terms, to a standard signal extraction problem when $\lambda^2 = \tau$, where τ is the signal-to-noise ratio of a standard signal extraction problem. The interested reader is directed to Appendix B for details.

[8] We chose the case of $\lambda < 0$ because the resulting exogenous process lends itself to a meaningful economic interpretation. In fact,

the high-λ case, the value of $\mathbb{E}\left(\varepsilon_t|x^t\right)$ is very close to the true innovation value of 1 on impact, whereas for the low-λ case, the underestimation is quite large. Second, in both cases the current innovation will persistently affect the prediction function several periods beyond impact. This is in contrast to the full information case where the impulse response is zero after impact (fundamentals). However for the weak confounding dynamics, the effect will be initially weaker and then it will only slowly decay. For strong confounding dynamics, the opposite is true: the effect is initially stronger and the decay is subsequently faster.

3 Information Equilibria: Main Theorem

This section establishes the main result of the paper: the existence of a new class of rational expectations equilibria for dynamic economies with dispersed information. We begin by presenting the full information solution to the equilibrium model (2.1) and then we state the main theorem of the paper.

3.1 Full Information Benchmark We define Full Information as the case when every buyer is endowed with perfect knowledge of the innovations history up to time t. Formally

$$U_t^i = \mathbb{V}_t(\varepsilon), \forall i \in [0,1]\,. \tag{3.1}$$

Here, and in the following analysis, we assume that agents always observe the endogenous information $\mathbb{V}_t(p) \vee \mathbb{M}_t(p)$. Under full information all the buyers will have the same information in equilibrium and so (2.1) can be written as the contemporaneous expectation of the discounted sum of future s_t's,

$$p_t = \sum_{j=0}^{\infty} \beta^j \mathbb{E}_t(s_{t+j}). \tag{3.2}$$

The solution of this model is well known and the equilibrium takes the form

$$p_t = \left[\frac{LA(L) - \beta A(\beta)}{L-\beta}\right]\varepsilon_t \tag{3.3}$$

which is the celebrated Hansen-Sargent formula [Hansen and Sargent (1991)]. Provided $|\beta| < 1$, equation (3.3) is the *unique* Information Equilibrium solution to (2.1) when information is specified as (3.1).

3.2 Dispersed Information Equilibrium The principal case of interest is one in which agents are endowed with dispersed information about the economic fundamentals, while they still observe the current and past history

later we will use a process similar to (2.4) to model a canonical S-shaped diffusion process. The prediction formula with $\lambda > 0$ would display the same response at impact but it would not exhibit the oscillatory pattern of Figure 1. Instead, the impulse response would turn negative at period 2 and gradually approaching zero from below from then onward. The three effects described above will all still be present, nonetheless.

of equilibrium prices. This case captures the informational setup of most of the recent literature on equilibrium models with dispersed information.[9] Our theorem therefore presents a class of rational expectations equilibria that can emerge in such models, but that have not been characterized so far.

In line with the dispersed information literature, we assume that all agents are identical in terms of the imperfect quality of information they possess. In particular, we assume each agent observes their own particular "window" of the world, as in Phelps (1969). Information is dispersed in the sense that, although complete knowledge of the fundamentals is not given to any one agent, by pooling the noisy signals across all agents it is possible to recover the full information about the state of the economy. The information set is formalized as follows. Consider a set of i.i.d. noisy signals specified as

$$\varepsilon_{it} = \varepsilon_t + v_{it} \quad \text{with} \quad v_{it} \overset{iid}{\sim} N\left(0, \sigma_v^2\right) \quad \text{for} \quad i \in [0,1]\,. \tag{3.4}$$

We assume that agents, in addition to observing the current and past realization of equilibrium prices, are endowed with the exogenous information

$$U_t^i = \mathbb{V}_t\left(\varepsilon_i\right) \quad \text{for} \quad i \in [0,1]\,. \tag{3.5}$$

The information set of an individual agent i can thus be written as

$$\Omega_t^i = \mathbb{V}_t(\varepsilon_i) \vee \mathbb{V}_t(p) \vee \mathbb{M}_t(p) \tag{3.6}$$

The following Theorem characterizes the equilibrium under dispersed information.

Theorem 1. *Let $\tau \equiv \sigma_\varepsilon^2/(\sigma_v^2 + \sigma_\varepsilon^2)$ be the signal-to-noise ratio associated with the signal ε_{it} in (3.4). Under the information assumption (3.6), a unique Information Equilibrium for (2.1) with $|\beta| < 1$ always exists and is determined as follows. Suppose that exactly one real scalar $|\lambda| < 1$ can be found such that*

$$A(\lambda) - \frac{\tau\beta A(\beta)(1-\lambda\beta)}{\lambda - (1-\tau(1-\lambda^2))\beta} = 0 \tag{3.7}$$

then the information equilibrium price is given by

$$p_t = \frac{1}{L-\beta}\left[LA(L) - \beta A(\beta)\frac{h(L)}{h(\beta)}\right]\varepsilon_t \tag{3.8}$$

[9] The informational setup of this section is especially common in the recent and fast growing literature on dispersed information and the business cycle; see, for example, Angeletos and La'O (2009b), Hellwig and Venkateswaran (2009), Lorenzoni (2009), Maćkowiak and Wiederholt (2007) and Rondina and Walker (2011).

with

$$h(L) \equiv \tau\lambda + (1-\tau)\left(\frac{\lambda - L}{1-\lambda L}\right). \tag{3.9}$$

If such a scalar $|\lambda| < 1$ *cannot be found, then the information equilibrium price is given by* (3.3).

Proof. See Appendix A. □

The theorem contains two fixed point conditions, which coincides with our definition of an Information Equilibrium (Definition IE): one that characterizes the information set, (3.7), and one that characterizes the equilibrium price, (3.8). Providing economic intuition and developing a deep understanding of these two equations is the sole purpose of the rest of the paper. Foreshadowing results, we show how (3.7) relates to the agents' information sets and in particular to knowledge of the model. We also derive a one-to-one mapping between this economy and one in which information is hierarchical. This mapping delivers an aggregation result that further enhances economic intuition. Notice also that the equilibrium price takes the form of a generalized Hansen-Sargent formula, with the term $\frac{h(L)}{h(\beta)}$ representing the departure from the standard formula of equation (3.3). In Section 4 we show that the term $h(L)$ emerges from the consideration that agents use the knowledge of the model, together with the history of the equilibrium price, to infer what is the market forecast of future prices.

Before proceeding to this analysis, we first want to argue that these information equilibria are empirically interesting objects by asking the question: How different is the equilibrium price of equation (3.8) from the full information price? Figure 2 reports the impulse response to a one time innovation in the fundamental process ε_t of the equilibrium price characterized in Theorem 1 when primitives of the model are such that a $|\lambda| < 1$ satisfying (3.7) can be found compared to the full information benchmark. The process for the economic fundamental is specified as $A(L) = \frac{1+\theta L}{1-\rho L}$, with $\theta = \sqrt{11}$ and $\rho = 0.9$, so that the effect of an innovation peaks after one period. The rest of the parameter values are set to $\beta = 0.9$ and $\tau = 0.02$. The full information equilibrium is given by plugging these numbers into equation (3.3).

As shown in figure 2, the full information price strongly reacts at impact by taking into account that the shock is persistent and will therefore affect the equilibrium price over the next several periods through its effect on the predictability of future price realizations. The effect will peak after one period and then decay monotonically. The behavior of the full information price essentially amplifies the behavior of the economic fundamentals through the forward looking nature of the equilibrium price equation.

The blue line represents the information equilibrium price of equation (3.8), with equation (3.7) providing the endogenous value for λ of -0.47. The information equilibrium price displays the confounding dynamics of section 2.2; in addition agents have noisy signals about the innovation and so they are not able to infer the value of the economic fundamentals. As a consequence they will under-react, roughly by 50% compared to the full information

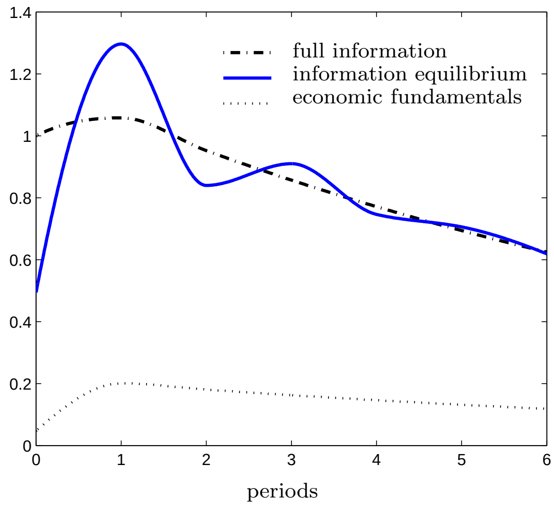

Figure 2: Impulse response of the full information price, the information equilibrium price of Theorem 1, and the process for economic fundamentals $s_t = 0.9s_{t-1} + \varepsilon_t + \sqrt{11}\varepsilon_{t-1}$. The parameter values are $\beta = 0.9$ and $\tau = 0.02$.

case. In the following period, as new information becomes available through the equilibrium price, agents realize that an innovation occurred, but because they initially under-reacted, they now infer from the equilibrium price that the innovation is larger than the actual value. This results in a very optimistic view of the fundamentals and an over-reaction of the equilibrium prices, of about 25% of the full information counterpart. In the subsequent period, observing the equilibrium price agents will think that they have over-estimated the innovation and they will correct downward their expectations, now again erring on the downside, causing the equilibrium price to under-react by 10% with respect to full information. As the over- and under-reactions subside, the equilibrium price response gets closer to the full information case. It is important to emphasize that this over- and under-reaction is optimal. Agents are fully rational and yet from the perspective of the true economic fundamentals, the market price presents what looks like waves of "optimism" and "pessimism" with respect to the full information benchmark. Given that many empirical time series (e.g., asset prices, business cycles) follow boom-bust cycles, we view the information equilibrium as a very interesting departure from the standard equilibrium.

4 Information Equilibria: Characterization

4.1 Equivalent Representation In order to develop intuition for Theorem 1, we begin by stating a powerful "aggregation" result and deriving an equivalence to an alternative information structure where instead of considering a continuum of agents with dispersed information, we assume that there are two types of buyers: fully informed and uninformed. The fully informed buyers observe the entire history of economic fundamentals ε up to time t; the uninformed buyers observe only the entire history of prices up to time t. The proportion of the fully informed buyers is denoted by $\mu \in [0,1]$, and, consequently the proportion of the uninformed buyers is

$1-\mu$. More formally, we consider a market with the following exogenous information structure:

$$U_t^i = \mathbb{V}_t(\varepsilon) \text{ for } i \in \mu \text{ and } U_t^i = \{0\} \text{ for } i \in 1-\mu. \tag{4.1}$$

Note that under this informational assumption, the market equilibrium equation (2.1) can be written as

$$p_t = \beta\left[\mu \mathbb{E}_t^{\mathcal{I}}\left(p_{t+1}\right) + (1-\mu)\,\mathbb{E}_t^{\mathcal{U}}\left(p_{t+1}\right)\right] + s_t. \tag{4.2}$$

where $\mathcal{I}$ is notation for the *fully informed*, while $\mathcal{U}$ is notation for the *uninformed*. The following theorem states the equivalence result.

Theorem 2. *Under the exogenous information assumption* (4.1)*, a unique Information Equilibrium for* (4.2) *with* $|\beta| < 1$ *always exists and is equivalent to the equilibrium characterized in Theorem 1 with* $\tau \equiv \mu$.

Proof. See Appendix A. □

The theorem states that in terms of the aggregate characterization of the equilibrium, the dispersed information setup is identical (i.e., same existence condition (3.7) and same equilibrium pricing function (3.8)) to the hierarchical information setup when the signal-to-noise ratio $\tau \equiv \sigma_\varepsilon^2/(\sigma_v^2 + \sigma_\varepsilon^2)$ is equal to the proportion of informed traders, μ. This equivalence result can be understood by thinking of the strategic behavior of the dispersedly informed buyers of Theorem 1. Each agent i receives a privately observed signal ε_{it} and a publicly observed signal p_t about the unobserved fundamental ε_t. The optimal behavior—in terms of forecast error minimization—is for the agent to act *as if* the signal ε_{it} contained no noise and thus was equal to the true state ε_t, in measure proportional to the informativeness of the signal τ. At the same time, it is certainly possible that the signal is pure noise and thus it would be optimal to ignore it and act just upon the public signal p_t, this in measure $(1-\tau) = \sigma_v^2/(\sigma_v^2 + \sigma_\varepsilon^2)$. Thus, in a dispersed information setting each agent behaves optimally by employing a "mixed" strategy approach: act as if they possess the full information of the informed buyers $\mathcal{I}$ of Theorem 2 with probability τ, and act as if they possess just the public information of the uninformed buyers $\mathcal{U}$ of Theorem 2 with probability $1-\tau$. Theorem 2 shows that the equilibrium price of a market with μ fully informed and $1-\mu$ fully uninformed buyers displays, in the aggregate, the "mixed" strategies of the individual buyers in the dispersed information environment. Theorem 2 is an aggregation result that allows one to study a market with only two representative buyers, one with full information in measure τ and one with just public information in measure $1-\tau$, knowing that the aggregate behavior of that market is equivalent to the aggregate behavior of a market with dispersedly informed (i.e. heterogeneous) buyers. We will make extensive use of this result when studying the aggregate properties of the equilibrium.

4.2 Informational Fixed Point

The informational fixed point condition (3.7) lies at the heart of the existence of an equilibrium with confounding dynamics, which is the key contribution of this paper. In this section we explore this condition and show under which economically relevant environments it can emerge. We argue that, even in a simple one-equation model like the one used in our analysis, condition (3.7) is easily obtained.

The following is a useful corollary to both Theorems 1 and 2 that helps in understanding the source of the confounding dynamics.

Corollary 1. *Let $\tau \to 0$ (or, equivalently, $\mu \to 0$), a unique Information Equilibrium for (2.1) with $|\beta| < 1$ always exists and is determined as follows: Suppose that exactly one real scalar $|\lambda| < 1$ can be found such that*

$$A(\lambda) = 0, \tag{4.3}$$

then the information equilibrium price process is

$$p_t = \frac{1}{L-\beta}\left[LA(L) - \beta A(\beta)\frac{\mathcal{B}_\lambda(L)}{\mathcal{B}_\lambda(\beta)}\right]\varepsilon_t \tag{4.4}$$

where $\mathcal{B}_\lambda(L) = \frac{L-\lambda}{1-\lambda L}$. If condition (4.3) does not hold for $|\lambda| < 1$, then the Information Equilibrium is given by (3.3).

Proof. See Appendix A. □

Condition (4.3) offers an important insight into the existence condition (3.7) in Theorem 1. It stipulates that, in order for the equilibrium price to display confounding dynamics as the informativeness of the signal goes to zero (or equivalently as the the proportioned of informed traders goes to zero), the supply process s_t must also possess confounding dynamics with respect to the structural innovations, ε_t. To see this more clearly, note that the supply process can be written as $s_t = (L-\lambda)\hat{A}(L)\varepsilon_t$—where $\hat{A}(L)$ has no zeros inside the unit circle—to satisfy (4.3). This supply process will contain the confounding dynamics described in section 2.2.

The key intuition behind this restriction and the existence condition (3.7) comes from the agents' knowledge of the model, $\mathbb{M}_t(p)$. This concept gets at the idea that in a rational expectations framework agents know that the price that clears the market must satisfy (2.1). Investors will use this information to reduce their forecast errors. Specifically as $\tau \to 0$, all agents will rationally believe that all market participants will have the same expectations about next period's price in equilibrium. Therefore whatever this expectation is, they know that it must satisfy

$$p_t - \beta\mathbb{E}_t(p_{t+1}) = s_t. \tag{4.5}$$

Recall that as $\tau \to 0$, only the public signal, p_t, is available to traders at t. Knowledge of the model implies that the left-hand side of (4.5) is in the information set of the traders, and therefore the entire history of s_t

must also be contained in the information set of all the agents in equilibrium, i.e. $\mathbb{M}_t(p) = \mathbb{V}_t(s)$. This suggests that in order for confounding dynamics to exist in equilibrium, the s_t itself must display such dynamics, which is exactly what condition (3.7) states.[10] In a simple representative agent economy, imposing confounding dynamics on the exogenous process is sufficient to generate *endogenous* confounding dynamics in equilibrium. The model's cross-equation restrictions ensure that endogenous variables will inherit the stochastic properties of the exogenous variables—confounding dynamics in this case.

In dynamic models with asymmetrically informed agents, conditions that guarantee agents remain heterogeneously informed in equilibrium (i.e., conditions which preserve confounding dynamics) are more difficult to derive and not easily interpretable, as evidenced by (3.7). However, the intuition behind knowledge of the model concept provides an unified way to proceed.[11]

Consider the case of Theorem 2 with $\mu > 0$. Condition (3.7) gives the condition that must hold for the uninformed agents to remain uninformed in equilibrium. Through knowledge of the model, the uninformed buyers will recognize that in equilibrium the following relationship must hold

$$p_t - \beta(1-\mu)\mathbb{E}_t^{\mathcal{U}}(p_{t+1}) = \beta\mu\mathbb{E}_t^{\mathcal{I}}(p_{t+1}) + s_t. \tag{4.6}$$

The difference between this existence condition and that of Corollary 1 is that the uninformed buyers are not able to back out the exact process for s_t given the history of prices and uninformed predictions, $\mathbb{E}^{\mathcal{U}}$. However, they are able to uncover the sum of the supply process s_t and the predictions of the fully informed buyers $\mathbb{E}^{\mathcal{I}}$. The question is whether this sum displays confounding dynamics that can be inherited by the equilibrium price. Condition (3.7) provides the answer to this question. Appendix A shows that (3.7) is equivalent to the right-hand side of (4.6) evaluated at λ. If this term vanishes at $|\lambda| < 1$, then the sum of the informed agents' expectation and the supply process has a non-fundamental moving average representation and is not invertible with respect to the information set of the uninformed agents. In other words, condition (3.7) implies the right-hand side of (4.6) will display confounding dynamics. Consequently the uninformed agents will only be able to see the sum but not the individual components of the sum. It is in this sense that models with disparately informed agents lead to *endogenous* signal extraction. Uninformed agents want to disentangle the effects on the equilibrium price of the informed agent's expectations from the supply process.

The above intuition is useful in interpreting the existence condition for the dispersed information case. When $\tau > 0$, knowledge of the model under dispersed information results in agents being able to infer the sum of the supply process s_t and the difference between the average market expectations $\bar{\mathbb{E}}$ and their individual expectations,

[10] The reasoning behind the result presupposes that all the agents at time t have access to the entire history of their expectations. If this was not the case, which for example could happen if one were to consider an overlapping generation structure of the market where a generation of agents is born in each period and dies the next period, then the new generation would only be able to observe the current realization of s_t and so the information equilibrium might not coincide with the one characterized by (4.4).

[11] Moreover, Rondina and Walker (2011) show that this concept overturns non-existence results thought to be pervasive in models with heterogeneously informed agents.

namely

$$p_t - \beta \mathbb{E}(p_{t+1}|\varepsilon_i^t, p^t) = \beta[\bar{\mathbb{E}}(p_{t+1}|\varepsilon_i^t, p^t) - \mathbb{E}(p_{t+1}|\varepsilon_i^t, p^t)] + s_t \tag{4.7}$$

The informational fixed point in Theorem 1 ensures that the process on the right hand side displays confounding dynamics for the information set (ε_{it}, p_t), so that, in equilibrium, knowledge of the model does not perfectly reveal the fundamental innovation ε_t.

Since condition (3.7) lies at the core of Theorem 1 it is important to ask whether it holds in economically relevant situations. Indeed, confounding dynamics can emerge in many interesting settings. For example, diffusion processes, such as the adoption of a new technology, normally display confounding dynamics. The diffusion pattern takes the typical "S" shape: an initial phase of low diffusion, a steep middle diffusion phase and final leveling-off phase [see Rogers (2003)]. Following Canova (2003), a diffusion process where an initial shock ε_t diffuses with the canonical "S" shape can be formalized by

$$s_t = s_{t-1} + \alpha\varepsilon_t + 2\alpha\varepsilon_{t-1} + .75\alpha\varepsilon_{t-2}, \tag{4.8}$$

with $0 < \alpha < 1$. The diffusion process (4.8) displays confounding dynamics.[12] For example, with $\beta = .5$, letting $\tau = 0$ Corollary 1 would then ensure that an information equilibrium is given by (4.4) with $\lambda = -2/3$, whereas, letting $\tau = .01$, Theorem 1 would ensure that an information equilibrium is given by (3.8) with $\lambda = -.7$.

One additional concern about (3.7) is that it could hold only for a combination of parameter values with measure zero, i.e. it could be a non-generic condition. This is clearly not the case. For simplicity consider the limiting condition (4.3). The equilibrium of the corollary is generic because $|\lambda|$ can be *anywhere* inside the unit circle, and $A(\lambda) = 0$ is the *only* restriction placed on $A(\cdot)$. The same argument can be immediately extended to (3.7) by continuity. This suggests that interesting information equilibria can easily emerge from standard rational expectations models. For example, from the diffusion process in (4.8) one can safely change the parameters along several dimensions without affecting the existence of a $|\lambda|$ satisfying (3.7) in Theorem 1. We provide additional examples of the non-generic behavior of the information equilibrium below.

4.3 Aggregate Characterization

Equipped with the results of Theorems 1 and 2 we turn now to the study of the properties of the price in an information equilibrium. We first notice that the price function in all the Theorems takes the form of a modified Hansen-Sargent formula (3.3). The Hansen-Sargent formula essentially represents an operator that "conditions down" from the full history of innovations (past, present and future) to

[12] Notice that we have specified a process with a unit root in (4.8), while we have previously stated that we focus on stationary equilibria. The unit root in the exogenous process can be easily dealt with by specifying an AR coefficient, solve for the equilibrium and then take the limit for the coefficient going to 1. The level of the price process will not have a well defined second moment, but the dynamics can be expressed in first differences. Alternatively, one could take the first difference of the market price using equation (2.1), which would eliminate the unit root due to s_t but not the confounding dynamics, and solve directly for the first difference.

a linear combination of innovations by subtracting off what is not contained in the information set of the agents. Corollary 2 formalizes the idea.

Corollary 2. *Under the assumptions of Theorem 2, if* $|\lambda| < 1$ *satisfying* (3.7) *exists, the information equilibrium price can be written as*

$$p_t = \left(\frac{LA(L)}{L-\beta}\right)\varepsilon_t - \left(\frac{\beta A(\beta)}{L-\beta}\right)\varepsilon_t - (1-\tau)\beta A^{\mathcal{U}}(\beta)\left(\frac{1-\lambda^2}{1-\lambda L}\right)\varepsilon_t, \tag{4.9}$$

where

$$A^{\mathcal{U}}(L) = \frac{A(L)}{L-\lambda-\tau\beta(1-\lambda^2)}. \tag{4.10}$$

Proof. Follows directly from Theorem 1. □

The Corollary represents the information equilibrium price as being comprised of three components. The first component of the RHS of (4.9) is the perfect foresight equilibrium,

$$p_t^f = \sum_{j=0}^{\infty} \beta^j s_{t+j} = \frac{LA(L)}{L-\beta}\varepsilon_t \tag{4.11}$$

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

The second component operates a first conditioning down that takes into account the fact that future values of ε_t are not known at t. This conditioning down amounts to subtracting off a particular linear combination of future values of ε_t, specifically

$$\beta A(\beta)\sum_{j=1}^{\infty}\beta^j \varepsilon_{t+j} \tag{4.12}$$

The third component is the novel part of the representation. It represents the conditioning down related to the uninformed buyers not being able to perfectly unravel the past realizations of ε_t from the equilibrium price—the confounding dynamics. The interpretation of this term offers important insights into the working of an information equilibrium at the aggregate level. To shed light on these insights we make use of the aggregation result of Theorem 2 and so we consider informed and uninformed buyers. Let $\mathbb{E}_t^{\mathcal{I}}(s_{t+1}) = \mathbb{E}[s_{t+1}|\mathbb{V}_t(\varepsilon)]$ denote prediction formula of a fully informed buyer, and $\mathbb{E}_t^{\mathcal{U}}(s_{t+1}) = \mathbb{E}[s_{t+1}|\mathbb{V}_t(p) \vee \mathbb{M}_t(p)]$ the prediction formula of an uninformed buyer in the information equilibrium of Corollary 2. Let us assume for the moment that $\mu \equiv \tau = 0$. In the equilibrium with only uninformed buyers, agents are concerned with forecasting the discounted, infinite sum of market fundamentals, i.e., $p_t = \sum_{j=0}^{\infty}\beta^j\mathbb{E}_t^{\mathcal{U}}(s_{t+j})$. Writing out the uninformed buyers expectations of future supply using the analytic form of the equilibrium price yields

$$\mathbb{E}_t^{\mathcal{U}}(s_{t+j}) = \mathbb{E}_t^{\mathcal{I}}(s_{t+j}) - A_{j-1}^{\mathcal{U}}\left(\frac{1-\lambda^2}{1-\lambda L}\right)\varepsilon_t. \tag{4.13}$$

The uninformed agents' expectations of fundamentals at each future date are given by the expectation of fully informed agents minus a term given by the linear combination of past ε_t's that the agents do not observe. This linear combination consists of the noise stemming from the confounding dynamics generated by $|\lambda| < 1$ (see Section 2.2, Equation (2.7)) multiplied by a coefficient that corresponds to the weight on the $(j-1)^{th}$ lag of the polynomial $A^{\mathcal{U}}(L)$ which represents the dynamics of the supply process s_t as *perceived* by the uninformed buyers in equilibrium. Uninformed agents would formulate predictions that are equal to those formulated by fully informed agents if it were not for the confounding dynamics. The information equilibrium price then contains the accumulated noise for the expectations at all horizons, namely

$$\sum_{j=1}^{\infty} \beta^j A_{j-1}^{\mathcal{U}}\left(\frac{1-\lambda^2}{1-\lambda L}\right)\varepsilon_t = \beta A^{\mathcal{U}}(\beta)\left(\frac{1-\lambda^2}{1-\lambda L}\right)\varepsilon_t. \tag{4.14}$$

Notice that as $|\lambda|$ gets closer to 1, the noise due to the confounding dynamics becomes smaller, disappearing in the limit.

When fully informed buyers are introduced into the market, so that $\mu > 0$, the noise due to confounding dynamics is affected through two channels. First, there are fewer uninformed buyers and so only a fraction $1-\mu$ of the cumulated noise (4.14) has to be subtracted off. Second, the presence of informed buyers changes the perceived supply process $A^{\mathcal{U}}(L)$ for the uninformed buyers as the equilibrium price now contains more information: both the polynomial $A^{\mathcal{U}}(L)$ and λ will reflect this change. As the proportion of informed buyers increases ($\mu \to 1$), the information equilibrium approaches the full information counterpart and the third term in (4.9) vanishes.

4.4 Dispersed Characterization While Theorem 2 guarantees equivalence with the informed-uninformed buyers setup at the aggregate level, there exist important differences between the two equilibria at the individual agent level. First, the dispersed information equilibrium displays a well defined cross sectional distribution of beliefs, as opposed to the degenerate distribution in the hierarchical case. Second, the cross-sectional variation is perpetual in the sense that the unconditional cross-sectional variance is positive. In other words, agents' beliefs are in perpetual disagreement. These two results are stated in terms of expectations about future prices in the following proposition.

Proposition 1. *Let $p_t = (L-\lambda)Q(L)\varepsilon_t$ be the information equilibrium characterized by Theorem 1, with $|\lambda| < 1$.*

The cross section of beliefs about future prices is given by

$$\mathbb{E}_t^i(p_{t+j}) = \mathbb{E}_t^{\mathcal{I}}(p_{t+j}) - (1-\tau)Q_{j-1}\left(\frac{1-\lambda^2}{1-\lambda L}\right)\varepsilon_t - \tau Q_{j-1}\left(\frac{1-\lambda^2}{1-\lambda L}\right)v_{it} \qquad \textit{for } j = 1, 2, ... \tag{4.15}$$

The implied unconditional cross-sectional variance in beliefs is given by

$$\tau^2\left(1-\lambda^2\right)(Q_{j-1})^2\sigma_v^2 \qquad \textit{for } j = 1, 2, ... \tag{4.16}$$

Proof. See Appendix A. □

If one considers the interpretation of the optimal signal extraction problem under dispersed information in terms of mixed strategies, the beliefs in (4.15) have an intuitive interpretation. If information was complete, the beliefs would coincide with the expectation $\mathbb{E}_t^{\mathcal{I}}(p_{t+j})$. The difference of the beliefs of agent i with respect to the full information has two components. One is common across agents, one is specific to each agent. The common component (the second term on the RHS of (4.15)) is the result of agent i acting *as if* uninformed with probability $1-\tau$. Similar to the uninformed buyers in the informed-uninformed case, agent i formulates her beliefs based on the common public information embedded into prices. As a result, her beliefs will differ from the full information case according to the noise due to confounding dynamics. The idiosyncratic component (the third term on the RHS of (4.15)) is the result of the agent acting *as if* they are fully informed. In acting as fully informed, the agent will condition on their private signal ε_{it}. In so doing she will inject an idiosyncratic error into her beliefs. As for the unconditional variance of the beliefs, Proposition 1 offers an analytical form that can be very useful in calibrating key parameters of the market if data on cross-sectional beliefs on prices are available.

4.5 Information Equilibrium: An Example

We conclude this section with a specific example which allows us to further analyze existence conditions and provide a sharper characterization of the resulting information equilibrium. Let the supply process s_t be given by

$$s_t = \rho s_{t-1} + \varepsilon_t + \theta\varepsilon_{t-1}, \qquad |\rho| \leq 1. \tag{4.17}$$

The full information solution to the equilibrium price is obtained by substituting (4.17) in (3.3), which results in

$$p_t - \rho p_{t-1} = \left(\frac{1+\theta\beta}{1-\rho\beta}\right)\varepsilon_t + \theta\varepsilon_{t-1}. \tag{4.18}$$

Suppose that the exogenous information for the buyers is specified as in (3.4) and the parameter values are such that exactly one $|\lambda| < 1$ that satisfies (3.7) exists, then Theorem 1 provides the closed form solution for the equilibrium price as

$$p_t^* - \rho p_{t-1}^* = \left(\frac{\lambda - L}{1-\lambda L}\right)\frac{1}{\lambda}\left[\frac{1+\theta\beta}{1-\rho\beta}\left(1+\frac{(1-\tau)\beta(1-\lambda^2)}{\lambda-(1-\tau(1-\lambda^2))\beta}\right)\varepsilon_t + \lambda^2\theta\varepsilon_{t-1}\right] \tag{4.19}$$

with λ being the solution to [13]

$$\frac{1+\theta\lambda}{1-\rho\lambda} = \tau\beta\left(\frac{1+\beta\theta}{1-\rho\beta}\right)\frac{1-\lambda\beta}{\lambda-(1-\tau(1-\lambda^2))\beta} \tag{4.20}$$

How do the two equilibria differ? Both equilibria share the autoregressive root ρ; however, the information equilibrium p_t^* contains an additional autoregressive root at λ. This is due to the presence of confounding dynamics in equilibrium: the learning effort of the uninformed buyers results in an additional persistent effect of past innovations. In addition, the process p_t^* also has an MA(2) representation, compared to the MA(1) of p_t.

To gain some insights on the different structure of the two equilibria at the aggregate level it is useful to look at the case when $\tau \to 0$. According to Corollary 1 the type of IE encountered hinges upon whether s_t spans the space of ε_t. The restriction $A(\lambda) = 0$ yields $(1+\theta\lambda)/(1-\rho\lambda) = 0$, which gives $\lambda = -1/\theta$. Therefore, if $|\theta| < 1$, then the s_t process spans ε_t. In this case equation (4.19) becomes

$$\tilde{p}_t - \rho\tilde{p}_{t-1} = \left(\frac{1+\theta L}{L+\theta}\right)\left[\left(\frac{\theta+\beta}{1-\rho\beta}\right)\varepsilon_t + \varepsilon_{t-1}\right]. \tag{4.21}$$

Figure 3 plots the impulse response functions for p_t and $\tilde{p}_t$ for two levels of confounding dynamics: $\lambda = -1/\theta = -1/\sqrt{11}$ in the left panel, and $\lambda = -1/\theta = -1/\sqrt{2}$ in the right panel.[14] 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 set to: $\beta = 0.985$, $\sigma_\varepsilon = 1$. We set $\rho = 1$ so that the process (4.17) can be interpreted as a diffusion process where innovations spread gradually but have a permanent effect. In response to an innovation, s_t will change permanently but such a change happens gradually over the course of two periods: at impact there is a jump to 1, after one period there is an additional jump of $1+\theta$ and then the process levels off at the new higher value. The source of confounding dynamics lies in the second jump being bigger than the first. This is common in diffusion processes where after an initial weak diffusion phase the diffusion gradient increases and becomes maximal before decreasing and leveling off once the diffusion is completed.

The full information price p_t reacts immediately to the innovation taking into account the accumulated permanent effect of the shock on the future values of the fundamentals s_t. The scale of the reaction at impact is dictated by the discount factor β. After the initial jump the dynamics follow that of the fundamentals and so the

[13] Condition (3.7) by construction has always a solution at $\lambda = \beta$; this particular solution can be disregarded as it is independent of the informational assumptions.

[14] These numbers are chosen so that the equivalent signal-to-noise ratios in a standard signal extraction problem correspond to 10 and 1, respectively.

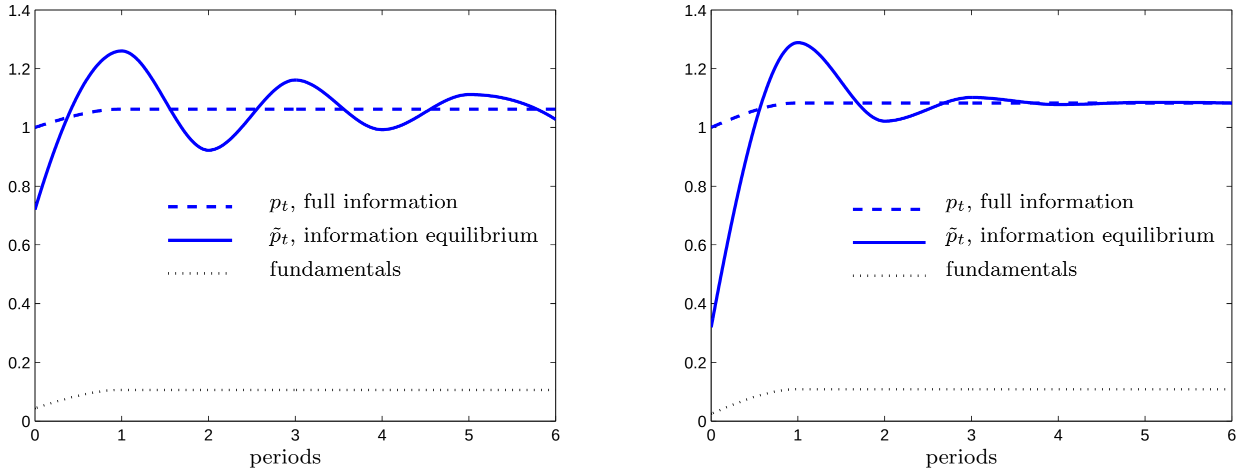

Figure 3a: strong confounding dynamics

Figure 3b: weak confounding dynamics

Figure 3: Impulse response of market price to one time innovation in ε_t. The dotted line represent the response of s_t; the dashed line is the response of the full information price p_t in Equation (4.18); the solid line is the response of the information equilibrium price $\tilde{p}_t$ in Equation (4.21). The responses are normalized so that the full information price has a unitary reaction at period 0; other parameters values are $\rho = 1$ and $\beta = .9$.

price levels off to the new permanent level. The market price with confounding dynamics $\tilde{p}_t$ displays substantially different behavior. First, because the agents cannot really be sure that a positive innovation has been realized, the price under-reacts at impact. The under-reaction is more pronounced for the strong confounding case (35% of the full information reaction) than for the weak one (75% of the full information reaction). At period 1, while the full information price reaches the new permanent plateau, the price with confounding dynamics overshoots the plateau by roughly 25% in both the strong and weak confounding case. After that, in the strong confounding case the price keeps fluctuating, but only slightly so, while the fluctuations are more persistent for the weak confounding case. The intuition for this is that the price is understood to be a bad signal in the strong confounding case, and so it gets discounted much quicker, which results in the innovation being given less relevance in the subsequent learning effort. In the weak confounding case, the price is a good signal of the innovation and so it remains important in the signal extraction problem, but in so doing the price remains affected by the learning effort for several periods in the future.

It bears reminding that there is no exogenously superimposed noise in the market generating the equilibrium price $\tilde{p}_t$. The dynamics of s_t are canonical diffusion dynamics, the market price is perfectly observed and agents are fully rational. And yet the market dynamics display waves of optimism and pessimism. This example is suggestive of the potential of the equilibria belonging to the class that we characterized in Theorems 1-2 for offering a rational explanation of apparently irrational market behavior, for example, market turbulence in periods of technological innovation.

We turn next to the analysis of the existence of the information equilibrium p_t^* when $\tau > 0$ in the context

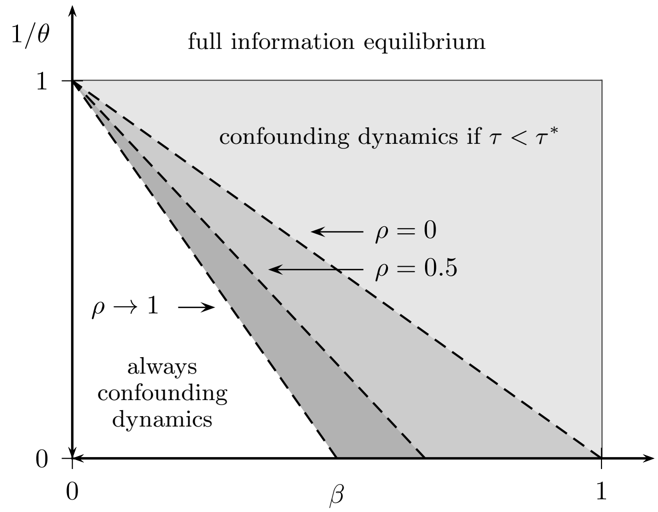

Figure 4: Existence space of Information Equilibrium with confounding dynamics as τ, ρ and θ are varied for the supply process $s_t = \rho s_{t-1} + \varepsilon_t + \theta\varepsilon_{t-1}$.

of the current example. This boils down to the existence of a $|\lambda| < 1$ that satisfies (4.20). The following result summarizes how the existence condition behaves as β, ρ, θ and τ are varied. The proof is reported in Appendix A.

Result *The model described by* (2.1) *and* (4.17) *with* $\beta, \rho \in (0, 1)$ *and* $\theta > 0$ *defines a space of existence for information equilibria with confounding dynamics of the form* (3.8) *characterized as follows:*

(R.1) *If* $\theta \leq 1$ *an IE with confounding dynamics does not exist.*

(R.2) *If* $\theta > 1$, *an IE with confounding dynamics exists if and only if* $\tau < \tau^*$ *with*

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

Figure 4 displays the existence conditions for an information equilibrium with confounding dynamics in (β, θ) space. Four points are noteworthy. First, as is evident from the figure and condition (**R.2**), if $\theta \leq 1$ an IE with confounding dynamics does not exist regardless of the other parameters in the model. Intuitively, if we interpret once again s_t as a diffusion process, when $\theta \leq 1$ there is no initial slow diffusion phase; the strongest diffusion takes place immediately and subsequently levels off.

Second, from condition (**R.2**), for a certain region of the parameter space (to the right of the dashed lines in figure 4) an IE with confounding dynamics exists only if the proportion of fully informed buyers is sufficiently small. The dashed lines represent the IE that prevails as $\tau \to 1$, plotted for various values of the autoregressive parameter ρ. To the left of the dashed line, confounding dynamics will *always* be preserved in equilibrium regardless the value of τ; from condition (**R.2**) this happens when $\theta \geq (1+\beta)/(1+\beta^2-\beta\rho(1-\beta))$. From section 2.2 we know that an increase in θ (a decrease in λ) corresponds to an increase in the noise associated with the confounding dynamics. The informational disparity between the fully informed and uninformed may become so large that no matter how many fully informed buyers participate in the market, the confounding dynamics will never be unraveled. How the discount factor β alters the space of existence is similar to that of the serial correlation parameter ρ, which is the third point to be made. As the serial correlation in the s_t process increases and β increases, it is more difficult to preserve confounding dynamics (the dashed line shifts to the left as ρ increases from 0 to 1). 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 fully informed and uninformed. Finally, the figure demonstrates the generic nature of the information equilibrium. The space of existence that preserves confounding dynamics is dense. Relatively small values of β and large values of θ *always* yield the IE given by Theorem 1 independent of τ and ρ.

5 Higher-Order Beliefs

In Section 3 and 4 we have characterized a class of rational expectations equilibria where agents remain differentially informed in equilibrium. It is well known that one way to describe the behavior of rational agents in such settings is in terms of engaging in higher-order thinking. Yet we have not discussed the form of such strategic thinking even though the equilibrium characterizations embed these dynamics. In fact, the rational expectations assumption implies that the solution to the equilibrium model 2.1 must be identical to the solutions of

$$p_t = \beta\bar{\mathbb{E}}_t\{\beta\bar{\mathbb{E}}_{t+1}p_{t+2} + s_{t+1}\} + s_t \tag{5.1}$$

$$= \beta^2\tau^2\mathbb{E}_t^{\mathcal{I}}p_{t+2} + \beta^2(1-\tau)^2\mathbb{E}_t^{\mathcal{U}}p_{t+2} + \beta\tau\mathbb{E}_{t+1}^{\mathcal{I}}s_{t+1} + \beta(1-\tau)\mathbb{E}_t^{\mathcal{U}}s_{t+1}$$

$$+ \beta^2\tau(1-\tau)\mathbb{E}_t^{\mathcal{I}}\mathbb{E}_{t+1}^{\mathcal{U}}p_{t+2} + \beta^2\tau(1-\tau)\mathbb{E}_t^{\mathcal{U}}\mathbb{E}_{t+1}^{\mathcal{I}}p_{t+2} + s_t \tag{5.2}$$

where we have used the aggregation result of Theorem 2, recursive substitution and the shorthand notation for the average expectations operator, $\bar{\mathbb{E}}_t = \tau\mathbb{E}_t^{\mathcal{I}}(\cdot) + (1-\tau)\mathbb{E}_t^{\mathcal{U}}(\cdot)$. These model specifications highlight the strategic interactions undertaken by agents. We interpret the equation from an aggregate point of view, and so we consider the two representative agents, informed and uninformed. The first two elements on the RHS of (5.2) follow from the law of iterated expectations, which must hold with respect to the individual agents' information sets. The first and second components of the second line in (5.2) encode the model's higher-order beliefs. Informed agents engage in forming the expectations of the uninformed agents' $t+1$ expectations of the price at $t+2$, and, similarly, uninformed agents engage in forming the expectations of the informed agents' $t+1$ expectations of the price at $t+2$. Substituting recursively for the future prices in (5.2) one obtains a representation of the equilibrium solution in terms of a weighted sum of higher-order beliefs (about the first moments) of the future supply realizations s_{t+j}. The strategic interaction amongst agents and the higher-order belief (HOBs) dynamics are usually considered mysterious objects since in many situations, especially in dynamic settings, it is hard to write the analytic form of the HOBs of any arbitrary order. Having a closed-form solution in hand, we are in able to study HOBs analytically. We derive the expectations found in (5.2) and show how they lead to the breakdown in the law of iterated expectations for the average expectations operator, as emphasized in (5.1). We also analyze the role of HOBs in information diffusion. In our analysis we focus on the higher-order beliefs in terms of future prices. An analogue analysis can be undertaken by focusing on the higher order beliefs about the underlying fundamental process s_t; we find that the resulting insights are equivalent under both approaches. In most of the following analysis, sections 5.1 and 5.2, we focus on the aggregate representation of higher-order-beliefs and so we focus on the informed-uninformed representation of the equilibrium. Theorem 2 ensures that this is without loss of generality, at least with respect to the aggregate behavior. Whenever we refer to the behavior of the informed (resp. uninformed) buyer, it is useful to think of it as the behavior of the dispersedly informed agent acting *as if* an informed buyer (resp. *as if* an uninformed buyer). In section 5.3 we report the analytical characterization of the higher order beliefs at the individual buyer level of the dispersed information case. We will show that the interpretation of the higher order thinking behavior of the dispersedly informed buyer is a straightforward application of the results of sections 5.1 and 5.2.

5.1 Higher-Order Beliefs Characterization

The first step in characterizing higher-order beliefs in an information equilibrium is to isolate the speculative component (agent $\mathcal{I}$'s belief about agent $\mathcal{U}'s$ belief and vice versa) by defining a stochastic process that takes as given the other agent's expectation. To achieve this we define $f_t^{\mathcal{I}} \equiv s_t - (1-\mu)\beta(p_{t+1} - \mathbb{E}_t^{\mathcal{U}}(p_{t+1}))$ for the informed agents and $f_t^{\mathcal{U}} \equiv s_t - \mu\beta(p_{t+1} - \mathbb{E}_t^{\mathcal{I}}(p_{t+1}))$ for the uninformed. Under these definitions we can state the following result.

Corollary 3. *If $|\lambda| < 1$, the IE described in Theorem 2 has the following representation,*

$$p_t = \frac{1}{L-\beta}\Big((1-\mu)\{LH^{\mathcal{U}}(L) - \beta H^{\mathcal{U}}(\beta)\kappa(L)\}\varepsilon_t + \mu\{LH^{\mathcal{I}}(L) - \beta H^{\mathcal{I}}(\beta)\}\Big)\varepsilon_t, \tag{5.3}$$

where $H^{\mathcal{U}}(L)\varepsilon_t = f_t^{\mathcal{U}}$, $H^{\mathcal{I}}(L)\varepsilon_t = f_t^{\mathcal{I}}$ and $\kappa(L) = \mathcal{B}_\lambda(L)\mathcal{B}_\lambda(\beta)^{-1}$.

Proof. See Appendix B.

Representation (5.3) makes clear the distinction between representative agent economies and models with heterogenous agents and heterogeneous beliefs. In a representative agent setting, the only rational expectations solution to this model would be some linear combination of market fundamentals, s_t. Introducing heterogenous beliefs allows for potentially substantial deviations from this traditional RE equilibrium. This representation makes clear that agents' beliefs about future prices are tied to the beliefs of other agents. For both agents, "market fundamentals" are a combination of the exogenous process, s_t, and the endogenous forecast errors of the other agent type.

Representation (5.3) also suggests that *both* informed and uninformed agents engage in some form of higher-order thinking as their behavior can be represented in terms of fundamentals that are function of the beliefs of the other agents. The extent to which agents are successful in learning from other agents' forecasts depends upon the information structure. The following proposition formalizes this concept, making clear the role of HOBs in an IE and demonstrating why HOBs lead to the break down in the law of iterated expectations for the average expectations operator.

Proposition 2. *If the information equilibrium given by Theorem 2 holds for* $|\lambda| < 1$*, then*

i. the informed agents form noiseless higher-order beliefs, while the uninformed form noisy higher-order beliefs;

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 B. □

The *average* expectation of the price at $t+1$ determines equilibrium according to (4.2). In turn, the agents recognize that the price at $t+1$ will be itself a function of the average expectations of the price at $t+2$. So if an agent could observe the average forecast of the price at $t+2$, her prediction performance of the price at $t+1$ would improve. Following this reasoning, the optimal expectation of both agent types must follow

$$\mathbb{E}_t^{\mathcal{I}} p_{t+1} = \mathbb{E}_t^{\mathcal{I}}[\beta \bar{\mathbb{E}}_{t+1} p_{t+2} + s_{t+1}], \qquad \mathbb{E}_t^{\mathcal{U}} p_{t+1} = \mathbb{E}_t^{\mathcal{U}}[\beta \bar{\mathbb{E}}_{t+1} p_{t+2} + s_{t+1}] \tag{5.4}$$

Following Theorem 2 the functional form of the equilibrium price is $p_t = (L-\lambda)Q(L)\varepsilon_t$ where $|\lambda| < 1$; the appendix shows that the time $t+1$ average expectation of the price at $t+2$ can be written as the actual price at $t+2$ minus the average market forecast error, namely

$$\bar{\mathbb{E}}_{t+1} p_{t+2} = p_{t+2} + \mu Q_0 \lambda \varepsilon_{t+2} - (1-\mu) Q_0 \mathcal{B}_\lambda(L) \varepsilon_{t+2} \tag{5.5}$$

The average market forecast error on the RHS of (5.5) has two components: the first term represents the error made by the informed agents, $Q_0 \lambda \varepsilon_{t+2}$, appropriately weighted by the mass of informed agents in the market, μ; the second term, $Q_0 \mathcal{B}_\lambda(L)\varepsilon_{t+2}$, represents the forecast error of the uninformed agents, weighted by the mass of uninformed agents in the market, $1-\mu$.

We know from the form of the lag polynomial $\mathcal{B}_\lambda(L) \equiv (L-\lambda)/(1-\lambda L)$ that the forecast error of uninformed agents contains a linear combination of current and past innovations (due to confounding dynamics), which makes the uninformed agents' error partially predictable for the informed agents. That is, the $t+2$ forecast error of the uninformed is correlated with respect to the time t information set of the informed agents. Hence, the informed agents will always achieve smaller forecast errors if they correct their expectation of the average price according to the forecast errors of the uninformed. More explicitly, the informed agents' time t expectation of the $t+1$ market average expectation takes the form

$$\mathbb{E}_t^{\mathcal{I}}\mathbb{E}_{t+1}^{\mathcal{U}}p_{t+2} = \mathbb{E}_t^{\mathcal{I}}p_{t+2} - Q_0\left(\frac{1-\lambda^2}{1-\lambda L}\right)\lambda\varepsilon_t. \tag{5.6}$$

In forming their expectations for the $t+2$ price conditional on time $t+1$ information, the uninformed agents incur the error $Q_0\left(\frac{1-\lambda^2}{1-\lambda L}\right)\varepsilon_{t+1}$. Informed agents can predict this error at time t by conditioning down with respect to their information set (all current and past innovations up to ε_t), which explains the multiplication by $\lambda\varepsilon_t$. While we have characterized first-order beliefs only, the autoregressive nature of the error incurred by the uninformed suggests that higher-order beliefs follow (5.6) closely with λ^j replacing λ, where j is the higher-order beliefs horizon (see Appendix B for explicit calculations).

The intuition that serially correlated forecast errors is driving the formation of higher-order beliefs seems to suggest that uninformed agents cannot engage in higher-order thinking. That is, uninformed agents possess strictly smaller information sets and are therefore unable to learn anything from the informed agents' forecast errors. This is false. The uninformed agents do engage in higher-order thinking. The uninformed agents form "noisy" HOBs because they are not able to disentangle the forecasts of the informed agents from the exogenous s_t process. The existence condition, (3.7), stipulates that the uninformed agents cannot completely separate out the effects of the informed agents' expectations from the exogenous process, s_t. These confounding dynamics ensure that the uninformed only observe the sum and not the individual components of the sum; being able to disentangle these two processes would imply a convergence to the full information equilibrium of (3.3). The uninformed agents therefore are solving an endogenous signal extraction problem as part of the formation of HOBs. However, the optimal expectation of the uninformed does not ignore the information coming from the informed agents' expectation. Appendix B shows that taking expectations in (5.4) delivers

$$\mathbb{E}_t^{\mathcal{U}}\mathbb{E}_{t+1}^{\mathcal{I}}p_{t+2} = \mathbb{E}_t^{\bar{\mathcal{U}}}p_{t+2} + Q_0\left(\frac{1-\lambda^2}{\lambda(1-\lambda L)}\right)\varepsilon_t \tag{5.7}$$

The term $\mathbb{E}_t^{\bar{\mathcal{U}}}p_{t+2}$ represents the prediction of the uninformed agents if they were to ignore the existence of informed agents in the market (and thus the information that is generated by the informed forecast errors). The second term represents the higher order thinking of uninformed agents as they recognize the presence of informed agents and benefit from the information contained in their forecast errors.[15] We show below just how much information the uninformed are learning by forming HOBs. We refer to the HOBs formed by the uninformed

[15] Notice that the term $\mathbb{E}_t^{\bar{\mathcal{U}}}p_{t+2}$ is defined in order to isolate the higher thinking process of the uninformed, but, strictly speaking, is an expectations that is not measurable with respect to the information set of uninformed agents in equilibrium when $\mu > 0$. The reason is that the information set of the uninformed agents is endogenous to the particular equilibrium we are considering. This is not true for the information set of the informed agents, and therefore there is no need to distinguish the informed information set in $\mathbb{E}_t^{\mathcal{I}}p_{t+2}$ of (5.6) from the equilibrium one as they always coincide.

agents as noisy HOBs to contrast the HOBs formed by informed agents, who observe the forecast errors of the informed directly.

An immediate consequence of agents forming HOBs is that the law of iterated expectations fails to hold with respect to the average expectations operator. This can easily be seen by substituting (5.5) into (5.1) and taking expectations, which delivers

$$p_t = \beta^2 \bar{\mathbb{E}}_t p_{t+2} + \beta \bar{\mathbb{E}}_t s_{t+1} + s_t - \beta\mu(1-\mu)Q_0\left(\frac{1-\lambda^2}{1-\lambda L}\right)\lambda\varepsilon_t + \beta\mu(1-\mu)Q_0\left(\frac{1-\lambda^2}{\lambda(1-\lambda L)}\right)\varepsilon_t \tag{5.8}$$

The last two components of (5.8) are due to the informed and uniformed agents' adjusting expectations due to HOBs, without these terms the law of iterated expectations would hold. The degree to which the law of iterated expectations fails is determined by the relevance of HOBs and is therefore related to the proportion of informed agents, μ, and to the extent of the confounding dynamics, measured by λ.

5.2 Higher Order Beliefs and Information Diffusion We are now in a position to study how the formation of HOBs affects the dissemination of information in equilibrium. Our aim is to compare the information equilibrium of Theorem 2 to an equilibrium where HOBs at all horizons are forcefully removed – we call this a "No-HOBs Equilibrium." Holding the same exogenous information assumption across the equilibria, a lower mean square forecast error will correspond to greater information diffusion. To conceptualize and then solve for the 'No-HOBs Equilibrium" we proceed as follows: First, using Proposition 2 we derive a representation for the information equilibrium price that isolates the noiseless HOBs of the informed agents and noisy HOBs of the uninformed. Next, we shut down these higher-order beliefs sequentially and solve for two No-HOBs equilibria–one that removes the informed agent's HOBs and one that removes both informed and uninformed agents' HOBs.

Iteratively applying Proposition 2 one can show that the equilibrium price is given by

$$\begin{aligned} p_t &= \sum_{j=1}^{\infty} (\mu\beta)^j \mathbb{E}_t^{\mathcal{I}}(s_{t+j}) + s_t \\ &+ \sum_{j=1}^{\infty} \beta^j (1-\mu)^j \mathbb{E}_t^{\mathcal{U}}\left(\beta\mu \mathbb{E}_{t+j}^{\mathcal{I}} p_{t+j+1} + s_{t+j}\right) + \mathbb{E}_t^{\mathcal{I}} \sum_{h=1}^{\infty} (\mu\beta)^h \mathbb{E}_{t+h}^{\mathcal{U}} \sum_{j=h}^{\infty} (1-\mu)^{j-h+1} \beta^{j-h+1} \left[\beta\mu \mathbb{E}_{t+j+1}^{\mathcal{I}} p_{t+j+2} + s_{t+j+1}\right] \end{aligned} \tag{5.9}$$

The last two terms in (5.9) capture the entire HOBs structure into the infinite future. When only fully informed buyers are present ($\mu = 1$), the expression coincides with the price under full information in (3.3). Likewise, when only uninformed buyers are present ($\mu = 0$), the expression coincides with the price under symmetric incomplete information in (4.21).

The weights assigned to the expectations in the three terms clarify the higher-order reasoning. The uninformed agents will form expectations of the sum of the s_t's and the entire path of future expectations of the informed agents, discounted at $\beta(1-\mu)$. The informed agents will form the "standard" discounted expectation of future s_t's with weight $\mu\beta$, but will also correct this forecast based upon the forecasts of the uninformed, which is the last term in (5.9). This term shows that the informed agents will correct the *entire path* of the uninformed agents' expectations, not just the time t forecast errors.

The formation of HOBs provides uninformed agents with two additional sources of information that they would not have otherwise. The first source comes from forming noisy HOBs themselves (the penultimate term of (5.9)) and the second source comes from the informed agents forming HOBs (the last term of (5.9)). Recall that the informed agents' HOBs correct for the serial correlation in the uninformed agents' forecast error. In equilibrium, this information gets impounded into the price and is partially revealed to the uninformed agents. The obvious question is: How much information is revealed through the formation of HOBs?

Given that we have an analytical solution at hand, we can answer this question by forcing each agent type to not engage in higher-order thinking. The following proposition solves for two boundedly rational equilibria to isolate the two sources of information coming from the HOBs. The first equilibrium solves (5.9) but sets the last term to zero, which isolates the role of HOBs formed by the informed buyers. The second equilibrium removes both the last term and the penultimate term in (5.9), which takes all HOBs out of the model. By taking the difference between the two equilibria, one can isolate the role of the HOBs formed by the uninformed buyers.

Proposition 3. No-Informed HOBs Equilibrium. *Assume that the fully informed buyers do not form higher order beliefs (i.e., solve* (5.9) *removing the last term). Under the exogenous information assumption* (4.1)*, i.e.* $U_t^i = \mathbb{V}_t(\varepsilon)$ *for* $i \in \mu$ *and* $U_t^i = \{0\}$ *for* $i \in 1-\mu$*, a unique boundedly-rational equilibrium always exists and is determined as follows. If there exists a* $|\tilde{\lambda}| < 1$ *such that*

$$A(\tilde{\lambda}) - \frac{\mu\beta A(\beta)}{\tilde{\lambda} - (1-\mu)\beta} = 0 \tag{5.10}$$

then the equilibrium price is given by

$$p_t = \frac{1}{L-\beta}\left(LA(L) - \beta A(\beta)\frac{k(L)}{k(\beta)}\right)\varepsilon_t \tag{5.11}$$

where $k(L) = \frac{\mu\tilde{\lambda}}{1-\tilde{\lambda}\beta} - (1-\mu)\frac{L-\tilde{\lambda}}{1-\tilde{\lambda}L}$*. If* (5.12) *does not hold for any* $|\tilde{\lambda}| < 1$*, the equilibrium is the full information equilibrium* (3.3)*.*

No-HOBs Equilibrium. *Assume that neither the informed nor uninformed buyers form higher-order beliefs (i.e., solve* (5.9) *removing the last term and setting the penultimate term to* $\sum_{j=1}^{\infty}\beta^j(1-\mu)^j\mathbb{E}_t^{\mathcal{U}}s_{t+j}$*). Under the exogenous information assumption* (4.1)*, i.e.* $U_t^i = \mathbb{V}_t(\varepsilon)$ *for* $i \in \mu$ *and* $U_t^i = \{0\}$ *for* $i \in 1-\mu$*, a unique boundedly-rational equilibrium always exists and is determined as follows. If there exists a* $|\lambda^*| < 1$ *such that*

$$A(\lambda^*) - \frac{\mu\beta A(\beta)}{\lambda^*} = 0 \tag{5.12}$$

then the equilibrium price given by

$$p_t = \frac{1}{L-\beta}\Big(LA(L) - \beta A(\beta)\kappa(L)\Big)\varepsilon_t \tag{5.13}$$

where $\kappa(L) = \mu + (1-\mu)\frac{L-\lambda^*}{1-\lambda^* L}$*. If* (5.13) *does not hold for any* $|\lambda^*| < 1$*, the equilibrium is the full information equilibrium* (3.3)*.*

Proof. See Appendix B. □

Proposition 3 allows us to state the main result of this section.

Corollary 4. *Assume* $A(L) = (1+\theta L)/(1-\rho L)$ *with* $\rho \in [0,1]$*. If an IE exists with* $\lambda \in (-1,1)$*, then higher-order beliefs always enhance information diffusion.*

Proof. See appendix B. □

The corollary essentially states that $|\lambda^*| < |\tilde{\lambda}| < |\lambda|$. Recall that as $|\lambda| \to 1$, confounding dynamics diminish and disappear altogether in the limiting case, as the discrepancy between the information set of the informed and uninformed gets smaller. We measure information diffusion as the relative difference between the informed and uninformed agents' variance of forecast error. Given that for each variant of (5.9) the price process can be written as $p_t = (L-\lambda)Q(L)\varepsilon_t$, it is straightforward to show that for each economy described in Corollary 4 and Theorem 2, the ratio of forecast errors is given by λ^2,

$$\frac{\mathbb{E}(p_{t+1} - \mathbb{E}_t^{\mathcal{I}} p_{t+1})^2}{\mathbb{E}(p_{t+1} - \mathbb{E}_t^{\mathcal{U}} p_{t+1})^2} = \lambda^2 \tag{5.14}$$

When the informed higher-order thinking is removed $|\lambda|$ declines to $|\tilde{\lambda}|$ which means that informed higher order thinking reduces the extent of the confounding dynamics and therefore has a positive effect on information diffusion in equilibrium. Intuitively, engaging in guessing the expectation of the average expectation of the average expectation and so on helps information diffusion because it forces informed agents to use their private information to guess the forecast errors of other agents. In so doing, more information is encoded into equilibrium prices and thus the variance of the forecast errors is reduced. When the uninformed higher order thinking is removed together with the informed higher order thinking, $|\lambda|$ decreases further to $|\lambda^*| < |\tilde{\lambda}|$. Even though uninformed agents form noisy HOBs, doing so increases their information and reduces their forecast errors.

To quantify the effects of higher order thinking, consider a variant of the numerical example presented in Section (4.5). Using the process for s_t specified in (4.17), let $\rho = 0.8$, $\theta = \sqrt{11}$ and $\mu = 0.06$. The ratio of the variance of the forecast errors, (5.14), is 0.84 when all HOBs are present. This value falls to 0.49 when only uninformed agents form HOBs, (5.10), and 0.137 when neither informed nor uninformed form HOBs, (5.12). By this measure, HOBs reduce the information discrepancy between the informed and uninformed agents by a factor

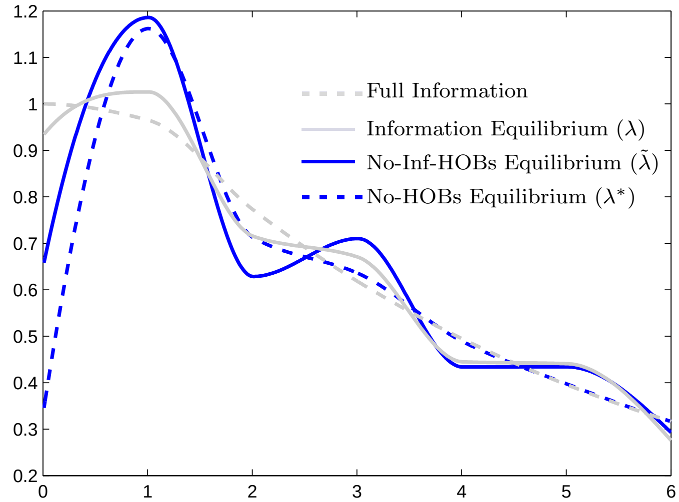

Figure 5: Impulse response of market price to one time unitary innovation in ε_t. The light dashed line is the price under full information; the solid light line is the price with all the HOBs present under confounding dynamics measured by λ; the solid dark line is the No-Informed HOB's equilibrium price under confounding dynamics measured by $\tilde{\lambda}$; the dashed dark line is the No-Hob's equilibrium price under confounding dynamics measured by λ^*. The parameter values are $s_t = 0.8s_{t-1} + \varepsilon_t + \sqrt{11}\varepsilon_{t-1}$, $\beta = 0.985$ and $\mu = 0.06$.

of seven. As a visual confirmation, figure 5 displays the impulse response of the full information equilibrium, the information equilibrium of Theorem 2 and the No-HOBs equilibria of Proposition 3 to a one time shock to the fundamentals ε_0. The impulse responses are normalized with respect to the response at impact of the full information price. The dynamics of the equilibrium with HOBs deviates only modestly from the full information; this is due to the informational effect of a small portion of agents being fully informed. How much of the informational effect is due to the higher order thinking of fully informed agents? The impulse response for the No-HOBs equilibria reveals that higher order thinking is remarkably important for informational diffusion. Without any HOBs, the market price would under-react at impact by approximately 70% of the price with HOBs, and it would over-react a period later of around $15-20\%$. Higher order thinking is therefore essential in keeping the market price from undergoing excessive fluctuations due to slow informational diffusion.

5.3 Higher Order Beliefs and Dispersed Information In this section we analytically characterize the higher order thinking of an individual agent in the dispersedly information setup of Theorem 1. The aggregation result of Theorem 2 will be helpful in applying the previous results at the dispersedly informed agent level. However, as we consider the individual agent there will be some crucial differences. For example, agent i will use her exogenous signal to forecast the forecasts of the *market* expectation and this forecast will be different from the direct forecast of agent i, and, for that matter of both the informed and uninformed buyers considered above. We summarize the description of the HOBs for the dispersed information case in the following proposition.

Proposition 4. *If the information equilibrium given by Theorem 1 holds for $|\lambda| < 1$, then:*

i. all agents form noisy higher-order beliefs;

ii. the average expectations operator does not satisfy the law of iterated expectations.

Proof. See Appendix A. □

The intuition developed for the equivalence result of Theorem 2 is helpful in indicating what is happening in the dispersed case. Take any arbitrary agent i. This agent is instructed by the optimality of signal extraction to act as informed with probability τ. She will recognize that a portion $1-\tau$ of agents is contemporaneously acting as uninformed. It follows that as an informed agent, agent i should forecast the forecast error of the agents acting as uninformed and embed it into her expectations about the future. At the same time, she is acting as uninformed as well, i.e. she is part of the portion of $1-\tau$ agents of whom she is forecasting the forecast errors. However, the relevance of her forecast error is infinitesimal and so it is irrelevant for her reasoning as informed. To formalize this intuition one can show that

$$\mathbb{E}_{it}p_{t+1} = \beta\mathbb{E}_{it}(p_{t+2}) + \tau\mathbb{E}_{it}s_{t+1} - \tau\beta Q_0\,(1-\tau)\,\mathbb{E}_{it}\left(\frac{L-\lambda}{1-\lambda L}\varepsilon_{t+2}\right) + (1-\tau)\mathbb{E}_{it}[Q_0\mu\beta\varepsilon_{t+2} + s_{t+1}] \tag{5.15}$$

In forming their predictions at $t+1$ agents acting as uninformed will incur in the prediction error $Q_0\left(\frac{L-\lambda}{1-\lambda L}\varepsilon_{t+2}\right)$. At time t agent i will take this into account and use her own information to forecast the forecast error and adjust her expectations of the average expectations accordingly, i.e. by weighting the forecast of the forecast errors by $(1-\tau)$. Similarly the last term shows that agents acting as informed will incur prediction error $\tau\beta Q_0$, and agent i will take this into account in forming forecasts of future prices.

Proposition 4 together with equation (5.15) offer intuition that applies more generally to signal extraction problems based on private and public noisy signals.[16] The equilibrium price in Theorem 1 in presence of confounding dynamics represents noisy public information. Such public information is common knowledge and, therefore, it represents the information set that all the agents possess if they were to disregard their private information. Take once again an arbitrary agent i. When agent i acts as fully informed, she engages in predicting the error of the other agents acting as fully uninformed. How will agent i determine the information set of the agents acting as uninformed? Common knowledge of rationality will suggest that "uninformed" acting agents just use the public signal as they act as if their private signal was fully uninformative. This reasoning will suggest to agent i that the forecast error to predict takes the form of the forecast error that would result by using only publicly available information.

[16] For instance all the literature on global games that was sparked by Morris and Shin (1999).

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. 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. The analytics, in turn, permitted insights into higher-order belief dynamics and the transmission of information in general. Given the generality of the forward-looking equation at the heart of our model, we expect the results presented in this paper to be relevant in the analysis of many dynamic economic applications under incomplete information.

References

Allen, F., S. Morris, and H. Shin (2006): "Beauty Contests and Iterated Expectations in Asset Markets," *Review of Financial Studies*, 19(3), 719–752.

Angeletos, G., and J. La'O (2009a): "Incomplete Information, Higher-Order Beliefs and Price Inertia," *Journal of Monetary Economics*, 56, S19–S37.

——— (2009b): "Noisy Business Cycles," *NBER Macroeconomics Annual*, 24.

Angeletos, G., and A. Pavan (2007): "Efficient Use of Information and Social Value of Information," *Econometrica*, 75(4).

Bacchetta, P., and E. van Wincoop (2006): "Can Information Heterogeneity Explain the Exchange Rate Puzzle?," *American Economic Review*, 96(3), 552–576.

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

Canova, F. (2003): *Methods for Applied Macroeconomic Research*. Princeton University Press, Princeton, New Jersey, first edn.

Conway, J. (1991): *The Theory of Subnormal Operators*. American Mathematical Society.

Futia, C. A. (1981): "Rational Expectations in Stationary Linear Models," *Econometrica*, 49(1), 171–192.

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

Hansen, L. P., and T. J. Sargent (1991): "Two Difficulties in Interpreting Vector Autoregressions," in *Rational Expectations Econometrics*, ed. by L. P. Hansen, and T. J. Sargent. Westview Press.

Hassan, T. A., and T. M. Mertens (2011): "The Social Cost of Near-Rational Investment," NBER Working Paper 17027.

Hellwig, C. (2006): "Monetary Business Cycle Models: Imperfect Information," *New Palgrave Dictionary of Economics*.

Hellwig, C., and V. Venkateswaran (2009): "Setting the Right Prices for the Wrong Reasons," *Journal of Monetary Economics*, 56, S57–S77.

Kasa, K. (2000): "Forecasting the Forecasts of Others in the Frequency Domain," *Review of Economic Dynamics*, 3, 726–756.

Kasa, K., T. B. Walker, and C. H. Whiteman (2008): "Asset Prices in a Time Series Model With Perpetually Disparately Informed, Competitive Traders," University of Iowa Working Paper.

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

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

Lorenzoni, G. (2009): "A Theory of Demand Shocks," *American Economic Review*, 99(5), 2050–2084.

Lucas, Jr., R. E. (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.

——— (1978): "Asset Prices in an Exchange Economy," *Econometrica*, 46(6), 1429–1445.

Maćkowiak, B., and M. Wiederholt (2007): "Business Cycle Dynamics under Rational Inattention," Working paper.

Morris, S., and H. S. Shin (1999): "Global Games: Theory and Applications," in *The Asian Financial Crisis: Causes, Contagion and Consequences*, ed. by P.-R. Agenor, M. Miller, D. Vines, and A. Weber. Cambridge University Press, Cambridge.

——— (2002): "The Social Value of Public Information," *American Economic Review*, 92, 1521–1534.

Nimark, K. (2007): "Dynamic Higher Order Expectations," Working paper.

——— (2008): "Dynamic Pricing and Imperfect Common Knowledge," *Journal of Monetary Economics*, 55(2), 365–382.

Pearlman, J. G., and T. J. Sargent (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*. Macmillan, London, second edn.

Rogers, E. M. (2003): *Diffusion of Innovations*. Free Press, New York, fifth edn.

Rondina, G. (2009): "Incomplete Information and Informative Pricing," Working Paper. UCSD.

Rondina, G., and T. B. Walker (2011): "An Information Equilibrium Model of the Business Cycle," Working Paper.

Rozanov, Y. A. (1967): *Stationary Random Processes*. Holden-Day, San Francisco.

Sargent, T. J. (1991): "Equilibrium with Signal Extraction from Endogenous Variables," *Journal of Economic Dynamics and Control*, 15, 245–273.

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*. University of Minnesota Press, Minneapolis.

Woodford, M. (2003): "Imperfect Common Knowledge and the Effects of Monetary Policy," in *Knowledge, Information, and Expectations in Modern Macroeconomics*, ed. by P. Aghion, R. Frydman, J. Stiglitz, and M. Woodford. Princeton University Press, Princeton, N.J.

A Proofs

A.1 Full Information Price

We want to solve for the infinite summation

$$p_t = \sum_{j=0}^{\infty} \beta^j \mathbb{E}_t(s_{t+j}). \tag{A.1}$$

In lieu of characterizing each term in the summation, we take advantage of the Riesz-Fischer Theorem and posit that the solution to (3.2) has the functional form $p_t = P(L)\varepsilon_t$.[17] Using the Wiener-Kolmogorov optimal prediction formula, expectations take the form $\mathbb{E}[p_{t+1}|\mathbb{V}_t(\varepsilon)] = L^{-1}[P(L) - P_0]\varepsilon_t$. Substituting the expectation into the equilibrium equation (2.1) yields a functional equation for $P(z)$.[18] 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

$$P(z) = \frac{zA(z) - \beta P_0}{z - \beta}. \tag{A.2}$$

[17] Note that there is no need to include in our guess the possibility of a zero $|\lambda| < 1$ as it would be informationally irrelevant given the full information provided to the agents.

[18] In our notation we distinguish between L and z to make clear that L is an operator, while z is a complex number.

Throughout the paper we always restrict our attention to stationary equilibria. Stationarity corresponds to the requirement that $P(z)$ has no unstable roots in the denominator. If $|\beta| \geq 1$, then (A.2) is stationary and the free parameter P_0 can be set arbitrarily. Uniqueness, then, requires $|\beta| < 1$, in which case the free parameter P_0 is set to ensure that the unstable root $|\beta| < 1$ cancels. Carrying out these steps one obtains equation (3.3) in the text.

A.2 Theorem 1

We report a statement of the theorem for the more general case of n possible λ's in the initial guess. The proof of theorem 1 in the main text is obtained by just setting $n = 1$ below.

Theorem 3. *Under the exogenous information assumption $U_t^i = \{0\}$ $\forall i$, a unique Information Equilibrium 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*

$$A(\lambda_i) = 0, \tag{A.3}$$

then the information equilibrium price process is

$$p_t = Q(L)\prod_{i=1}^{n}(L-\lambda_i)\varepsilon_t = \frac{1}{L-\beta}\left\{LA(L) - \beta A(\beta)\frac{\prod_{i=1}^{n}\mathcal{B}_{\lambda_i}(L)}{\prod_{i=1}^{n}\mathcal{B}_{\lambda_i}(\beta)}\right\}\varepsilon_t \tag{A.4}$$

where

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

If condition (4.3) does not hold for any $|\lambda_i| < 1$, then the IE is given by (3.3).

Proof. Substituting the conditional expectation (**??**) into the equilibrium equation 2.1 yields the z-transform in ε_t-space

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

Working out the algebra yields

$$Q(z)(z-\beta)\prod_{i=1}^{n}(z-\lambda_i) = zA(z) - Q_0\prod_{i=1}^{n}\mathcal{B}_{\lambda_i}(z) \tag{A.5}$$

For $|\beta| < 1$, stationarity 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)]. Evaluating at $z = \lambda_i$ gives the restriction on the $A(\cdot)$ process, $A(\lambda_i) = 0$ for all i, which corresponds to (4.3). By Proposition 10.4 of Conway (1991), this restriction guarantees that the knowledge of the model does not reveal any additional information than the posited price sequence. Finally, evaluating at $z = \beta$ gives

$$Q_0 = \frac{\beta A(\beta)}{\prod_{i=1}^{n}\mathcal{B}_{\lambda_i}(\beta)} \tag{A.6}$$

Substituting this into (A.5) yields (A.4).

□

A.3 Theorem 2

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

$$\mathbb{E}_t^{\mathcal{I}}(p_{t+1}) = L^{-1}[(L-\lambda)Q(L) + \lambda Q_0]\varepsilon_t$$
$$\mathbb{E}_t^{\mathcal{U}}(p_{t+1}) = L^{-1}[(L-\lambda)Q(L) - Q_0\mathcal{B}_\lambda(L)]\varepsilon_t$$

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

$$(z-\lambda)Q(z) = \beta\mu z^{-1}[(z-\lambda)Q(z)+\lambda Q_0] + \beta(1-\mu)z^{-1}[(z-\lambda)Q(z) - Q_0\mathcal{B}_\lambda(z)] + A(z) \tag{A.7}$$

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

$$\begin{aligned}(z-\beta)Q(z) &= \frac{1}{z-\lambda}\left\{zA(z) + \beta Q_0[\mu\lambda - (1-\mu)\mathcal{B}_\lambda(z)]\right\} \\ &= \frac{1}{z-\lambda}\left\{zA(z) + \beta Q_0 h(z)\right\}\end{aligned} \tag{A.8}$$

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

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

which is (3.7). Substituting this into (A.8) delivers (3.8).

A.4 Theorem 1 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

$$\begin{pmatrix} \varepsilon_{it} \\ p_t \end{pmatrix} = \begin{pmatrix} \sigma_\varepsilon & \sigma_v \\ (L-\lambda)\,p\,(L) & 0 \end{pmatrix} \begin{pmatrix} \hat{\varepsilon}_t \\ \hat{v}_{it} \end{pmatrix} = \Gamma(L) \begin{pmatrix} \hat{\varepsilon}_t \\ \hat{v}_{it} \end{pmatrix}. \tag{A.9}$$

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

$$\begin{pmatrix} \varepsilon_{it} \\ p_t \end{pmatrix} = \Gamma^*(L) \begin{pmatrix} w_{it}^1 \\ w_{it}^2 \end{pmatrix}. \tag{A.10}$$

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

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

where

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

The vector of fundamental innovations is then given by

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

The expectation term for agent i is provided by the second row of the Wiener-Kolmogorov prediction formula applied to the fundamental representation (A.10), which is

$$\mathbb{E}(p_{t+1}|\varepsilon_i^t, p^t) = [\Gamma_{21}^*(L) - \Gamma_{21}^*(0)]\,L^{-1} w_{it}^1 + [\Gamma_{22}^*(L) - \Gamma_{22}^*(0)]\,L^{-1} w_{it}^2. \tag{A.11}$$

It is straightforward to show that

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

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

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

The average market expectation is then

$$\bar{\mathbb{E}}(p_{t+1}) = [(L-\lambda)p(L)+\lambda p_0]\,L^{-1}\frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2+\sigma_v^2}\hat{\varepsilon}_t + [(1-\lambda L)p(L)-p_0]\,L^{-1}\frac{\sigma_v^2}{\sigma_\varepsilon^2+\sigma_v^2}\frac{L-\lambda}{1-\lambda L}\hat{\varepsilon}_t. \tag{A.12}$$

Now, if we let

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

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 (A.7) 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

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

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

$$\begin{pmatrix} \varepsilon_{it} \\ p_t \\ m_{it} \end{pmatrix} = \begin{pmatrix} \sigma_\varepsilon & \sigma_v \\ (L-\lambda)\,p(L) & 0 \\ A(L)\,\sigma_\varepsilon & \frac{\sigma_\varepsilon\sigma_v}{\sigma_\varepsilon^2+\sigma_v^2}\left(\frac{1-\lambda^2}{1-\lambda z}\right)\beta p_0 \end{pmatrix}\begin{pmatrix} \hat{\varepsilon}_t \\ \hat{v}_{it} \end{pmatrix}. \tag{A.13}$$

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

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

Using the fact that $p_0 = Q_0\sigma_\varepsilon$ one can immediately see that this condition is equivalent to (3.7). Therefore, in a dispersed information equilibrium, it is always true that the enlarged information matrix (A.13) carries the same information as the information matrix (A.9). This completes the proof of Theorem 3.

A.5 Proposition 1 Once the analytic form for $\Gamma^*_{21}(L)$ and $\Gamma^*_{22}(L)$ are known one can compute $\mathbb{E}(p_{t+j}|\varepsilon^t_i, p^t)$ for any $j = 1, 2, ...$ We show the $j = 1$ case here. Substitute $\Gamma^*_{21}(L)$ and $\Gamma^*_{22}(L)$ into (A.11) and collecting the terms that constitute (A.12), one gets

$$\begin{aligned}\mathbb{E}(p_{t+1}|\varepsilon^t_i, p^t) &= \bar{\mathbb{E}}(p_{t+1}) + \frac{\sigma_\varepsilon}{\sigma^2_\varepsilon + \sigma^2_v} L^{-1}[(L-\lambda)p(L) + \lambda p_0 - (L-\lambda)p(L) + p_0 \frac{L-\lambda}{1-\lambda L}]\sigma_v \hat{v}_{it} \\ &= \bar{\mathbb{E}}(p_{t+1}) + \frac{\sigma_\varepsilon}{\sigma^2_\varepsilon + \sigma^2_v} L^{-1}[\lambda p_0 + p_0 \frac{L-\lambda}{1-\lambda L}]\sigma_v \hat{v}_{it} \\ &= \bar{\mathbb{E}}(p_{t+1}) + \mu Q_0 \frac{1-\lambda^2}{1-\lambda L} v_{it}, \end{aligned} \tag{A.14}$$

which completes the proof for the first statement of the theorem for $j = 1$. The variance of the term $\mu Q_0 \frac{1-\lambda^2}{1-\lambda L} v_{it}$ can be readily computed since the innovations v_{it} are independently distributed with variance σ^2_v.

A.6 Result in Example The proof follows immediately from restriction (3.7) in Theorem (2). Condition (**R.1**) is derived by taking the limit of (3.7) as $\mu \to 0$. Substituting the parameters of the example, condition (3.7) with $\mu = 0$ is given by $(1+\theta\lambda)/(1-\rho\lambda) = 0$. Clearly, $|\lambda| < 1$ will not be a possibility when $\theta \in (0, 1)$, hence (**R.1**). Notice that, because $\theta > 0$, then $\lambda < 0$ from (3.7). It follows that $\lambda = -1$ will be the critical value to dictate whether an equilibrium with confounding dynamics exists or not. Taking (3.7) and setting $\lambda = -1$ one obtains the expression for μ^*. For any $\mu \geq \mu^*$ one has $\lambda < -1$, while for $\mu < \mu^*$ one has $0 > \lambda > -1$ which is (**R.2**).

B Online Appendix (Not for Publication)

B.1 Asset Demand Derivation and Market Clearing The ubiquitous equilibrium equation (2.1) can be derived from many micro-founded models. It falls from the Lucas (1978) asset pricing model where agents are risk neutral and the shares are traded cum-dividend. Alternatively, Futia (1981) envisioned the equilibrium arising from land speculation. He assumed a fixed quantity of land and two types of traders–speculative and nonspeculative. Nonspeculative demand is assumed to arise from noise traders; that is, traders whose demand is independent of current and past prices. This demand never exceeds total supply, and therefore the difference between total supply and the nonspeculative demand is the market fundamental, s_t.

The demand for the speculative trader can be derived from a myopic investor who may choose to hold wealth in either a riskless asset which earns the return r or a risky asset. The wealth of agent i evolves according to

$$w_{i,t+1} = z_{i,t}(p_{t+1}) + (w_{i,t} - z_{i,t}p_t)(1+r)$$

where p_t is the price of the risky asset at time t and $z_{i,t}$ is the number of units of the risky asset held at time t.

The speculative agents seeks to maximize, by choice of z_{it}, the expected value of a constant absolute risk aversion (CARA) utility function

$$-E_t^i \exp(-\gamma w_{i,t+1}), \tag{B.1}$$

where γ is the risk aversion parameter, and E_t^i denotes the time t conditional expectation of agent i. All random variables in the model are assumed to be distributed normally, so that (B.1) can be calculated from the (conditional) moment generating function for the normal random variable $-\gamma w_{i,t+1}$. That is,

$$-E_t^i \exp(-\gamma w_{i,t+1}) = -\exp\{-\gamma E_t^i(w_{i,t+1}) + (1/2)\gamma^2 \mathrm{v}_t(w_{i,t+1})\}$$

where v_t denotes conditional variance. Note that $\mathrm{v}_t(w_{t+1}) = z_{i,t}^2 \mathrm{v}_t(p_{t+1})$. Stationarity implies the conditional variance term will be a constant; thus write $\mathrm{v}_t(w_{i,t+1}) \equiv z_{i,t}^2\delta$. The agent's demand function for the risky asset follows from the first-order necessary conditions for maximization and is given by

$$z_{i,t} = \frac{1}{\gamma\delta}[E_t^i p_{t+1} - \alpha p_t] \tag{B.2}$$

where $\alpha \equiv 1 + r > 1$.

Market clearing equates supply and demand, which yields

$$p_t = \alpha^{-1}\int_0^1 E_t^i p_{t+1} di - \alpha^{-1}\gamma\delta s_t \tag{B.3}$$

This relates to (2.1) by $\alpha^{-1} = \beta$ and one can think of s_t in (2.1) as being scaled by the risk aversion coefficient, γ, the opportunity cost associated with investing in the risky asset α, and the conditional variance term, δ. Clearly, δ is an endogenous object, but we abstract from this complication to make the analysis as transparent as possible.

B.2 Equivalence between Confounding Dynamics and Standard Signal Extraction It is helpful to establish a connection between the information contained in $\tilde{\varepsilon}_t$ when $|\lambda| < 1$ and a signal extraction problem cast in a more familiar setting. Suppose that agents observe an infinite history of the signal

$$z_t = \varepsilon_t + \eta_t, \tag{B.4}$$

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

$$x_t = \varepsilon_t + \theta\varepsilon_{t-1}, \tag{B.5}$$

Proposition 5. *The information content of* (B.5) *is equivalent to* (B.4)*, where equivalence is defined as equality of variance of the forecast error conditioned on the infinite history of the observed signal, i.e.*

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

when

$$\theta^2 = \frac{1}{\tau} \tag{B.6}$$

and where $\tau = \sigma_\varepsilon^2/(\sigma_\varepsilon^2 + \sigma_\eta^2)$.

Proof. We need to show that the representations (B.5) and (B.4) are equivalent in terms of unconditional forecast error variance

$$\mathbb{E}\left[\left(\varepsilon_t - \mathbb{E}\left(\varepsilon_t|x^t\right)\right)^2\right] = \mathbb{E}\left[\left(\varepsilon_t - \mathbb{E}\left(\varepsilon_t|z^t\right)\right)^2\right] \tag{B.7}$$

when $\theta^2 = 1 + \sigma_\eta^2/\sigma_\varepsilon^2$.

The optimal forecast $\mathbb{E}[\varepsilon_t|z^t]$ is given by weighting z_t according to the relative variance of ε, $\mathbb{E}(\varepsilon_t|z^t) = \left(\frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2+\sigma_\eta^2}\right)z_t$ and therefore,

$$\mathbb{E}\left[\left(\varepsilon_t - \mathbb{E}\left(\varepsilon_t|z^t\right)\right)^2\right] = \frac{\sigma_\varepsilon^2\sigma_\eta^2}{\sigma_\varepsilon^2 + \sigma_\eta^2} \tag{B.8}$$

Calculating the optimal expectation for ε_t conditional on x^t requires more careful treatment. While there are many moving average representations for x_t that deliver the same observed autocorrelation structure (which is essentially all the information contained in x^t), there exists only one that minimizes the variance of the forecast error in the LHS of (B.7). We first need to take the conditional expectation $\mathbb{E}[\varepsilon_t|x^t]$. This expectation is found by deriving the fundamental moving-average representation and using the Wiener-Kolmogorov optimal prediction formula. The fundamental representation is derived through the use of Blaschke factors

$$x_t = (1+\theta L)\left(\frac{L+\theta}{1+\theta L}\right)\left(\frac{1+\theta L}{L+\theta}\right)\varepsilon_t = (L+\theta)\tilde{\varepsilon}_t \tag{B.9}$$

where $\tilde{\varepsilon}_t$ is defined as in (2.7). Given that (B.9) is an invertible representation then the Hilbert space spanned by current and past x_t is equivalent to the space spanned by current and past $\tilde{\varepsilon}_t$. This implies

$$\mathbb{E}(\varepsilon_t|\tilde{\varepsilon}^t) = \mathbb{E}(\varepsilon_t|x^t) \tag{B.10}$$

To show (B.10) notice that (B.9) can be written as

$$\varepsilon_t = C(L)\tilde{\varepsilon}_t = \left[\frac{(\theta^{-1} + L^{-1})}{1 - (-\theta L)^{-1}}\right]\tilde{\varepsilon}_t \tag{B.11}$$

Thus, while (B.9) does not have an invertible representation in current and past $\tilde{\varepsilon}$ it does have a valid expansion in current and future $\tilde{\varepsilon}$. Notice that

$$\varepsilon_t = (\theta^{-1} + L^{-1})\sum_{j=0}^{\infty}(-\theta)^{-j}\tilde{\varepsilon}_{t+j} = (\theta^{-1} + L^{-1})[\tilde{\varepsilon}_t + (-\theta)^{-1}\tilde{\varepsilon}_{t+1} + \cdots]$$

The optimal prediction formula yields

$$\mathbb{E}(\varepsilon_t|\tilde{\varepsilon}^t) = \left[C(L)\right]_+\tilde{\varepsilon}_t = \theta^{-1}\tilde{\varepsilon}_t = \left[\frac{1}{\theta^2(1+\theta^{-1}L)}\right]x_t \tag{B.12}$$

We must now calculate

$$\mathbb{E}\left[\left(\varepsilon_t - \mathbb{E}\left(\varepsilon_t|x^t\right)\right)^2\right] = \mathbb{E}\left(\varepsilon_t^2\right) + \mathbb{E}\left(\varepsilon_t|x^t\right)^2 - 2\mathbb{E}\left(\varepsilon_t\mathbb{E}\left(\varepsilon_t|x^t\right)\right) \tag{B.13}$$

$$= \sigma_\varepsilon^2 + \frac{1}{\theta^2}\mathbb{E}(\tilde{\varepsilon}_t^2) - \frac{2}{\theta}\mathbb{E}(\varepsilon_t\tilde{\varepsilon}_t) \tag{B.14}$$

Notice that the squared modulo of the Blaschke factor is equal to 1, $\left(\frac{1+\theta z}{z+\theta}\right)\left(\frac{1+\theta z^{-1}}{z^{-1}+\theta}\right) = 1$, and therefore $\mathbb{E}(\tilde{\varepsilon}^2) = \sigma_\varepsilon^2$.

To calculate $\mathbb{E}(\varepsilon_t\tilde{\varepsilon}_t)$ we use complex integration and the theory of the residue calculus,

$$\mathbb{E}(\varepsilon_t\tilde{\varepsilon}_t) = \frac{\sigma_\varepsilon^2}{2\pi i}\oint \frac{1+\theta z}{z+\theta}\frac{dz}{z} = \sigma_\varepsilon^2\left[\lim_{z\to 0}\frac{1+\theta z}{z+\theta}\right] = \frac{\sigma_\varepsilon^2}{\theta}. \tag{B.15}$$

Equations (B.14) and (B.15) give the desired result

$$\mathbb{E}\left[\left(\varepsilon_t - E\left(\varepsilon_t|x^t\right)\right)^2\right] = \left(1 - \frac{1}{\theta^2}\right)\sigma_\varepsilon^2$$

□

To substantiate the claim in the main text one needs just to recognize that by setting $\lambda = 1/\theta$ the result stated follows immediately.

B.3 Full Information Price and Theorem 1 In this section we show that the full information price (3.3) is not an Information Equilibrium under the assumption of Theorem 1. We do this for the case of $s_t = \varepsilon_t + \theta\varepsilon_{t-1}$, with $|\theta| > 1$, in order to keep notation at a minimum. The exogenous information is specified as in Theorem 1, namely $U_t^i = \{0\}\ \forall i$. Plugging the functional form for the exogenous supply in (3.3) we have $p_t = (1+\theta\beta)\,\varepsilon_t + \theta\varepsilon_{t-1}$ (C.1). We show that this price cannot possibly be consistent with our definition of an Information Equilibrium. The argument is by contradiction. Suppose that (C.1) is indeed an Information Equilibrium as defined in 2.1.3 under the assumption that $U_t^i = \{0\}\ \forall i$. First, because expectations are symmetric across agents, information from the model will always reveal the s_t process. The information set in equilibrium is therefore given by the bivariate process for (p_t, s_t). For the endogenous information to be consistent with the equilibrium equation we need to show that there exists a square-summable linear combination of the observable variables that corresponds to the expectations of future price in equilibrium as implied by the closed form solution of the model. The latter is $\mathbb{E}[p_{t+1}|\mathbb{M}_t(p) \vee \mathbb{V}_t(p)] = \mathbb{E}\left(\theta\varepsilon_t|s^t, p^t\right) = \theta\varepsilon_t$, since no information about the future ε_{t+1} is available in the information set, other than the unconditional distribution. The linear combination of observable variables that deliver the above expectation is $\mathbb{E}\left(\theta\varepsilon_t|s^t, p^t\right) = \theta s_t - \theta\frac{1}{\beta}\left(p_{t-1} - s_{t-1}\right)$. This relationship has to hold in equilibrium and when substituted into the equilibrium equation results in $p_t = -\theta p_{t-1} + (1+\theta\beta)\, s_t + \theta s_{t-1}$ (C.2). If p_t defined by (C.1) above is to be an Information Equilibrium, then it must satisfy the dynamic equation (C.2). Notice that this equation contains the explosive autoregressive root $|\theta| > 1$. Remember that 2.1.3 requires the price process to be stationary: how can (C.2) be reconciled with (C.1)? This is possible if the explosive root happens to exactly cancel with the moving average root of s_t. In order for this to be the case, $U_t^i = \{0\}\ \forall i$ must be violated, which results in the contradiction that we were looking for. To see why the exogenous information assumption must be violated, suppose that the asset market takes place at time 0. The equilibrium price for the asset market a period earlier, p_{-1}, is not defined. What would then $(p_{-1} - s_{-1})$ be? To ensure that the unstable root exactly cancels with the moving average root it must be that $(p_{-1} - s_{-1}) = \beta\varepsilon_0$ (C.3). Any other assumption will result in an explosive path for the price process. However, because p_{-1} is not well defined, assuming (C.3) is essentially equivalent to exogenously providing the agents with the knowledge of the initial state ε_0, which violates the assumption that $U_t^i = \{0\}\ \forall i$. Therefore, the equilibrium (C.1) is not an information equilibrium under the assumptions of Theorem 1. Given the above argument one might think that the same contradiction applies to the IE characterized in Theorem 1. This is not the case. One can in fact show that, under the supply process assumed above, the dynamic equation of the IE price characterized by Theorem 1 in terms of observables is now $p_t = -(1/\theta)p_{t-1} + (\theta+\beta)/(\theta)s_t + (1/\theta)s_{t-1}$. Notice that the autoregressive root in this equation implies stationarity, and so there is no need for it to cancel with the moving average root of s_t. The implication is that there is no need for a precise initial condition on p_{-1} to ensure stationarity, which means that $U_t^i = \{0\}\ \forall i$ does not need to be violated and the contradiction argument does not go through.

B.4 Proofs of Propositions for Section 5

B.4.1 Corollary 3

If we redefine the fundamentals process for the informed as

$$\begin{aligned} f_t^I = s_t - (1-\mu)\beta[p_{t+1} - E_t^U p_{t+1}] &= A(L)\varepsilon_t - (1-\mu)\beta[(L-\lambda)Q(L)L^{-1} - L^{-1}[(L-\lambda)Q(L) - Q_0\mathcal{B}_\lambda(L)] \\ &= A(L)\varepsilon_t - (1-\mu)\beta L^{-1}Q_0\mathcal{B}_\lambda(L) \end{aligned} \tag{B.16}$$

and solve the model

$$p_t = \beta E_t^I p_{t+1} + f_t^I \tag{B.17}$$

using the z-transform methodology described in the paper, then by guessing that $p_t = \pi(L)\varepsilon_t$, we have

$$\begin{aligned} \pi(z) &= \beta z^{-1}[\pi(z) - \pi_0] + A(z) - z^{-1}\mathcal{B}_\lambda(z)(1-\mu)\beta Q_0 \\ (z-\beta)\pi(z) &= zA(z) - \mathcal{B}_\lambda(z)(1-\mu)\beta Q_0 - \beta\pi_0 \end{aligned} \tag{B.18}$$

Evaluating the RHS at β, $A(\beta) - \mathcal{B}_\lambda(\beta)(1-\mu)Q_0 = \pi_0$.

The equilibrium is

$$\begin{aligned} p_t &= \frac{1}{L-\beta}\Big(LM^I(L) - \beta M^I(\beta)\Big) \\ &= \frac{1}{L-\beta}\Big(LA(L) - (1-\mu)\beta Q_0\mathcal{B}_\lambda(L) - \beta A(\beta) + (1-\mu)\beta Q_0\mathcal{B}_\lambda(\beta)\Big) \\ &= \frac{1}{L-\beta}\Big(LA(L) - \beta A(\beta) - (1-\mu)\beta Q_0(\mathcal{B}_\lambda(L) - \mathcal{B}_\lambda(\beta))\Big) \end{aligned} \tag{B.19}$$

Recall that $Q_0 = -A(\beta)/h(\beta)$, substituting this in to (B.19) and a bit of algebra delivers the equilibrium (3.8).

For the uninformed, we have a guess of $p_t = \tilde{\pi}(L)\tilde{\varepsilon}_t$, therefore the model is solved in $\tilde{\varepsilon}$ space

$$\begin{aligned} p_t - \beta E[p_{t+1}|p^t, E_t^I] &= f_t^U \\ \pi(L) = \beta L^{-1}[\pi(L) - \pi_0] &+ (1-\lambda L)\left(\frac{A(L) + L^{-1}\mu\beta\lambda Q_0}{L-\lambda}\right) \end{aligned} \tag{B.20}$$

The equilibrium is

$$\pi(L) = \frac{1}{L-\beta}\left(\frac{(1-\lambda L)(LA(L) + \mu\beta\lambda Q_0)}{L-\lambda} - \frac{(1-\lambda\beta)(\beta A(\beta) + \mu\beta\lambda Q_0)}{\beta-\lambda}\right)\mathcal{B}_\lambda(L)\varepsilon_t \tag{B.21}$$

Recall that $Q_0 = -A(\beta)/h(\beta)$, substituting this in to (B.21) and a bit of algebra delivers the equilibrium (3.8).

B.4.2 Proposition 2

Write the equilibrium price as $p_t = (L-\lambda)Q(L)\varepsilon_t$ where $|\lambda| < 1$ and $Q(L)$ satisfies (3.8). For $j = 1$, the time $t+1$ average expectation of the price at $t+2$ is given by

$$\begin{aligned} \overline{\mathbb{E}}_{t+1}p_{t+2} &= \mu\mathbb{E}_{t+1}^{\mathcal{I}}p_{t+2} + (1-\mu)\mathbb{E}_{t+1}^{\mathcal{U}}p_{t+2} \\ &= L^{-1}(L-\lambda)Q(L)\varepsilon_{t+1} + L^{-1}Q_0[\mu\lambda - (1-\mu)\mathcal{B}_\lambda(L)]\varepsilon_{t+1} \\ &= p_{t+2} + L^{-1}Q_0[\mu\lambda - (1-\mu)\mathcal{B}_\lambda(L)]\varepsilon_{t+1} \end{aligned} \tag{B.22}$$

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

$$\mathbb{E}_t^{\mathcal{I}}\overline{\mathbb{E}}_{t+1}p_{t+2} = \mathbb{E}_t^{\mathcal{I}}p_{t+2} + \mu\lambda Q_0\mathbb{E}_t^{\mathcal{I}}\varepsilon_{t+2} - Q_0(1-\mu)\mathbb{E}_t^{\mathcal{I}}\mathcal{B}_\lambda(L)\varepsilon_{t+2}. \tag{B.23}$$

Clearly $\mathbb{E}_t^{\mathcal{I}}\varepsilon_{t+2} = 0$, whereas the expectation in the last term of (B.23) is given by

$$\mathbb{E}_t^{\mathcal{I}}\mathcal{B}_\lambda(L)\varepsilon_{t+2} = L^{-2}\{\mathcal{B}_\lambda(L) - \mathcal{B}_\lambda(0) - \mathcal{B}_\lambda(1)L\}\varepsilon_t \tag{B.24}$$

where the notation $\mathcal{B}_\lambda(j)$ stands for "the sum of the coefficients of L^j ". If we write

$$\mathcal{B}_\lambda(L) = (L-\lambda)(1 + \lambda L + \lambda^2 L^2 + \lambda^3 L^3 + \cdots).$$

it is straightforward to show that $\mathcal{B}_\lambda(0) = -\lambda$ and $\mathcal{B}_\lambda(1) = (1-\lambda)(1+\lambda) = (1-\lambda^2)$, from which follows

$$\mathcal{B}_\lambda(L) - \mathcal{B}_\lambda(0) - \mathcal{B}_\lambda(1)L = \frac{L-\lambda}{1-\lambda L} + \lambda - (1-\lambda^2)L = \frac{\lambda(1-\lambda^2)L^2}{1-\lambda L}.$$

Putting things together, the informed agent's expectation of the average expectation is

$$\mathbb{E}_t^{I}\overline{\mathbb{E}}_{t+1}p_{t+2} = \mathbb{E}_t^{I}p_{t+2} - (1-\mu)Q_0\lambda\left(\frac{1-\lambda^2}{1-\lambda L}\right)\varepsilon_t \tag{B.25}$$

For the uninformed, we need to evaluate the following expectation

$$\mathbb{E}_t^{\mathcal{U}}p_{t+1} = (1-\mu)\beta\mathbb{E}_t^{\mathcal{U}}p_{t+2} + \mathbb{E}_t^{\mathcal{U}}[\mu\beta\mathbb{E}_{t+1}^{\mathcal{I}}p_{t+2} + s_{t+1}] \tag{B.26}$$

Writing out the term in brackets gives

$$\begin{aligned}
\mu\beta\mathbb{E}_{t+1}^{\mathcal{I}}p_{t+2} + s_{t+1} &= \mu\beta L^{-1}[(L-\lambda)Q(L) + \lambda Q_0]\varepsilon_{t+1} + A(L)\varepsilon_{t+1} \\
&= \mu\beta(L-\lambda)Q(L)\varepsilon_{t+2} + G(L)\varepsilon_{t+2} \\
&= \mu\beta p_{t+2} + G(L)\varepsilon_{t+2}
\end{aligned}$$

where $G(L)\varepsilon_{t+2} = [\mu\beta\lambda Q_0 + LA(L)]\varepsilon_{t+2}$. Note that the existence condition implies that $G(L)$ must vanish at $L = \lambda$. Therefore, we may rewrite $G(L)\varepsilon_{t+2}$ as $(L-\lambda)\hat{G}(L)\varepsilon_{t+2}$, where $\hat{G}(L)$ has no zeros inside the unit circle. This implies that $G_0 = -\hat{G}_0\lambda$, $G_i = \hat{G}_{i-1} - \lambda\hat{G}_i$, for $i = 1, \ldots$ and therefore $\hat{G}_0 = -\mu\beta Q_0$, $\hat{G}_1 = (\hat{G}_0 - G_1)/\lambda = -(\mu\beta Q_0 + A_0)/\lambda$

Evaluating (B.26) yields

$$\begin{aligned}
\mathbb{E}_t^{\mathcal{U}}p_{t+1} &= \beta\mathbb{E}_t^{\mathcal{U}}p_{t+2} + \mathbb{E}_t^{\mathcal{U}}(L-\lambda)\hat{G}(L)\varepsilon_{t+2} \\
&= \beta\mathbb{E}_t^{\mathcal{U}}p_{t+2} + L^{-2}[(L-\lambda)\hat{G}(L) - \{\hat{G}_0 + (\hat{G}_1 - \lambda\hat{G}_0)L\}\mathcal{B}_\lambda(L)]\varepsilon_t \\
&= \beta\mathbb{E}_t^{\mathcal{U}}p_{t+2} + s_{t+1} + \frac{A_0}{\lambda}\mathcal{B}_\lambda(L)\varepsilon_{t+1} + \left(\frac{\mu\beta Q_0(1-\lambda^2)}{\lambda(1-\lambda L)}\right)\varepsilon_t
\end{aligned} \tag{B.27}$$

Now we define $\mathbb{E}_t^{\bar{\mathcal{U}}}p_{t+1}$ as (B.27) without the last term, so that

$$\mathbb{E}_t^{\bar{\mathcal{U}}}p_{t+1} = \beta\mathbb{E}_t^{\mathcal{U}}p_{t+2} + s_{t+1} - \frac{A_0}{\lambda}\mathcal{B}_\lambda(L)\varepsilon_{t+1} \tag{B.28}$$

(B.28) would hold if the uninformed agents ignored the information coming from the informed agent's forecast errors. Therefore the difference between (B.27) and (B.28) must be due to HOBs. This difference is given by the last term in (B.27). Using the above definitions in conjunction with (B.26) equation (5.7) follows.

B.4.3 Proposition 3, No HOBs Equilibrium If we were to assume that informed agents acted irrationally and ignored information coming from the model, then the informed would not form HOBs and their expectations would satisfy the law of iterated expectations,

$$E_t^I(p_{t+1}) = \beta E_t^I(p_{t+2}) + E_t^I s_{t+1} \tag{B.29}$$

That is, the higher-order beliefs component $(1-\mu)(1-\lambda^2)\left(\frac{Q_0\lambda}{1-\lambda L}\right)\varepsilon_t$ (which was derived in the proof of Proposition 2) is removed from the informed agents' expectation.

Assuming $|\lambda| < 1$ and $p_t = (L-\lambda)Q(L)\varepsilon_t$, then

$$E_t^I(p_{t+1}) = \beta L^{-2}[(L-\lambda)Q(L) + \lambda Q_0 - LQ_0 + L\lambda Q_1]\varepsilon_t + L^{-1}[A(L) - A_0]\varepsilon_t$$

and equilibrium in z-transforms can be written as

$$(z-\lambda)(z-\beta)(z+\mu\beta)Q(z) = z(z+\mu\beta)A(z) - z\mu\beta A_0 + \beta G(z)Q_0 + \mu\beta^2\lambda z Q_1 \tag{B.30}$$

where $G(z) = \mu\beta(\lambda - z) - (1-\mu)z\mathcal{B}_\lambda(z)$. To remove the pole at $z = -\mu\beta$, Q_1 must satisfy

$$(\mu\beta)^2 A_0 + \beta G(-\mu\beta)Q_0 - \mu^2\beta^3\lambda Q_1 = 0$$

substituting this into (B.30) gives

$$(z-\lambda)(z-\beta)Q(z) = \left\{zA(z) + \frac{\beta}{1+\mu\lambda\beta}Q_0 g(z)\right\}$$

where $g(z) = \mu\lambda(1+\lambda\beta) - (1-\mu)\mathcal{B}_\lambda(z)$. Removing the poles at λ and β implies the restrictions

$$A(\lambda) + \frac{\beta Q_0\mu(1+\lambda\beta)}{1+\mu\lambda\beta} = 0, \qquad Q_0 = \frac{-A(\beta)(1+\mu\lambda\beta)}{g(\beta)}$$

This delivers the equilibrium conditions

$$p_t = \frac{1}{L-\beta}\left(LA(L) - \beta A(\beta)\frac{g(L)}{g(\beta)}\right)\varepsilon_t \tag{B.31}$$

and $A(\cdot)$ must satisfy

$$A(\lambda) = \frac{\beta\mu A(\beta)(1+\lambda\beta)}{g(\beta)} \tag{B.32}$$

The equilibrium conditions (B.31) and (B.32) is the boundedly rational equilibrium assuming *first*-order higher order beliefs are removed. To remove the first- and second-order higher-order beliefs requires the informed agents' expectation to be given by

$$\begin{aligned} E_t^I p_{t+1} &= \beta^2 E_t^I(p_{t+3}) + E_t^I s_{t+1} + \beta E_t^I(s_{t+2}) \\ &= \beta^2 L^{-3}[(L-\lambda)Q(L) + \lambda Q_0 - (Q_0 - \lambda Q_1)L - (Q_1 - \lambda Q_2)L^2]\varepsilon_t \\ &+ L^{-1}[A(L) - A_0]\varepsilon_t + \beta L^{-2}[A(L) - A_0 - A_1 L]\varepsilon_t \end{aligned} \tag{B.33}$$

This assumes the law of iterated expectations applies to time $t+1$ and $t+2$ for the informed agents.

Substituting this expression into equilibrium yields

$$\begin{aligned} (z-\lambda)Q(z) = \beta\mu\{&\beta^2 z^{-3}[(z-\lambda)Q(z) + \lambda Q_0 - (Q_0 - \lambda Q_1)z - (Q_1 - \lambda Q_2)z^2] \\ &+ z^{-1}[A(z) - A_0] + \beta z^{-2}[A(z) - A_0 - A_1 z]\} \\ &+ \beta(1-\mu)z^{-1}[(z-\lambda)Q(z) - Q_0\mathcal{B}_\lambda(z)] + A(z) \end{aligned}$$

Some tedious algebra delivers

$$\begin{aligned}(z-\lambda)(z-\beta)(z^2+\mu\beta z+\mu\beta^2)Q(z) &= \mu\beta^3[\lambda Q_0-(Q_0-\lambda Q_1)z-(Q_1-\lambda Q_2)z^2]\\ &+\mu\beta z^2[A(z)-A_0]+\mu\beta^2 z[A(z)-A_0-A_1 z]\\ &-\beta(1-\mu)z^2[Q_0\mathcal{B}_\lambda(z)]+z^3A(z)\\ &=(z^2+\mu\beta z+\mu\beta^2)zA(z)+\beta J(z)Q_0+\mu\beta^3 z(\lambda-z)Q_1\\ &+\mu\beta^3\lambda z^2Q_2-z(\mu\beta^2-\mu\beta z)A_0-\mu\beta^2z^2A_1\end{aligned}$$

where $J(z)=\mu\beta^2(\lambda-z)-z^2(1-\mu)\mathcal{B}_\lambda(z)$.

Term hitting $Q(z)$ does not factor but it is easy to show that both zeros are inside unit circle. Write the zeros as $(z^2+\mu\beta z+\mu\beta^2)=(z-\xi_1)(z-\xi_2)$.

$$\begin{aligned}(z-\lambda)(z-\beta)(z-\xi_1)(z-\xi_2)Q(z) &= (z-\xi_1)(z-\xi_2)zA(z)+\beta J(z)Q_0+\mu\beta^3z(\lambda-z)Q_1\\ &+\mu\beta^3\lambda z^2Q_2-z(\mu\beta^2-\mu\beta z)A_0-\mu\beta^2z^2A_1\end{aligned}$$

Using Q_2 and Q_1 to remove $z=\{\xi_1,\xi_2\}$ gives two restrictions and two unknowns.

$$\beta J(\xi_1)Q_0+\mu\beta^3\xi_1(\lambda-\xi_1)Q_1+\mu\beta^3\lambda\xi_1^2Q_2-\xi_1(\mu\beta^2-\mu\beta\xi_1)A_0-\mu\beta^2\xi_1^2A_1=0$$
$$\beta J(\xi_2)Q_0+\mu\beta^3\xi_2(\lambda-\xi_2)Q_1+\mu\beta^3\lambda\xi_2^2Q_2-\xi_2(\mu\beta^2-\mu\beta\xi_2)A_0-\mu\beta^2\xi_2^2A_1=0$$

Substituting in these values, dividing by $(z-\xi_1)(z-\xi_2)$ and tedious algebra delivers

$$(z-\lambda)(z-\beta)Q(z)=zA(z)+\frac{\beta}{\kappa}Q_0 j(z) \tag{B.34}$$

where $j(z)=\mu\lambda(1+\lambda\beta+(\lambda\beta)^2)-(1-\mu)\mathcal{B}_\lambda(z)$, and κ is a complicated constant of $\xi_1,\xi_2,\lambda,\beta$ and μ. To remove the pole at $z=\lambda$, $A(\cdot)$ must satisfy$A(\lambda)+\frac{\beta Q_0\mu(1+\lambda\beta+(\lambda\beta)^2)}{\kappa}=0$. To remove the pole at $z=\beta$, Q_0 must satisfy $Q_0=\frac{-A(\beta)\kappa}{j(\beta)}$.

Substituting in Q_0 delivers the result

$$p_t=\frac{1}{L-\beta}\left(LA(L)-\beta A(\beta)\frac{j(L)}{j(\beta)}\right)\varepsilon_t \tag{B.35}$$

where

$$j(L)=\mu\lambda(1+\lambda\beta+(\lambda\beta)^2)-(1-\mu)\mathcal{B}_\lambda(L) \tag{B.36}$$

and $A(\cdot)$ must satisfy

$$A(\lambda)=\frac{\beta\mu A(\beta)(1+\lambda\beta+(\lambda\beta)^2)}{j(\beta)} \tag{B.37}$$

By induction, we are converging to

$$p_t=\frac{1}{L-\beta}\left(LA(L)-\beta A(\beta)\frac{k(L)}{k(\beta)}\right)\varepsilon_t \tag{B.38}$$

where

$$k(L)=\mu\lambda(\sum_{j=0}^{n}(\lambda\beta)^j)-(1-\mu)\mathcal{B}_\lambda(L) \tag{B.39}$$

and $A(\cdot)$ must satisfy

$$A(\lambda) = \frac{\beta\mu A(\beta)(\sum_{j=0}^{n}(\lambda\beta)^j)}{k(\beta)} \tag{B.40}$$

Letting $n \to \infty$ delivers the desired result.

Removing Both HOBs Corollary 2 shows that removing both the informed and uninformed HOBs will lead to an equilibrium in which each agent only forecasts the sum of future s_t's. The boundedly-rational equilibrium in this setup will therefore be a convex combination of the fully informed equilibrium given by (3.3) and the fully uninformed equilibrium of Theorem 1.

B.4.4 Corollary 4 Given the ARMA(1,1) specification, the roots determining λ for the information equilibrium are given by the following quadratic,

$$f(\lambda) = \beta\mu\theta(1-\rho\beta)\lambda^2 + [\beta\mu(1+\theta\beta) - \theta(1-\rho\beta)]\lambda - (1+\beta\mu\theta) + \rho\beta(1-\mu) = 0 \tag{B.41}$$

If we remove the higher-order beliefs of the informed, the

$$g(\bar{\lambda}) = \theta(1-\rho\beta)\bar{\lambda} + 1 + \beta\mu\theta - \rho\beta(1-\mu) = 0 \tag{B.42}$$

which gives

$$\bar{\lambda} = \frac{-(1+\mu\theta\beta) + \rho\beta(1-\mu)}{\theta(1-\rho\beta)} \tag{B.43}$$

Removing both the informed and uniformed's HOBs gives

$$h(\lambda^*) = \theta(1-\rho\beta)\lambda^{*2} + [1 - \rho\beta + \beta\rho\mu(1+\theta\beta)]\lambda^* - \mu\beta(1+\theta\beta) = 0 \tag{B.44}$$

The proof consists of two parts:

We will first show that $|\lambda| > |\bar{\lambda}|$ for all $\lambda \in (-1, 1)$.

From Result IE (figure 3), an IE with $|\lambda| < 1$ requires $\theta > 1$. Notice also that $\theta > 0$ implies the quadratic (B.41) is convex and $f(\lambda)\big|_{\lambda=0} = \rho\beta - 1 - \beta\mu(\theta + \rho) < 0$. To prove the result we show that evaluating (B.41) at the root of (B.42) delivers a negative value. Evaluating (B.41) at $\bar{\lambda}$ yields

$$\frac{\beta^2\mu[1 - \rho\beta + \mu\beta(\rho+\theta)](\rho+\theta)(\mu-1)}{\theta(1-\rho\beta)} < 0 \tag{B.45}$$

which proves $|\lambda| > |\bar{\lambda}|$.

We now prove that $|\bar{\lambda}| > |\lambda^*|$. Removing all HOBs could yield an equilibrium with two roots inside the unit circle. The product of the two roots of (B.44) is $-\mu\beta(1+\theta\beta)/(\theta(1-\rho\beta))$, which is always less than (B.43) in absolute value when

$$\beta\{\mu[1 - \theta(1-\beta)] + (1-\mu)\rho\} < 1 \tag{B.46}$$

which holds given the restrictions on the parameter values.

B.4.5 Proposition 4 The notation of the proof is that of Theorem 1 unless otherwise specified. We begin by noticing that

$$\mathbb{E}_{it}\overline{\overline{\mathbb{E}}}_{t+1}p_{t+2} = \mu\mathbb{E}_{it}\mathbb{E}^{\mathcal{I}}_{t+1}p_{t+2} + (1-\mu)\mathbb{E}_{it}\mathbb{E}^{\mathcal{U}}_{t+1}p_{t+2}. \tag{B.47}$$

From the hierarchical equilibrium case we know that $\mathbb{E}^{\mathcal{U}}_{t+1}p_{t+2} = \mathbb{E}^{\mathcal{I}}_{t+1}p_{t+2} - Q_0\frac{1-\lambda^2}{1-\lambda L}\varepsilon_{t+1}$. We also notice that, because the information set of an arbitrary agent i is strictly smaller than the information set of an informed agent of the hierarchical equilibrium and because the law of iterated expectations holds at the single agent level, we have $\mathbb{E}_{it}\mathbb{E}_{it+1}\mathbb{E}^{\mathcal{I}}_{t+1}p_{t+2} = \mathbb{E}_{it}p_{t+2}$. Because of the second property we also have that $\mathbb{E}_{it}\mathbb{E}^{\mathcal{U}}_{t+1}p_{t+2} = \mathbb{E}_{it}\mathbb{E}_{it+1}\mathbb{E}^{\mathcal{U}}_{t+1}p_{t+2}$. Therefore

$$\mathbb{E}_{it}\overline{\mathbb{E}}_{t+1}p_{t+2} = \mu\mathbb{E}_{it}p_{t+2} + (1-\mu)\mathbb{E}_{it}p_{t+2} - (1-\mu)Q_0\mathbb{E}_{it}\frac{1-\lambda^2}{1-\lambda L}\varepsilon_{t+1}. \tag{B.48}$$

The crucial step in the proof is then to show that the expectation in the last term is non-zero. In order to do so we first notice that $\frac{L-\lambda}{1-\lambda L}\varepsilon_{t+2} = \frac{1-\lambda^2}{1-\lambda L}\varepsilon_{t+1} - \lambda\varepsilon_{t+2}$ and so

$$\mathbb{E}\left(\frac{1-\lambda^2}{1-\lambda L}\varepsilon_{t+1}|\varepsilon_i^t, p^t\right) = \mathbb{E}\left(\frac{L-\lambda}{1-\lambda L}\varepsilon_{t+2}|\varepsilon_i^t, p^t\right). \tag{B.49}$$

Then, the crucial step in the proof is to show that

$$\mathbb{E}\left(\frac{L-\lambda}{1-\lambda L}\varepsilon_{t+2}|\varepsilon_i^t, p^t\right) = \mu\lambda\frac{(1-\lambda^2)}{1-\lambda L}\varepsilon_{it}. \tag{B.50}$$

where $\mu \equiv \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2+\sigma_v^2}$. Remember that we defined

$$\tilde{\varepsilon}_t = \mathcal{B}(L)\varepsilon_t. \tag{B.51}$$

To ease notation, let $\tilde{\varepsilon} = y$, then we look for $\mathbb{E}\left(y_{t+2}|\varepsilon_i^t, p^t\right) = \pi_1(L)\varepsilon_{it} + \pi_2(L)p_t$. From Theorem 1 in Rondina (2009) we know that

$$\begin{bmatrix} \pi_1(L) & \pi_2(L) \end{bmatrix} = \left[L^{-2}g_{y,(\varepsilon,p)}(L)\left(\Gamma^*(L^{-1})^T\right)^{-1}\right]_+ \Gamma^*(L)^{-1} \tag{B.52}$$

where $\Gamma^*(L)$ and $\left(w_{it}^1, w_{it}^2\right)$ are defined in (A.10) 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

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

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

$$\left[L^{-2}\frac{L-\lambda}{1-\lambda L}\right]_+ = \frac{\lambda\left(1-\lambda^2\right)}{1-\lambda L}, \tag{B.53}$$

Summarizing we have shown that

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

Notice that

$$\Gamma^*(L)^{-1}\begin{bmatrix} \varepsilon_{it} \\ p_t \end{bmatrix} = \begin{bmatrix} w_{it}^1 \\ w_{it}^2 \end{bmatrix}$$

so that

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

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 of statement (i). The proof can be generalized to expectations of order higher than 1. For statement (ii) the proof follows exactly the proof of Proposition 2 since it concerns only aggregate variables, which we know from the proof of Theorem 1 follow the same patter as those of the hierarchical case.