

DYNAMIC HIGHER ORDER EXPECTATIONS

KRISTOFFER P. NIMARK

ABSTRACT. In models where privately informed agents interact, agents may need to form higher order expectations, i.e. expectations of other agents' expectations. This paper explores dynamic higher order expectations in a setting where it is common knowledge that agents are rational Bayesians. This structure is a natural generalization of full information rational expectations and allows any order of expectation at any horizon to be determined recursively. The usefulness of the approach is illustrated by solving a version of Singleton's (1987) asset pricing model with disparately informed traders but without assuming that shocks can be observed perfectly with a lag. In the context of Singleton's asset pricing model, we prove that both the impact of expectations on the asset's price and the variance of expectations are decreasing as the order of expectation increases. We use these results to derive a finite dimensional state representation that can be made arbitrarily accurate. The solution method exploits the Euler-type structure of the asset pricing function and should be applicable to a variety of settings where privately informed agents optimize intertemporally.

Keywords: Dynamic Higher Order Expectations, Private Information, Asset Pricing

1. INTRODUCTION

Most economic models involve some type of interaction between multiple agents where the payoff of one agent depends not only on the actions taken by him, but also on the actions taken by other agents. When agents' preferences and environment are identical and all share the same information, an individual agent can infer the actions that others will take by introspection, since all agents will choose the same action in equilibrium. If agents have access to different information, this is no longer possible since individual agents cannot know with certainty what other agents know and therefore also not know with certainty what actions they will take. It then becomes necessary for agents to form expectations about the actions of others. Additionally, to predict the behavior of agents that form expectations about the actions of others, one need to form expectations about other agents' expectations about the actions of others, and so on, leading to the well-known infinite regress of expectations.¹ The idea that agents observe different pieces of information has a lot of appeal and has been applied to a variety of settings, including general equilibrium models of the business cycle

Date: October 2007. The author thanks Thomas Sargent and seminar participants at New York University, Birkbeck College, Goethe University Frankfurt, the 2007 annual meeting of the Society for Economic Dynamics in Prague and the Workshop for Macro Dynamics in Adelaide for useful comments and suggestions.

Address: RESEARCH DEPARTMENT, RESERVE BANK OF AUSTRALIA, 65 MARTIN PLACE, N.S.W. 2001, SYDNEY, AUSTRALIA. E-MAIL: NIMARKK@RBA.GOV.AU
The views expressed in this paper are those of the author and not necessarily those of the Reserve Bank of Australia.

¹Townsend (1983) and Sargent (1991).

and asset pricing models.² However, as a consequence of the infinite regress problem one could characterize most existing models of private information and strategic interaction as efforts to avoid modelling higher order expectations explicitly, and instead find alternative representations where higher order expectations do not occur as state variables. Notable exceptions are Woodford (2002), Morris and Shin (2002) and Adam (forthcoming) who by restricting their attention to models of static decisions are able to analyze higher order expectations explicitly. This paper explores the properties of higher order expectations in a setting where agents make *dynamic choices* and shows how such models can be solved in a linear setting with an explicit role for higher order expectations.

Throughout this paper, it is assumed that agents are rational Bayesians and that this is common knowledge. ‘Rational’ means that agents do not make systematic mistakes given their information sets. ‘Bayesian’ means that agents update their expectations when new information arrives. That this is common knowledge means that all agents know that all agents know, and so on, that all agents are rational Bayesians. Common knowledge of rationality gives enough structure to expectations to allow any order of expectation at any horizon to be determined recursively.³

We also use the structure imposed on expectations by common knowledge of rationality to define an *average one-period-ahead expectations operator*. In the context of Singleton’s (1987) asset pricing model with disparately but symmetrically informed investors, the operator is used to compute the equilibrium price of the asset. The usefulness of the operator comes from the fact that it allows us to iterate Euler equations with ‘average expectations’ terms forward in time without relying on the law of iterated expectations. The solved price equation then resembles a discounted sum of expected future fundamentals.

Deriving the dynamics of higher order expectations does not solve the problem of how to model the infinite regress of expectations in practise. Again in the context of Singleton’s (1987) model, we prove two important results towards this end. First, we show analytically that the impact of expectations on the price of the asset decreases as the order of expectation increases.⁴ This result holds under the same conditions that guarantee that a solution exists when agents are perfectly informed, i.e. that the eigenvalue of the fundamental process multiplied by the discount rate is smaller than unity in absolute value. Second, we prove that the variance of expectations decreases as the order of expectation increases. The second result is a direct implication of common knowledge of rational expectations and should apply to linear models with private information more generally. Taken together, the two results constitute the main contribution of the paper: The dynamics of an infinite horizon model with private information and with agents that optimize intertemporally can be approximated to an arbitrary accuracy by a finite state representation. This result is derived without relying on the common strategy of making additional assumptions to ensure that private information is short lived.

²Some examples are Townsend (1983), Sargent (1991), Woodford (2002), Lorenzoni (2005), Bacchetta and Van Wincoop (2005), Kasa, Walker and Whiteman (2006) and Cespa and Vives (2007).

³A similar assumption is used implicitly in the static decision models of both Woodford (2002) and Morris and Shin (2002) to construct higher order expectations of the current state of the economy.

⁴A result with similar implications for a strategic one period game without endogenous signals can be found in Weinstein and Yildiz (forthcoming).

One way to make private information short-lived is to impose that all shocks are observed perfectly by all agents with a lag. This assumption was first introduced by Townsend (1983) as a way to restrict the dimension of the relevant state for ‘forecasting the forecasts of others’. Before Townsend, Lucas (1975) assumed that the inhabitants of his island economy pooled their information between periods in order to circumvent the infinite regress problem. More recently, Bacchetta and van Wincoop (2006) used a similar assumption to Townsend, to analyze exchange rate dynamics. In their model, fundamentals are perfectly observed contemporaneously, but individual investors receive a private signal of future fundamentals. Together with projection techniques, short lived information makes it possible to solve dynamic models with private information, but may result in kinks in the impulse response functions at the lag when shocks become common knowledge. In many settings it is also arguably unrealistic to assume that the shocks can ever be observed, not even with very long lags.

Other strategies to restrict the state dimension can be found in for instance Allen, Morris and Shin (2006) who set up a finite horizon model to investigate the effects of private information on the price of an asset with a terminal liquidation date. Cespa and Vives (2007) introduce long-term traders in a model that resembles that of Allen, Morris and Shin (2006). Lorenzoni (2006) presents a dynamic general equilibrium model where agents are subject to idiosyncratic productivity shocks and need to infer aggregate productivity to optimally choose consumption. The model is used to explain the origin of demand shocks. Lorenzoni assumes that an observation far enough in the past is uninformative and truncates the state space in the time domain.

A different approach to solve a dynamic model with private information, that does not rely on restricting the dimension of the state, is taken by Kasa, Walker and Whiteman (2006). They present a simple asset price model with risk neutral traders where the fundamentals of the asset are driven by two mutually orthogonal stochastic processes. Traders are divided into two ‘types’ depending on which of the stochastic processes that they can observe. Both types observe the equilibrium price. Kasa *et al* then derive conditions for when the observation of the equilibrium price does or does not reveal the information held by the other type of trader. They also show how a solution can be found analytically, which is made possible by conducting the analysis entirely in the frequency domain. It is not clear whether their approach can be generalised to a setting with a large number of traders (or types) or where traders receive information about the same underlying process, i.e. a setting with non-orthogonal private signals, but it does offer some analytical elegance.

It was the paper by Townsend (1983) that coined the popular term ‘forecasting the forecasts of others’ to describe the infinite regress problem discussed above. In an ironic twist to the history of the topic, subsequent research, i.e. Kasa (2000) and Pearlman and Sargent (2005), have showed that in the model studied by Townsend, private information is not preserved when agents observe equilibrium prices so there is actually no need for agents to ‘forecast the forecasts of others’. Walker (forthcoming) claims that this also applies to the model of Singleton studied here. However, we demonstrate below that this result is due to additional (and special) assumptions made by Walker that are not part of Singleton’s original set up and that in general, private information is preserved in the Singleton set up.

The next section defines the concept of dynamic higher order expectations and fixes notation. This is followed in Section 3 by what is hopefully an intuitive (and in any case a detailed) account of how common knowledge of rationality can be used to construct a law of motion for higher order expectations. Section 3 also introduces the average one-period-ahead expectations operator mentioned above. Section 4 presents Singleton's model that is used as a vehicle for the rest of the analysis. Section 5 contains the main results of the paper. There, we demonstrate how common knowledge of rationality can be used to construct the law of motion for higher order expectations when agents observe endogenous signals and how the law of motion can be used to compute the price of the asset in Singleton's model. In Section 5 we also prove the two results that allow us to work with a finite dimensional state without assuming that shocks can be observed with a lag. Section 6 examines the relationship between model parameters and the importance of higher order expectations and Section 7 discusses how observing equilibrium prices reveal information to individual traders and under what circumstances the underlying state will be perfectly revealed. Section 8 concludes.

2. CONCEPTS AND NOTATION

Before analyzing the dynamics of higher order expectations, it is necessary to invest some time in a notational machinery as well as to define exactly what is meant by a dynamic higher order expectation.

2.1. Contemporaneous expectations. We start by defining *contemporaneous higher order expectations*, i.e. higher order expectations held today of the current value of a random variable. There is a continuum of agents indexed by $j \in (0, 1)$. Agent j 's first order expectation of the variable θ_t is agent j 's best estimate of the value of the variable given his information set $I_t(j)$. We denote agent j 's first order expectation of θ_t at time t

$$\theta_{t|t}^{(1)}(j) \equiv E[\theta_t | I_t(j)] \quad (2.1)$$

The average first order expectation is obtained by taking averages of (2.1) across agents

$$\theta_{t|t}^{(1)} \equiv \int E[\theta_t | I_t(j)] \, dj \quad (2.2)$$

The average second order expectation is obtained by taking the average of agents' expectations of (2.2)

$$\theta_{t|t}^{(2)} \equiv \int E[\theta_{t|t}^{(1)} | I_t(j)] \, dj \quad (2.3)$$

The average contemporaneous second order expectation of θ_t thus is the average expectation at time t of the average expectation at time t of the value of θ_t . We can generalize this notation to the k^{th} order expectation of θ_t

$$\theta_{t|t}^{(k)} \equiv \int E[\theta_{t|t}^{(k-1)} | I_t(j)] \, dj \quad (2.4)$$

Define the zero order expectation of θ_t as the actual value of the variable

$$\theta_t^{(0)} \equiv \theta_t \quad (2.5)$$

In general

$$\theta_{t|t}^{(k)} \neq \theta_{t|t}^{(k+l)} \quad (2.6)$$

for $l \neq 0$. We call a sequence of contemporaneous expectations, for instance from order zero to k , a *hierarchy* of expectations from order zero to k . Vectors consisting of a hierarchy of expectations are denoted

$$\theta_{t|t}^{(0:k)} = \left[\theta_{t|t}^{(0)} \quad \theta_{t|t}^{(1)} \quad \dots \quad \theta_{t|t}^{(k)} \right]' \quad (2.7)$$

2.2. Dynamic higher order expectations. In this section we define and set notation for two types of dynamic higher order expectations that differ in how easily they can be reduced to a function of the contemporaneous expectation hierarchy of the same variable. This difference in turn depends on the fact that the law of iterated expectations applies to only one of them. The accompanying notation should make clear (i) the order of expectation and (ii) on what time period information each order of expectation is conditioned on.

The law of iterated expectations can loosely be attributed to that agents do not believe that they have ‘incorrect’ expectations so that they do not expect to receive information in the future that will make them want to revise their own expectations in a particular direction. The same is not true about expectations about other’s expectations. For instance, an investor may believe that the average ‘market expectation’ of the fundamental value of an asset is incorrect, but as more information becomes available to others over time the ‘market expectation’ will be revised towards what the investor believes is the asset’s true value. It is the fact that it can be rational to expect others to revise their expectations in a certain direction that makes the law of iterated expectations inapplicable to higher order expectations.

2.2.1. Higher order expectations conditional on current information only. The simplest form of dynamic higher order expectations are expectations of what others expect *today* that the value of a variable will be s periods ahead, i.e higher order expectations based only on the information that others are believed to have access to today. The law of iterated expectations can then be applied directly to the current contemporaneous expectation of any order, since agents do not believe that other agents expect to revise their own expectations.

Let the one-period-ahead expectation of the random variable θ_t conditional on the current value of θ_t be

$$E[\theta_{t+1} | \theta_t] = f(\theta_t) \quad (2.8)$$

so that the expectation of horizon s conditional on θ_t is

$$E[\theta_{t+s} | \theta_t] = f^s(\theta_t). \quad (2.9)$$

That agents do not expect to revise their own expectations as time passes implies that we can apply the transition function f to the contemporaneous expectation of a given order to find the dynamic expectation of the same order. The expectation of order k at horizon s of this ‘simple’ type of dynamic higher order expectation is thus given by

$$f^s(\theta_{t|t}^{(k)}) \quad (2.10)$$

2.2.2. *Higher order expectations conditional on expected future information sets.* The static notation of contemporaneous expectations together with the transition function f for the random variable provide sufficient notation for the simple type of dynamic higher order expectations that are conditioned only on the expected current information sets of other agents. However, this is not the form of dynamic higher order expectations that arises naturally in economic applications. An example of a more interesting form of dynamic higher order expectations are expectations in period t of average expectations in period $t + 1$ of the average expectation of a random variable in period $t + 2$ when other agents are expected to receive new information in period $t + 1$. This latter type of dynamic expectations are relevant for optimal decision making when privately informed agents trade with each other sequentially and each agent need to forecast what others will be willing to pay for an asset or a good at some future date when other agents may have received new information and updated their expectations. We thus need a notation that indicates on what time period information sets the expectations at each order are conditioned on. Denote the average expectation in period t of the average expectation at $t + 1$ of θ_{t+2}

$$\theta_{t+2|t+1|t}^{(2)} \equiv \int E \left[\int E[\theta_{t+2} | I_{t+1}(j)] dj | I_t(j) \right] dj \quad (2.11)$$

More generally, denote the k order expectation of θ at horizon k

$$\begin{aligned} \theta_{t+k|\dots|t}^{(k)} &\equiv \int E \left[\int E[\dots \int E[\theta_{t+k} | I_{t+k-1}(j)] dj | I_{t+k-2}(j) dj] \right. \\ &\quad \left. \dots | I_{t+1}(j) dj | I_t(j) \right] dj \end{aligned} \quad (2.12)$$

The superscript on the term on the left hand side of (2.12) denotes the order of the expectation, i.e. the number of times averages are taken of expectations (and equals the number of integral signs in the expression on the right hand side). The subscripts denotes on what time period information the different orders of expectations are conditioned on. In words, (2.12) thus says “the average expectation in period t conditional on the information available to agents in period t of the average expectation in period $t + 1$ conditional on the information available to agents at $t + 1$ of the average expectation at period $t + 2$of the average expectation in period $t + k - 1$ conditional on the information available to agents in period $t + k - 1$ of the value of θ_{t+k} ”. This will in general differ from the k th order expectation of what the variable θ_t will be k periods ahead conditional only on the information other agents are believed to have access to today. I.e.

$$\theta_{t+k|\dots|t}^{(k)} \neq f^k(\theta_{t|t}^{(k)}) \quad (2.13)$$

3. THE ARITHMETIC OF DYNAMIC HIGHER ORDER EXPECTATIONS

This section gives a detailed account of the dynamics of higher order expectations. We use a simple set up with a continuum of agents that are estimating an unobservable process. Agents are rational Bayesians in the sense that they form optimal estimates of the process given their information sets and update their estimates when new information arrives by applying Bayes’ law through the Kalman filter. This is common knowledge, i.e. all agents know that all agents know, and so on, that all agents are rational Bayesians. In this section we

provide a more formal definition of this assumption and we demonstrate that it gives enough structure to expectations to pin down all orders of expectations at any horizon for given initial conditions. We also introduce an *average one-period-ahead expectations operator* that maps one agent's expectation of the current hierarchy of expectations into what he expects the average expectation of the same hierarchy will be in the next period. In this section, agents' sole concern is to estimate an unobservable process and the average of other agents' estimate of the same process. This set up is transparent (but void of any economic meaning) and the hope is that it will provide some intuition for the economically more interesting sections that follows.

3.1. Estimating a persistent process. Agents are indexed by $j \in (0, 1)$ and estimate the unobservable persistent process

$$\theta_t = \rho\theta_{t-1} + v_t \quad (3.1)$$

$$v_t \sim N(0, \sigma_v^2)$$

In period t agent j observes the unbiased but noisy signal $s_t(j)$ of the true value of θ_t

$$s_t(j) = \theta_t + \eta_t(j) \quad (3.2)$$

$$\eta_t(j) \sim N(0, \sigma_\eta^2) \quad \forall j \quad (3.3)$$

Equations (3.1) and (3.2) form a state space system that can be estimated using the Kalman filter. Agent j 's optimal estimate of θ_t in period t is given by the updating equation

$$\theta_{t|t}^{(1)}(j) = (1 - g_1)\rho\theta_{t-1|t-1}^{(1)}(j) + g_1s_t(j) \quad (3.4)$$

$$g_1 = \frac{p}{p + \sigma_\eta^2} < 1 \quad (3.5)$$

$$p = \sigma_v^2 + p\rho^2 - \frac{(p\rho)^2}{p + \sigma_e^2} \quad (3.6)$$

The interpretation of (3.4) is the following. The current estimate $\theta_{t|t}^{(1)}$ is a weighted average of the prior $\rho\theta_{t-1|t-1}$ and the observation $s_t(j)$. Intuitively, less weight is put on a noisy observation so the Kalman gain g_1 is decreasing in the variance of the noise term η_t . The limit cases $\sigma_\eta^2 = 0$ and $\sigma_\eta^2 = 1$ implies that $g_1 = 1$ and $g_1 = 0$ respectively so that $\sigma_\eta^2 = 0$ implies that the signal is a perfect indicator of the underlying variable since $s_t(j) = \theta_t \quad \forall t, j$.

3.2. Higher order estimates. In order to derive a law of motion for higher order estimates, take averages of the updating equation (3.4) across agents and combine it with the actual process (3.1) to get

$$\begin{bmatrix} \theta_t \\ \theta_{t|t}^{(1)} \end{bmatrix} = \begin{bmatrix} \rho & 0 \\ g_1\rho & (1 - g_1)\rho \end{bmatrix} \begin{bmatrix} \theta_{t-1} \\ \theta_{t-1|t-1}^{(1)} \end{bmatrix} + \begin{bmatrix} 1 \\ g_1 \end{bmatrix} v_t \quad (3.7)$$

The system (3.7) is the joint law of motion for the actual and average first order expectation of θ_t . Higher order estimates can be added recursively by recognizing that individual agents

can form an estimate of the system (3.7) by using the Kalman filter. The relevant state space system is then the transition (state) equation (3.7) and the measurement equation (3.8)

$$s_t(j) = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} \theta_t \\ \theta_{t|t}^{(1)} \end{bmatrix} + \eta_t(j) \quad (3.8)$$

which is just a reformulation of (3.2). Applying standard formulas for a multivariate Kalman filter yields the updating equation for agent j 's estimate of the system (3.7)

$$\begin{bmatrix} \theta_{t|t}^{(1)}(j) \\ \theta_{t|t}^{(2)}(j) \end{bmatrix} = \left(I - \begin{bmatrix} g_1 \\ g_2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix}' \right) \begin{bmatrix} \rho & 0 \\ g_1\rho & (1-g_1)\rho \end{bmatrix} \begin{bmatrix} \theta_{t-1|t-1}^{(1)}(j) \\ \theta_{t-1|t-1}^{(2)}(j) \end{bmatrix} \\ + \begin{bmatrix} g_1 \\ g_2 \end{bmatrix} \left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}' \begin{bmatrix} \theta_t \\ \theta_{t|t}^{(1)} \end{bmatrix} + \eta_t(j) \right) \quad (3.9)$$

where g_1 and g_2 are the elements of the Kalman gain matrix. Taking averages of (3.9) across agents and amending it to (3.1) yields the law of motion for estimates of θ_t from order zero to two.

$$\begin{bmatrix} \theta_t^{(0)} \\ \theta_{t|t}^{(1)} \\ \theta_{t|t}^{(2)} \end{bmatrix} = \begin{bmatrix} \rho & 0 & 0 \\ g_1\rho & (1-g_1)\rho & 0 \\ g_2\rho & (g_1-g_2)\rho & (1-g_1)\rho \end{bmatrix} \begin{bmatrix} \theta_{t-1}^{(0)} \\ \theta_{t-1|t-1}^{(1)} \\ \theta_{t-1|t-1}^{(2)} \end{bmatrix} + \begin{bmatrix} 1 \\ g_1 \\ g_2 \end{bmatrix} v_t \quad (3.10)$$

We could in principle repeat this procedure to form ever larger state space systems, including higher and higher orders of estimates. However, the three dimensional system (3.10) is sufficient to illustrate how common knowledge of rationality can be used to construct a law of motion for higher order expectations.

3.3. Expectations dynamics and common knowledge of rationality. Above, we implicitly assumed that rational expectations were common knowledge to derive the law of motion for the hierarchy of contemporaneous expectations (3.10): Individual agents used their knowledge of the process (3.1) - (3.2) and the information available to them to form their best estimates of θ_t . To estimate the average estimate of θ_t individual agents used their knowledge of the actual updating equation of other agents (3.4) to form a second order estimate of θ_t . Common knowledge of rationality simply means that agents' first order expectations are rational expectations of the actual value of the variable, that second order expectations are rational expectations of average first order expectations, that third order expectations are rational expectation of average second order expectations, and so on. Assumption 1 formalises this notion and extends it to also include higher order expectations of endogenous variables.

Assumption 1: It is common knowledge that agents' expectations are rational (model consistent). Let $\mathcal{M} : \mathbb{R}^\infty \rightarrow \mathbb{R}^\infty$ be a mapping from the hierarchy of contemporaneous expectations of θ_t in period t to the expected hierarchy of contemporaneous expectations in period $t+1$ conditional on actual hierarchy in period t

$$E \left[\theta_{t+1|t+1}^{(0:\infty)} \mid \theta_{t|t}^{(0:\infty)} \right] \equiv \mathcal{M} \left(\theta_{t|t}^{(0:\infty)} \right) \quad (3.11)$$

and let $\mathcal{M}^s \equiv \mathcal{M} \left(\mathcal{M} \left(\dots \mathcal{M} \left(\theta_{t|t}^{(0:\infty)} \right) \right) \right)$ so that

$$E \left[\theta_{t+s|t+s}^{(0:\infty)} \mid \theta_{t|t}^{(0:\infty)} \right] = \mathcal{M}^s \left(\theta_{t|t}^{(0:\infty)} \right) \quad (3.12)$$

Common knowledge of rational expectations then implies that

$$E \left[\theta_{t+s|t+s}^{(0:\infty)} \mid \theta_{t|t}^{(k:\infty)} \right] = \mathcal{M}^s \left(\theta_{t|t}^{(k:\infty)} \right) \quad \forall k, s \geq 0 \quad (3.13)$$

Let $\mathcal{A} : \mathbb{R}^\infty \rightarrow \mathbb{R}$ be a mapping from the complete hierarchy of contemporaneous expectations of θ_t in period t to the endogenous variable z_t in period t

$$z_t^{(0)} = \mathcal{A} \left(\theta_{t|t}^{(0:\infty)} \right) \quad (3.14)$$

Common knowledge of rational expectations then implies that

$$E \left[z_{t+s}^{(0)} \mid \theta_{t|t}^{(k:\infty)} \right] = \mathcal{A} \left(\mathcal{M}^s \left(\theta_{t|t}^{(k:\infty)} \right) \right) \quad \forall k, s \geq 0 \quad (3.15)$$

Assumption 1 is a natural generalisation of the full (or common) information rational expectations assumption. To see this, note that (3.11) - (3.15) imply that if agents share the same contemporaneous expectation hierarchy of the fundamental θ_t , they will also share the same expectations of future hierarchies of contemporaneous expectations and of future endogenous variables.

The mapping \mathcal{M} represents the actual law of motion for the contemporaneous expectations hierarchy. For something to be ‘common knowledge’ it is not enough that something is commonly believed, it must also be true. Setting $k = 1$ in (3.13) and (3.15) makes agents’ expectations rational. That (3.13) applies to all $k \geq 0$ makes it *common* knowledge.

3.4. A one-period ahead average expectations operator. In this section we define an operator that can be used to compute dynamic higher order expectations conditional on the expected future information sets of other agents. Let

$$\theta_{t|t}^{(0:\infty)} = M \theta_{t-1|t-1}^{(0:\infty)} + N v_t \quad (3.16)$$

describe the law of motion for the contemporaneous expectation hierarchy $\theta_{t|t}^{(0:\infty)}$. The difference equation (3.16) is the infinite dimensional equivalent of (3.10). Common knowledge of rational expectations then implies that

$$\int E \left[\theta_{t+1|t+1}^{(0:\infty)} \mid I_t(j) \right] dj = \int M \theta_{t|t}^{(1:\infty)}(j) dj \quad (3.17)$$

or equivalently

$$\int E \left[\theta_{t+1|t+1}^{(0:\infty)} \mid I_t(j) \right] dj = M \left[\mathbf{0}_{\infty \times 1} \quad I \right] \theta_{t|t}^{(0:\infty)} \quad (3.18)$$

Define a new operator $\overline{M} : \mathbb{R}^\infty \rightarrow \mathbb{R}^\infty$

$$\overline{M} \equiv M \left[\mathbf{0}_{\infty \times 1} \quad I \right] \quad (3.19)$$

Applying the operator \overline{M} to an expectation hierarchy does two things: It moves expectations one step ‘up’ in orders of expectations since

$$\theta_{t|t}^{(k+1:\infty)} = \left[\mathbf{0}_{\infty \times 1} \quad I \right] \theta_{t|t}^{(k:\infty)} \quad (3.20)$$

and one step forward in time by (3.16). The difference between M and \overline{M} is that M is a mapping from an agent's current estimate of $\theta_{t|t}^{(0:\infty)}$ to the *same agent's* estimate of what $\theta_{t|t}^{(0:\infty)}$ will be tomorrow, while \overline{M} maps an agent's current estimate of $\theta_{t|t}^{(0:\infty)}$ into what he believe *others* (on average) expect $\theta_{t|t}^{(0:\infty)}$ to be tomorrow. \overline{M} can therefore be used to compute dynamic higher order expectations of the form

$$\theta_{t+k|\dots|t}^{(k)} \equiv \int E\left[\int E\left[\dots \int E\left[\int E\left[\theta_{t+k} \mid I_{t+k-1}(j)\right] dj \mid I_{t+k-2}(j) dj\right] \right. \right. \right. \right. \right. \right. \quad (3.21)$$

$$\dots \mid I_{t+1}(j) dj \mid I_t(j) dj \quad (3.22)$$

for a given current hierarchy contemporaneous expectations $\theta_{t|t}^{(0:\infty)}$ since

$$\theta_{t+k|\dots|t}^{(k)} = e_1' \overline{M}^k \theta_{t|t}^{(0:\infty)} \quad (3.23)$$

where e_1' is a row vector that picks out the first element of a column vector.

The two operators M and \overline{M} can thus be used to compute any order of expectation at any horizon, and the choice between the two depends on the particular question at hand. In the simple set up of this section we did not have a model to motivate why agents were interested in forecasting other agents' expectations. In the next section, we present a model where privately informed traders buy and sell a long lived asset. In that model, traders will need to form higher order expectations about expectations that other traders will hold at some point in the future when new signals have arrived and the information sets of other traders have been updated. In that setting, the operator \overline{M} will be used to compute the current equilibrium price of the asset.⁵

4. THE SINGLETON ASSET PRICING MODEL

In the previous section, higher order expectations were discussed without any references to a model to motivate why agents needed to form higher order expectations. In this section we present a version of the model of Singleton (1987) with disparately informed traders that will serve as the vehicle for the argument of the rest of the paper. Having an explicit model not only make the analysis more interesting, it also allows us to introduce *endogenous signals*.

Singleton presents and solves a number of models that differ slightly in their patterns of persistence and assumed structural parameter values. In what he refers to as Models 1-7, the unobservable fundamental process follows an MA(2) process and in Models 8-12 it follows an AR(1). In the first class of models, a finite dimensional state representation can be found without making strong assumptions about the revelation of the shocks since a private signal about a MA(2) process does not carry information that is useful for forecasts beyond a two period horizon. Private information about an AR(1) process on the other hand is long lived. To solve the second class of models, Singleton assumes that the innovations to the AR(1)

⁵Allen, Morris and Shin (2006) defines an *average belief operator* $\overline{E} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$. The operator \overline{E} maps the average k order expectations of the average signal vector into $k+1$ order expectations of the same vector and can be used to compute higher order expectations of the state since the static setting results in a proportional relationship between higher order beliefs. In our model, the elements of N in the law of motion (3.16) could be generated by a similar operator if θ_t was a non-persistent process.

process are perfectly and publicly observed with a two period lag. This allows him to derive an accurate and finite dimensional state representation. The rest of this paper uses the same set up as in Singleton's Models 8-12 as a vehicle to show how dynamic models with private information can be solved *without* assuming that the shocks to the hidden process ever become common knowledge.

4.1. Model Set Up. There is a continuum of competitive traders indexed by $j \in (0, 1)$ who at time t divide their wealth between a risky asset with price p_t and coupon payment c_t and a risk free asset with return \bar{r} . The wealth of trader j then evolves according to

$$w_{t+1}(j) = z_t(j) [p_{t+1} + c_{t+1}] - [z_t(j)p_t - w_t(j)] (1 + \bar{r}) \quad (4.1)$$

where $z_t(j)$ is the asset holdings of trader j who chooses his portfolio to maximize

$$E [-e^{-\gamma w_{t+1}(j)} | I_t(j)] \quad (4.2)$$

and $I_t(j)$ is the information set of trader j at time t (defined below). The coupon payments follow the known autoregressive process

$$\begin{aligned} c_t &= \bar{c} + \psi c_{t-1} + u_t \\ u_t &\sim N(0, \sigma_u^2) \end{aligned} \quad (4.3)$$

Maximizing (4.2) subject to (4.1) yields agent j 's optimal demand for the risky asset

$$z_t^d(j) = \frac{(E[p_{t+1} | I_t(j)] - (1 + \bar{r})p_t) + (\bar{c} + \psi c_t)}{\gamma \delta} \quad (4.4)$$

where δ is the conditional variance of $(p_{t+1} + c_{t+1})$. The supply of the asset at time t , z_t^s , depends linearly on the price p_t and additively on the persistent stochastic shock θ_t and the i.i.d. disturbance ϵ_t

$$z_t^s = \xi p_t + \theta_t + \epsilon_t \quad (4.5)$$

$$\theta_t = \rho \theta_{t-1} + v_t \quad (4.6)$$

$$[\epsilon_t \ v_t]' \sim N\left(\mathbf{0}, \begin{bmatrix} \sigma_\epsilon^2 & 0 \\ 0 & \sigma_v^2 \end{bmatrix}\right) \quad (4.7)$$

Equating net demand and supply

$$\int z_t^d(j) = z_t^s \quad (4.8)$$

yields the equilibrium price

$$p_t = \lambda \left(\int E[p_{t+1} | I_t(j)] dj \right) + \lambda \psi c_t - \delta \gamma \lambda [\theta_t + \epsilon_t] \quad (4.9)$$

where

$$\lambda \equiv \frac{1}{\xi \gamma \delta + (1 + \bar{r})}. \quad (4.10)$$

4.2. Traders' Information Sets. The basic structure of the model described above is identical to Model 8-12 in Singleton (1983). Where the present paper differ is in the assumption on what traders can observe. In Singleton's paper the information set of trader j at time t is given by

$$I_t^S = \{s_{t-T}(j), p_{t-T}, c_{t-T} : T \geq 0; v_{t-T}, \epsilon_{t-T} : T \geq 2\} \quad (4.11)$$

where

$$s_t(j) = \theta_t + \eta_t(j) \quad (4.12)$$

$$\eta_t(j) \sim N(0, \sigma_\eta^2) \quad \forall j \quad (4.13)$$

Each trader observes the price of the asset, p_t , and the coupon payment, c_t , perfectly. The persistent component θ_t of the supply process is not perfectly revealed by the observation of the price due to the unobservable transitory supply shock ϵ_t . Trader j also observes a private signal $s_t(j)$ of the persistent supply process θ_t and it is due to the private measurement error $\eta_t(j)$ that the need to 'forecast the forecasts of others' arises. Singleton uses a similar method to overcome the infinite dimension of the state as Townsend (1983), i.e. he assumes that the shocks to the supply process become known to all traders after a finite number of periods (which in Singleton's case is after two periods). This allows for a finite dimensional time series representation of the model.

While the assumption of public revelation of shocks with a lag is convenient from a modeling perspective, it is not an assumption that is always realistic. We want to solve the model without imposing that all shocks are observed perfectly after a finite number of periods. The information set of our trader is therefore given by

$$I_t(j) = \{s_{t-T}(j), p_{t-T}, c_{t-T} : T \geq 0\} \quad (4.14)$$

Traders thus form expectations about the future price of the asset by observing the private signal $s_t(j)$, the commonly observable price p_t and the coupon payment c_t . It is common knowledge that all traders choose their portfolio to maximize (4.2) subject to the structural equations (4.3) - (4.6).

4.3. The common information solution. To solve the model we need to integrate out the average expectations term $\int E[p_{t+1} | I_t(j)] dj$. Under full information, this could be done by iterating (4.9) forward, yielding the solved price equation

$$p_t = \frac{\lambda\psi}{1 - \lambda\psi} c_t - \frac{\delta\gamma\lambda}{1 - \lambda\rho} \theta_t - \delta\gamma\lambda\epsilon_t \quad (4.15)$$

if $|\lambda\psi| < 1$ and $|\lambda\rho| < 1$. With imperfect but common information sets, the solution has the similar form

$$p_t = \frac{\lambda\psi}{1 - \lambda\psi} c_t - \delta\gamma\lambda\theta_t - \frac{\delta\gamma\lambda^2\rho}{1 - \lambda\rho} \theta_{t|t}^{(1)} - \delta\gamma\lambda\epsilon_t \quad (4.16)$$

where the actual supply disturbance θ_t in the geometric sum is replaced by the common first order estimate. Both of these solutions exploits the law of iterated expectations, but as demonstrated in the previous section, the law of iterated expectations is not generally applicable to higher order expectations. The next section derives an algorithm to solve the model when traders have private information.

5. A SOLUTION ALGORITHM

This section presents the solution algorithm for the model described above. To solve the model we need to find an appropriate representation of the state and the law of motion of the state. We also need to derive a mapping from the state to the price of the asset, using the state's law of motion. Finding the price function also includes computing the conditional variance δ of the sum of the price and the coupon payment, since this partly determines the rate at which expected future prices of the asset are discounted. The algorithm describes a fixed point problem in three steps.

5.1. Step 1: Computing the price. The first step of the solution is to find the price of the bond as a function of the contemporaneous expectation hierarchy of the supply disturbance $\theta_{t|t}^{(0:\infty)}$ for a given law of motion of the hierarchy and a given conditional variance of $(p_t + c_t)$, δ . The hierarchy of expectations is thus, together with the perfectly observable c_t and the transitory shock ϵ_t , the state of the model and the price function resembles a standard discounted expected sum of future fundamentals.

Proposition 1. *Conjecture (to be verified below) that the expectation hierarchy $\theta_{t|t}^{(0:\infty)}$ follows a vector AR(1) process of the form*

$$\theta_{t|t}^{(0:\infty)} = M\theta_{t-1|t-1}^{(0:\infty)} + N \begin{bmatrix} v_t \\ \epsilon_t \end{bmatrix} \quad (5.1)$$

The price function

$$p_t = \lambda \left(\int E[p_{t+1} | I_t(j)] \, dj \right) + \lambda\psi c_t - \delta\gamma\lambda[\theta_t + \epsilon_t] \quad (5.2)$$

can then be equivalently represented by a linear function of the contemporaneous expectation hierarchy, i.e.

$$p_t = \mathbf{a}\theta_{t|t}^{(0:\infty)} + \frac{\lambda\psi}{1 - \lambda\psi}c_t - \delta\gamma\lambda\epsilon_t \quad (5.3)$$

$$= \lambda \left(\int E[p_{t+1} | I_t(j)] \, dj \right) + \lambda\psi c_t - \delta\gamma\lambda[\theta_t + \epsilon_t] \quad (5.4)$$

where

$$\mathbf{a} = -\delta\gamma\lambda e_1 [I - \lambda\bar{M}]^{-1} \quad (5.5)$$

and

$$\bar{M} \equiv M \begin{bmatrix} \mathbf{0}_{\infty \times 1} & I \end{bmatrix} \quad (5.6)$$

Proof. In the Appendix. □

The computation of \mathbf{a} is conceptually similar to computing the expected discounted sum of future fundamentals under perfect (or imperfect but common) information. The difference in a set up with private information is that when (5.2) is iterated forward, we take repeated averages of expectations across investors. The fact that investors re-optimize their asset holdings in every period means that the price today depends on the average expectation of the price tomorrow, but the price tomorrow does not depend on what agents expect today

that the price of the asset will be the day after tomorrow. Instead, the price tomorrow depends on what agents will believe tomorrow that the price will be the day after tomorrow. Agents therefore need to form expectations *today* of the average expectation *tomorrow* of the price of the asset *the day after tomorrow* and so on. This is the reason why we need to use the operator \bar{M} (rather than M) that moves expectations both forward in time and ‘upwards’ in the hierarchy of expectations to compute the current price.

5.2. Step 2: The dynamics of the expectation hierarchy. This section shows how to find the law of motion for the hierarchy of expectations (5.1) for a given vector \mathbf{a} and conditional variance δ . The traders estimate the state of the model recursively, by applying the Kalman filter to the current price and the private signal of the supply disturbance. Computing the Kalman gain requires that the law of motion of the state that is being estimated is known. The state consists of the actual supply disturbance $\theta_{t|t}^{(0)}$ and the hierarchy of expectations of the supply disturbance $\theta_{t|t}^{(1:\bar{k})}$, so the law of motion of the state is determined by the actual supply process (4.6) and the law of motion of the higher order estimates. The Kalman filter thus plays a dual role: it both determines the traders’ estimate of the state as well as the law of motion of the very same state that the traders are estimating.

Conceptually, the procedure to construct the law of motion for the hierarchy of expectations here is the same as that of the simple set up in Section 3. The only difference is that traders now also observe an endogenous signal of the state, i.e. the price of the asset. The observable variables in the measurement equation then depend on the complete hierarchy of expectations from order zero to infinity, so we cannot add orders of expectations to the system one by one as we did in Section 3. Instead, we need to find an expression for the law of motion of the complete hierarchy of expectations from order zero to infinity simultaneously.

Trader j estimates the hierarchy of contemporaneous expectations recursively, using the Kalman filter updating equation (5.7)

$$\theta_{t|t}^{(1:\infty)}(j) = M\theta_{t-1|t-1}^{(1:\infty)}(j) + K \left(S_t(j) - LM\theta_{t-1|t-1}^{(1:\infty)}(j) - Qc_t \right) \quad (5.7)$$

$$S_t(j) = \begin{bmatrix} s_t(j) \\ p_t \end{bmatrix} \quad (5.8)$$

$$Q = \begin{bmatrix} 0 \\ \frac{\lambda\psi}{1-\lambda\psi} \end{bmatrix} \quad (5.9)$$

where $S_t(j)$ is the vector containing the private signal of the supply process $s_t(j)$ and the current price of the asset, p_t . M is the transition matrix from the conjectured law of motion (5.1). The term Qc_t subtracts the impact of the known coupon payment process on the price and the term in parenthesis in (5.7) is thus the surprise component of the observation, i.e. the innovation in the updating equation. The matrix L maps the expected state $M\theta_{t-1|t-1}^{(1:\bar{k})}$ into an expected observation and is given by

$$L = \begin{bmatrix} e'_1 \\ \mathbf{a} \end{bmatrix} \quad (5.10)$$

K is the Kalman gain matrix (defined below).

We want to find the conjectured vector AR(1) law of motion (5.1) for the hierarchy of average contemporaneous expectations, that is, we want to find the matrices M and N . We thus need to integrate the state updating equation (5.7) across traders and express all remaining terms as functions of the lagged expectation hierarchy $\theta_{t-1|t-1}^{(0:\infty)}$ and the two structural innovations v_t and ϵ_t . Use the definition of the private signal $s_t(j)$ (4.12), the price equation (5.3) and that the idiosyncratic noise ‘washes out’ in aggregation since $\int \eta_t(j) dj = 0$ to write the average signal S_t as

$$S_t = LM\theta_{t-1|t-1}^{(0:\infty)} + LN \begin{bmatrix} v_t \\ \epsilon_t \end{bmatrix} + \begin{bmatrix} 0 \\ -\delta\gamma\lambda\epsilon_t \end{bmatrix} \quad (5.11)$$

Substituting the average signal (5.11) into the updating equation (5.7) gives the law of motion of the average of traders’ estimate of the state

$$\theta_{t|t}^{(1:\infty)} = (I - KL)M\theta_{t-1|t-1}^{(1:\infty)} + KLM\theta_{t-1|t-1}^{(0:\infty)} + KLN \begin{bmatrix} v_t \\ \epsilon_t \end{bmatrix} + K \begin{bmatrix} 0 \\ -\delta\gamma\lambda\epsilon_t \end{bmatrix} \quad (5.12)$$

The final step to get the conjectured form (5.1) is to collect terms and append the actual supply disturbance process

$$\theta_t = \rho\theta_{t-1} + v_t \quad (5.13)$$

to the updating equation (5.12). The matrices M and N in the conjectured law of motion (5.1) are then given by

$$M = \begin{bmatrix} \rho & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ KLM \end{bmatrix} + \begin{bmatrix} 0 & \mathbf{0} \\ \mathbf{0} & (I - KL)M \end{bmatrix} \quad (5.14)$$

$$N = \begin{bmatrix} e'_1 \\ KLN \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ \mathbf{0} & -K_2\delta\gamma\lambda \end{bmatrix} \quad (5.15)$$

and the equilibrium Kalman gain matrix for this system is

$$K = PL'(LPL' + \Sigma_{\eta\epsilon})^{-1} \quad (5.16)$$

$$P = M(P - PL'(LPL' + \Sigma_{\epsilon\epsilon})^{-1}LP)M' + \Sigma_{\theta\theta} \quad (5.17)$$

$$\Sigma_{\eta\epsilon} = \begin{bmatrix} \sigma_\eta^2 & 0 \\ 0 & \sigma_\epsilon^2 \end{bmatrix} \quad (5.18)$$

$$\Sigma_{\theta\theta} = N \begin{bmatrix} \sigma_v^2 & 0 \\ 0 & \sigma_\epsilon^2 \end{bmatrix} N' \quad (5.19)$$

Proposition 2. *The transition matrix M is lower triangular, all elements are of the same sign as ρ and the elements of each row sum to ρ .*

Proof. In the Appendix □

In the simple example of Section 3, we could see that the transition matrix for the state equation was lower triangular. That this is also the case when agents can observe endogenous signals is not obvious, since higher order expectations affect the observed price and could therefore in principle affect lower order expectations. That the dynamics of lower order estimates are independent of higher order estimates will make it easier to show that the

infinite dimensional system considered so far can be approximated to an arbitrary accuracy by a finite dimensional system. That all elements are of the same sign as ρ is a direct implication of common knowledge of rational expectations and the properties of the Kalman filter. The third part of the proposition that the rows of M sum to ρ is also an implication of that rational expectations are common knowledge: If all orders of expectations coincide, then all orders of expected future values of θ_t must coincide as well. This will only be the case if the elements of the each row of M sum to ρ .

Proposition 3. *If the actual supply process $\{\theta_t\}$ follows an AR(1) with coefficient ρ , then the largest eigenvalue of M is ρ .*

Proof. Proposition 3 is a direct implication of Proposition 2. The eigenvalues of a lower triangular matrix are the elements along the diagonal (see Strang 1988), which by Proposition 2 are restricted to lie in the interval $(0, \rho)$. \square

Proposition 3 implies that (5.1) is a stable process and that a solution to the price equation exists when a fixed point for the law of motion exists.

5.3. Step 3: The conditional variance. The conditional variance of $(c_{t+1} + p_{t+1})$, δ , is the variance of investors' forecast error of the sum $c_{t+1} + p_{t+1}$ based on their period t information sets. The conditional forecast error is given by

$$\delta = \widehat{\mathbf{a}}\widehat{P}\widehat{\mathbf{a}}' + \left[1 + 2\frac{\lambda\psi}{1 - \lambda\psi} + \left(\frac{\lambda\psi}{1 - \lambda\psi} \right)^2 \right] \sigma_u^2 \quad (5.20)$$

$$\widehat{\mathbf{a}} = \left[\mathbf{a} \quad -\delta\gamma\lambda \right] \quad (5.21)$$

where \widehat{P} is the one period ahead joint forecast error covariance matrix of ϵ_t and the hierarchy of expectations of θ_t . Details on how to compute \widehat{P} are given in the Appendix.

5.4. A finite dimensional state representation. The derivations above contain infinite dimensional matrices and vectors. Below we present and prove a proposition that the infinite dimensional model solution derived above can be approximated to an arbitrary accuracy by a finite dimensional representation. The ground work for the proof is provided by two preceding lemmas. The first proves that the impact of contemporaneous expectations approaches zero as the order of expectation increases. The second proves that the variance of contemporaneous expectations decreases as the order of expectation increases. Taken together, these two results imply that there is an upper bound on the variance of the approximation error introduced by limiting the model to include only a finite number of orders of expectations. The upper bound of the approximation error variance is decreasing (at least) geometrically as the maximum number of orders included in the model increases. The approximation error introduced by truncating the state at a maximum order of expectation \bar{k} can thus be made arbitrarily small by choosing a large enough \bar{k} .

Lemma 1. *Denote the elements of the vector $\mathbf{a} = [a_1 \ a_2 \ a_3 \ \dots \ a_\infty]$. The k^{th} element of \mathbf{a} is smaller than*

$$|a_k| \leq \left| \frac{\delta\gamma\lambda}{1 - \lambda\rho} (\lambda\rho)^{k-1} \right| \quad (5.22)$$

and

$$\lim_{k \rightarrow \infty} a_k = 0 \quad (5.23)$$

Proof. That the matrix M in the law of motion (5.1) is lower triangular implies that a k th order contemporaneous expectation does not affect expectations in the forward iteration of (5.2) (using the operator \bar{M}) of lower order than k . That the elements of the rows of M sum to ρ imply that future higher order expectations of θ_t with horizon s are convex combinations of lower order expectations times ρ^s . The maximum impact a k th order contemporaneous expectation can have on the current price is therefore if the k th order contemporaneous expectation completely determines all expectations of order and horizon $> k$ so that

$$\theta_{t+k+s|\dots|t}^{(k+s)} = \rho^s \theta_{t|t}^{(k)} \quad \forall s > 0 \quad (5.24)$$

Iterating the price function forward, substituting in (5.24) and discounting by λ yields (5.22). That the sequence $\{a_k\}_{k=0}^{\infty}$ converges to zero is implied by $0 \leq |\lambda\rho| < 1$. Both δ and λ are endogenous variables, but the product $\delta\gamma\lambda$ is bounded by the definition of λ (4.10) to lie in the interval $(0, \frac{1}{\gamma})$ so the limit (5.23) holds for all permissible values of δ and λ . \square

Note that the condition $0 \leq |\lambda\rho| < 1$ that is required for the limit of the sequence $\{a_k\}_{k=1}^{\infty}$ to be zero is the same as what is required in the full (or common) information case for the model to be stable.

Lemma 2. *Common knowledge of rational expectation imply that the unconditional variance of expectations decreases as the order of expectation increases, i.e.*

$$E \left[\theta_{t|t}^{(k)} \theta_{t|t}^{(k)} \right] \geq E \left[\theta_{t|t}^{(k+1)} \theta_{t|t}^{(k+1)} \right] \quad (5.25)$$

Proof. Any two orders of expectation obey the identity

$$\theta_{t|t}^{(k)} \equiv \theta_{t|t}^{(k+1)} - e_t^{(k+1)} \quad (5.26)$$

where $e_t^{(k+1)}$ is defined as the $k+1$ order expectation error. The variance of the left hand side of (5.26) must equal the variance of the right hand side

$$E \left[\theta_{t|t}^{(k)} \theta_{t|t}^{(k)} \right] = E \left[\theta_{t|t}^{(k+1)} \theta_{t|t}^{(k+1)} \right] + E \left[e_t^{(k+1)} e_t^{(k+1)} \right] - 2E \left[\theta_{t|t}^{(k+1)} e_t^{(k+1)} \right]$$

The proof follows from the fact that variances are non-negative and that the covariance between the estimate and the error must be zero

$$E \left[\theta_{t|t}^{(k+1)} e_t^{(k+1)} \right] = 0. \quad (5.27)$$

That the covariance between the estimate and the error is zero is implied by rationality. To see why, consider a candidate estimate $\theta_{t|t}^{*(k+1)}$. If the covariance was not zero, a better estimate $\tilde{\theta}_{t|t}^{(k+1)}$ could be found by subtracting the projection of the error on the candidate estimate

$$\tilde{\theta}_{t|t}^{(k+1)} = \theta_{t|t}^{*(k+1)} - \frac{E \left[\theta_{t|t}^{(k+1)} e_t^{(k+1)} \right]}{E \left[\theta_{t|t}^{(k+1)} \theta_{t|t}^{(k+1)} \right]} \theta_{t|t}^{*(k+1)}$$

so if the candidate $\theta_{t|t}^{*(k+1)}$ is an optimal estimate of $\theta_{t|t}^{(k)}$, then (5.27) must hold. \square

Lemma 1 and 2 provide the ground work for the next proposition that when proved allows us to work with a finite dimensional representation of the model, where the required state dimension depends on the desired accuracy of the solution.

Proposition 4. *When a solution to the pricing equation (4.9) exists, it can be approximated to an arbitrary accuracy by*

$$p_t = \mathbf{a}_{0:\bar{k}} \theta_{t|t}^{(0:\bar{k})} + \frac{\lambda\psi}{1 - \lambda\psi} c_t - \delta\gamma\lambda\epsilon_t \quad (5.28)$$

where \bar{k} is a finite positive integer and

$$\theta_{t|t}^{(0:\bar{k})} \equiv \left[\theta_{t|t}^{(0)} \quad \theta_{t|t}^{(1)} \quad \dots \quad \theta_{t|t}^{(k)} \quad \dots \quad \theta_{t|t}^{(\bar{k})} \right]' \quad (5.29)$$

$$\mathbf{a}_{0:\bar{k}} \equiv \left[a_0 \quad a_1 \quad \dots \quad a_k \quad \dots \quad a_{\bar{k}} \right] \quad (5.30)$$

i.e. there exists a $\bar{k} \in \mathcal{N}$ such that

$$E \left(\mathbf{a} \theta_{t|t}^{(0:\infty)} - \mathbf{a}_{0:\bar{k}} \theta_{t|t}^{(0:\bar{k})} \right)^2 < \varepsilon : 0 < \varepsilon$$

Proof. The approximation error variance is given by

$$\begin{aligned} & E \left(\mathbf{a} \theta_{t|t}^{(0:\infty)} - \mathbf{a}_{0:\bar{k}} \theta_{t|t}^{(0:\bar{k})} \right)^2 \quad (5.31) \\ &= E \left(\mathbf{a}_{\bar{k}:\infty} \theta_{t|t}^{(\bar{k}:\infty)} \right)^2 \\ &= \mathbf{a}_{\bar{k}:\infty}' E \left[\theta_{t|t}^{(\bar{k}:\infty)} \theta_{t|t}^{(\bar{k}:\infty)'} \right] \mathbf{a}_{\bar{k}:\infty} \end{aligned}$$

which by lemma 1 and 2 can be made arbitrarily small by choosing a large enough \bar{k} . \square

By (5.22), the elements of \mathbf{a} decreases at least geometrically so the sum of the sequence $\{a_k\}_{k=1}^{\infty}$ also converges, but we also know what the sum converges to. Common knowledge of rational expectations imply that the sum must converge to

$$\sum_{k=1}^{\infty} a_k = \frac{\delta\gamma\lambda}{1 - \lambda\rho} \quad (5.32)$$

i.e. the sum of the elements in \mathbf{a} must be equal to the coefficient on θ_t in the full information solution (4.15) of the price equation since if by chance all orders of contemporaneous expectations of θ_t coincide, then so must all dynamic higher order expectations θ_t which means that the law of iterated expectations apply and that present discounted value of future expected values of θ_t must be given by (5.32).

5.5. Finding a fixed point. Solving the model implies finding a fixed point of equations (5.5),(5.14),(5.15),(5.16),(5.17) and (5.20). In practise we also need to choose a maximum order of expectations to include in the representation of the model. A tolerance criteria can be formulated as a ratio of the maximum approximation error variance (5.31) over the total variance of the asset's price. By Proposition 4, this ratio will approach zero as the maximum order \bar{k} is increased.

6. PRIVATE INFORMATION AND ASSET PRICE DYNAMICS

In the previous section we showed that the impact of higher order expectations on the price of the asset decreases as the order of expectation increases. If the impact decreases too quickly, higher order expectations may become irrelevant. Indeed, Singleton found that what mattered most in his model was that agents had *imperfect* information, rather than *private* information per se. In this section we show that this is not a general result, but highly dependent on how the model is parameterized. Broadly speaking, the privateness of information will matter more when expectations about the future matter more. That is, parameterizations that imply little discounting of the future or where the unobservable supply process is highly persistent will imply a larger difference in the price dynamics between the assumptions of private signals versus the case of equally precise but common signals.⁶ Obviously, parameterisations that imply little information imperfections at all will also result in small differences between private and common information price dynamics.

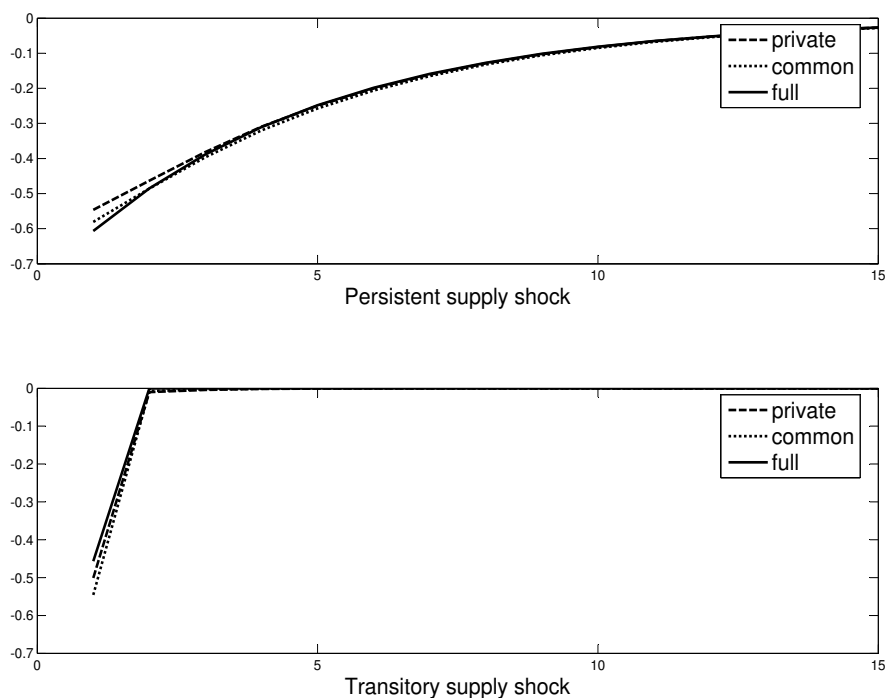


FIGURE 1. Price impulse responses under Singleton's baseline calibration

⁶Under common information it is assumed to be common knowledge that all traders observe the same unbiased signal $s_t = \theta_t + \eta_t$ where $\eta_t \sim N(0, \sigma_\eta^2)$.

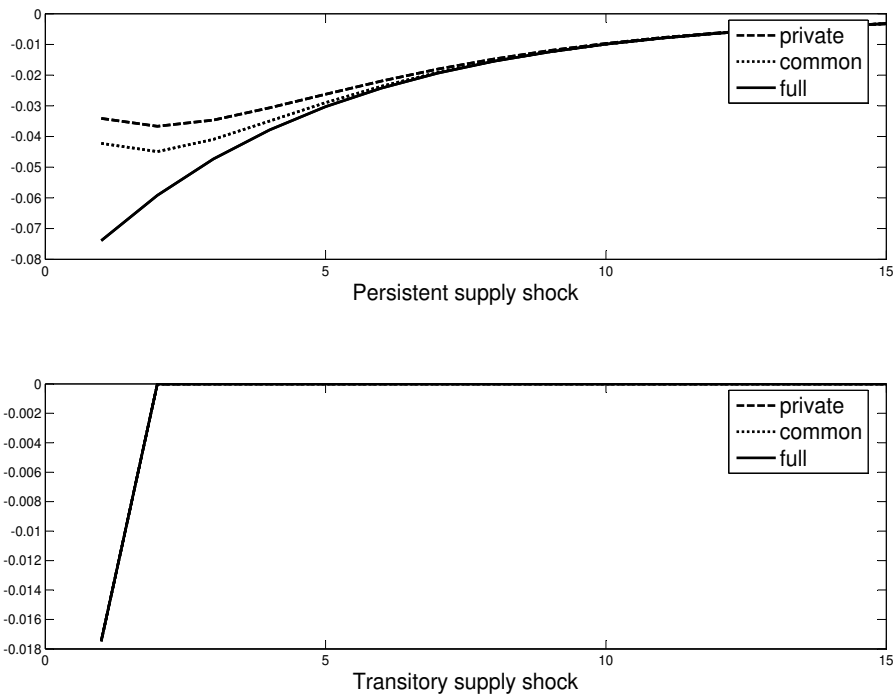


FIGURE 2. Price impulse responses with low exogenous variances

6.1. The role of the discount rate and the variances of shocks. Singleton's baseline calibration was to set $\{\gamma, \xi, \psi, \rho, \bar{r}, \sigma_u^2, \sigma_v^2, \sigma_\epsilon^2, \sigma_{\eta_i}^2\} =$

$\{2, 1.5, 0.7, 0.8, 0.02, 0.1, 1, 1, 2\}$. Figure 1 shows that this parameterisation does not imply very large differences between the full (solid line), private (dashed line), common (dotted line) information cases. This is partly due to the large specified variances. Since the conditional variance δ enters into the parameter λ that discounts the impact of higher order expectations on the price of the asset, so choosing variances are not just a matter of normalization. With the above parameterisation λ equals 0.27, or that a unit increase in the average expected price of the asset in period $t + 1$ implies an increase of only 0.27 in the price in period t . Multiplying the variances of the Singleton's baseline calibration by 0.01, increases the value of λ to 0.96. Figure 2 shows that with lower exogenous variances we get much larger differences between the price response to a persistent supply shock under full, private, and common information. This is because the higher order expectations in the sum (5.5) are discounted less, implying that the differences between first and higher order expectations matter more.

For both parameterisations above, the responses to the transitory shocks were similar across the three information structures. We can choose parameters such that we get large differences between the full, private and common information structures for both persistent

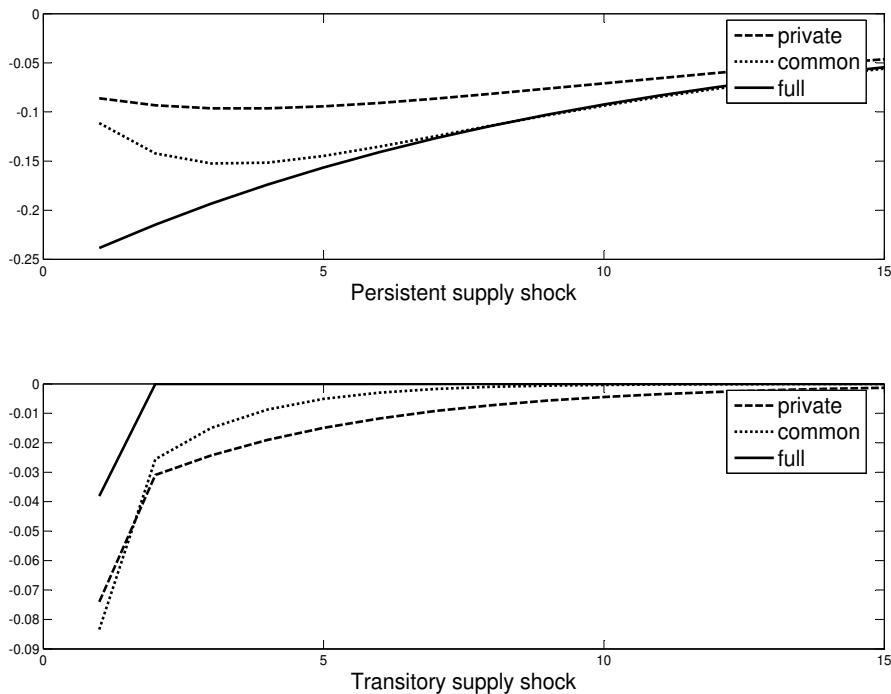


FIGURE 3. Price impulse responses

and transitory shocks. The key is to not choose the variance of either the persistent or transitory supply shock to be too large relative to the other, since traders will attribute most (of the surprise component) of what they observe to the shock with the dominant innovation variance. A parameterisation that achieves this balance and result in large differences in the responses to both persistent and transitory shocks is to set $\{\gamma, \xi, \psi, \rho, \bar{r}, \sigma_u^2, \sigma_v^2, \sigma_\epsilon^2, \sigma_{\eta_i}^2\}$ equal to $\{1, 1.5, 0.5, 0.9, 0.01, 0.01, .1, 0.001, 1\}$. The impulse of this parameterisation is displayed in Figure 3 which demonstrates that the different information structures imply very different price dynamics. Both private and common imperfect information results in weaker initial responses to a persistent supply shock compared to the full information case. The responses to the transitory shock has the opposite pattern in the impact period, i.e. information imperfections increases the first period response of the asset's price to a transitory shock. The reason for this is that the conditional variance of $c_t + p_t, \delta$, is lower when traders have imperfect information since this reduces the volatility of the price response to persistent shocks. (The impact of the transitory shock is the shock multiplied by $-\delta\gamma\lambda$.) Imperfect information also makes the price response to a transitory shock persistent and the persistence is stronger with private signals than with an equally precise common signal.

7. ENDOGENOUS SIGNALS AND REVELATION OF THE STATE

In the model presented above, traders observe the price of the asset and the price conveys some information to individual traders not only about the persistent supply process θ_t but also about the (higher order) expectations of other traders. Walker (forthcoming) has recently argued that in an environment like the Singleton model, equilibrium prices completely reveal the state of the model and therefore there is no need for traders to ‘forecast the forecast of others’, since all agents will share the same expectations. However, Walker does not actually prove this result using any of the original models presented in Singletons’s paper. Instead, Walker enlarges the information sets of the traders by making the additional assumption that traders can observe one component of the supply disturbance directly. In equilibrium, the price of the asset then reveals the second component of the supply disturbance perfectly.

Walker uses frequency domain methods to derive conditions for when prices reveal the state to the traders in the Singleton model. These conditions amount to checking whether the equilibrium MA representation of the vector of traders’ signals has any roots inside the unit circle. If all roots are outside the unit circle, then traders can back out all past innovations to the system perfectly. There is a simpler way: If equilibrium prices reveal the state perfectly, then the equilibrium in question must be the full information equilibrium. It is thus sufficient to check if the matrix that maps the period t state into trader j ’s vector of period t observations using the full information solution is invertible. If it isn’t, then there exists no fully revealing equilibrium.

In our model, each trader observes the coupon payment c_t , the price p_t and the private signal $s_t(j)$. From the full information solution (4.15) and the information set of trader j (4.14) we can write down trader j ’s observation equation as

$$\begin{bmatrix} c_t \\ p_t \\ s_t(j) \end{bmatrix} = \underbrace{\begin{bmatrix} \frac{\lambda\psi}{1-\lambda\psi} & 0 & 0 & 0 \\ 0 & -\frac{\delta\gamma\lambda}{1-\lambda\rho} & -\delta\gamma\lambda & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}}_{\widehat{L}} \begin{bmatrix} c_t \\ \theta_t \\ \epsilon_t \\ \eta_t(j) \end{bmatrix} \quad (7.1)$$

Clearly, the matrix \widehat{L} is not invertible. By assuming that either θ_t or ϵ_t is observable, Walker in effect adds a row to \widehat{L} which makes it invertible. Below, we argue that Walker’s result can also be understood as an outcome of a special parameterisation of Singleton’s model.

In the framework presented here, a perfectly revealing equilibrium price manifests itself by making all orders of expectations of θ_t coincide at all times so that $\theta_{t|t}^{(k)} = \theta_{t|t}^{(0)} \forall k, t$. The top panel of Figure 4 illustrates the response of the hierarchy of expectations of θ_t from order zero to 50 to a unit innovation in θ_t . (The parameterisation is the same as that used for Figure 3.) The thick solid line is the response of the actual shock, or $\theta_{t|t}^{(0)}$, the dashed line immediately beneath it is the first order expectations, the dotted line next is the second order expectation and so on.⁷ The transitory supply shock ϵ_t functions as aggregate noise that prevents the price from perfectly revealing θ_t . If we decrease its variance, equilibrium

⁷The top panel of Figure 4 also provides a nice illustration of Lemma 2.

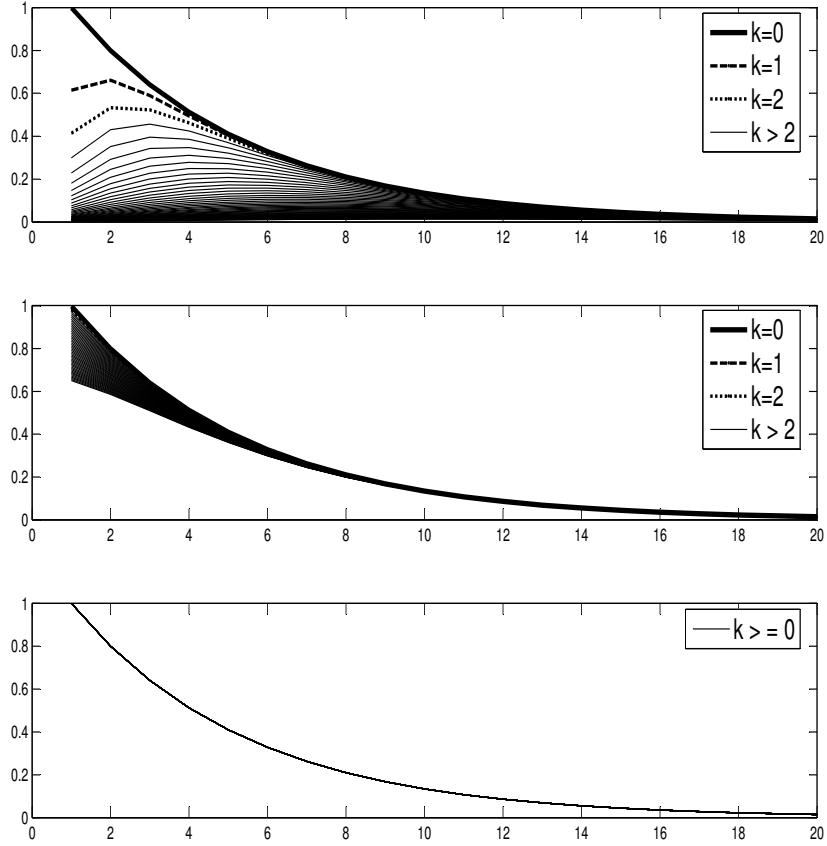


FIGURE 4. Impulse responses of $\theta_{t|t}^{(0:50)}$ to a unit innovation v_t

prices will be more informative about θ_t and other traders' (higher order) expectations of θ_t . This can be seen in the mid panel of Figure 5, where we have plotted a second impulse response function for the hierarchy $\theta_{t|t}^{(0:50)}$. The variance of ϵ_t in the mid panel is set to 1/10 of that in the top panel. It is clear that decreasing the variance of the transitory shock makes all orders of expectations move closer together, i.e. making traders better informed about all orders of expectations of θ_t .

From a filtering perspective, setting the variance of ϵ_t equal to zero is equivalent to making it perfectly observable. The bottom panel of Figure 4 demonstrates that the model with $\sigma_\epsilon^2 = 0$ replicates the result of Walker: Equilibrium prices perfectly reveal the value of θ_t so that all orders of expectations coincide and the graph collapses to a single line. However, this

is not a general property of Singleton’s model, but an artefact of the additional assumptions that $\sigma_\epsilon^2 = 0$, or equivalently, that traders can observe ϵ_t perfectly.

8. CONCLUSIONS

In this paper we derive a method for solving dynamic models with private information. The principal difficulty of solving models in this class is the infinite regress of expectations arising from agents’ need to ‘forecast the forecasts of others’. Here, we demonstrate how the infinite regress problem can be made tractable by imposing some structure on expectations. Specifically, it is common knowledge that agents are rational Bayesians. This assumption allows us to derive the dynamics of higher order expectations explicitly and transparently.

We use the structure imposed on expectations by common knowledge of rationality to solve a version of Singleton’s (1987) asset pricing model with privately informed traders. By defining an *average one-period-ahead expectations operator*, we derive an expression for the price of the asset as a geometric sum that resembles the present discounted value of expected future fundamentals. The word ‘resembles’ is used with care to indicate that while the functional form is similar to the corresponding expression in a full information model, there is an important difference since the price function is not derived by relying on the law of iterated expectations. Instead, the operator is used to compute a convergent sequence of higher order expectations of the future price of the asset. The current price of the asset is given by the discounted sum of this sequence.

Determining the dynamics of higher order expectations and how these map into the price of an asset does not by itself solve the infinite regress problem. However, it does provide us with a framework that is tractable enough to derive conditions under which the model can be approximated to an arbitrary accuracy by a finite dimensional state representation. Incidentally, this is the same condition that guarantees that a stable solution exists in the full (or common) information case: If the discount rate multiplied by the eigenvalue of the fundamental process is smaller than unity in absolute value, we only need to model a finite number of orders of expectations to achieve any required degree of accuracy.

It was Townsend’s (1983) paper that coined the term ‘forecasting the forecasts of others’. Subsequent research, i.e. Kasa (2000) and Pearlman and Sargent (2005), have showed that in the model studied by Townsend, private information is not preserved when agents observe equilibrium prices so there is actually no need for agents to ‘forecast the forecasts of others’. Walker (forthcoming) claims that this also applies to the model of Singleton studied here. However, above we showed that this result is due to additional (and special) assumptions made by Walker that are not part of Singleton’s original set up. In the model presented here, the number of innovations in each period relative to the number of signals are sufficient to ensure that equilibrium prices do not fully reveal the underlying state. We proposed a simple way to verify this based on that if equilibrium prices are fully revealing, then the equilibrium in question must be the full information equilibrium. (The same method can be used to verify that prices are also not fully revealing in Singleton’s original information structure.)

The literature has to date produced a wealth of qualitative results derived from the interactions that arise between agents when each individual has access to his own piece of

information. A natural next step is to test whether these qualitative results hold up when subjected to quantitative scrutiny. The solution method proposed in this paper allows us to solve dynamic models with private information accurately (and quickly) without making some of the modelling compromises previously thought to be necessary. In addition, the method delivers the solved model in a form that can be estimated directly by maximum likelihood methods. The hope is that this paper will help shorten the step from qualitative to quantitative results by opening up the possibility of using dynamic models with privately informed agents that are realistic enough to use for empirical work.

REFERENCES

- [1] Adam, Klaus, "Optimal Monetary Policy with Imperfect Common Knowledge", forthcoming *Journal of Monetary Economics*.
- [2] Allen, F. S. Morris and H.S., 2006, "Shin, Beauty Contests and Iterated Expectations in Asset Markets", *Review of Financial Studies*, 19, pp719 – 752.
- [3] Cespa, Giovanni and Xavier Vives, 2007, "Dynamic Trading, Asset Prices and Bubbles", mimeo, Università di Salerno and IESE Business School.
- [4] Bacchetta, Phillippe and Eric van Wincoop, (2006) "Can Information Heterogeneity Explain the Exchange Rate Determination Puzzle?", *American Economic Review* vol 96, pp552-576.
- [5] Kasa, Kenneth, 2000, "Forecasting the forecast of others in the frequency domain", *Review of Economic Dynamics*, 3, pp726-756.
- [6] Kasa, K., T. Walker, and C. Whiteman, 2007, "Asset Prices in a Time Series Model with Perpetually Disparately Informed, Competitive Traders", mimeo, Indiana University.
- [7] Lorenzoni, Guido, 2005, "A Theory of Demand Shocks" NBER working paper.
- [8] Morris, S. and H.S. Shin, 2002, "The social value of public information", *American Economic Review* 92, pp1521-1534.
- [9] Pearlman, J.G. and T.J. Sargent, 2005, "Knowing the forecasts of others", *Review of Economic Dynamics*, Volume 8, pp480-497.
- [10] Sargent, Thomas J., 1991, "Equilibrium with Signal Extraction from Endogenous Variables", *Journal of Economic Dynamics and Control* 15, pp245-273.
- [11] Singleton, Kenneth J., 1987, "Asset prices in a time series model with disparately informed, competitive traders", in *New Approaches to Monetary Economics*, Eds. W.A. Burnett and K.J. Singleton, Cambridge University Press.
- [12] Strang, Gilbert, 1988, "Linear Algebra and Its Applications", 3rd ed., Saunders.
- [13] Townsend, Robert M., 1983, Forecasting the Forecasts of Others, *Journal of Political Economy*, vol 91, pp546-588.
- [14] Walker, Todd, 2006, "How Equilibrium Prices Reveal Information in Time Series Models with Disparately Informed, Competitive Traders", forthcoming, *Journal of Economic Theory*.
- [15] Weinstein, J. and M. Yildiz, forthcoming, "Impact of Higher-Order Uncertainty", *Games and Economic Behavior*.
- [16] Woodford, M. 2002, "Imperfect Common Knowledge and the Effects of Monetary Policy," in P. Aghion, R. Frydman, J. Stiglitz, and M. Woodford, eds., Knowledge, Information, and Expectations in Modern Macroeconomics: In Honour of Edmund S. Phelps, Princeton: Princeton University Press.

APPENDIX A. PROOF OF PROPOSITION 1

Let the expectation hierarchy $\theta_{t|t}^{(0:\infty)}$ follow a stable VAR(1) process of the form

$$\theta_{t|t}^{(0:\infty)} = M\theta_{t-1|t-1}^{(0:\infty)} + N \begin{bmatrix} v_t \\ \epsilon_t \end{bmatrix} \quad (\text{A.1})$$

The price function

$$p_t = \lambda \left(\int E[p_{t+1} | I_t(j)] \, dj \right) + \lambda\psi c_t - \delta\gamma\lambda [\theta_t + \epsilon_t] \quad (\text{A.2})$$

can then be equivalently represented by a linear function of the contemporaneous expectation hierarchy, i.e.

$$p_t = \mathbf{a}\theta_{t|t}^{(0:\infty)} + bc_t - d\epsilon_t \quad (\text{A.3})$$

$$= \lambda \left(\int E[p_{t+1} | I_t(j)] \, dj \right) + \lambda\psi c_t - \delta\gamma\lambda [\theta_t + \epsilon_t] \quad (\text{A.4})$$

where

$$\mathbf{a} = \delta\gamma\lambda e_1 [I - \lambda\bar{M}]^{-1} \quad (\text{A.5})$$

and

$$\bar{M} \equiv M \begin{bmatrix} \mathbf{0}_{\infty \times 1} & I \end{bmatrix} \quad (\text{A.6})$$

Proof. Iterate (4.9) forward

$$\begin{aligned} p_t &= \frac{\lambda\psi}{1 - \lambda\psi} c_t - \delta\gamma\lambda\epsilon_t - \delta\gamma\lambda\theta_t \\ &\quad - (\delta\gamma\lambda)\lambda \left(\int E[\theta_{t+1} | I_t(j)] \, dj \right) \\ &\quad - (\delta\gamma\lambda)\lambda^2 \int E \left[\int E[\theta_{t+2} | I_{t+1}(j)] \, dj \mid I_t(j) \right] dj - \dots \\ &\quad \dots - (\delta\gamma\lambda)\lambda^n \int E \left[\int E \left[\int E[\theta_{t+n} | I_{t+n}(j)] \, dj \mid I_{t+k-1}(j) \right] \right. \\ &\quad \left. \mid I_{t+1}(j) \right] dj \mid I_t(j) \right] dj \dots \end{aligned} \quad (\text{A.7})$$

i.e. the price is a function of the average expectation in period t of the value of the supply shock θ in period $t+1$ and the average expectation in period t of the average expectation of the value of θ in period $t+2$ and so on. We can use the operator $\bar{M} = M \begin{bmatrix} \mathbf{0} & I \end{bmatrix}$ from Section 3 to write the price function (A.7) as

$$p_t = \frac{\lambda\psi}{1 - \lambda\psi} c_t - \delta\gamma\lambda\epsilon_t - \delta\gamma\lambda \sum_{s=0}^{\infty} \lambda^s e_1 \bar{M}^s \theta_{t|t}^{(0:\infty)} \quad (\text{A.8})$$

or equivalently

$$p_t = \frac{\lambda\psi}{1 - \lambda\psi} c_t - \delta\gamma\lambda\epsilon_t - \delta\gamma\lambda e_1 [I - \lambda\bar{M}]^{-1} \theta_{t|t}^{(0:\infty)} \quad (\text{A.9})$$

For the inverse $[I - \lambda \overline{M}]^{-1}$ to exist, the eigenvalues of $\lambda \overline{M}$ must lie inside the unit circle. From Proposition 2 and the definition of \overline{M} we know that the elements of the rows of \overline{M} are of the same sign as, and sum to, ρ . The Perron-Frobenius Theorem (see Strang 1988) gives a bound on the absolute magnitude of the largest eigenvalue of \overline{M} . It states if \mathbf{M} is an n -dimensional positive matrix, it has a unique positive eigenvalue Λ_1 such that

$$\Lambda_1 > |\Lambda_i| : i = 2, 3, \dots, n \quad (\text{A.10})$$

$$\min_i \sum_{j=1}^n m_{ij} \leq \Lambda_1 \leq \max_i \sum_{j=1}^n m_{ij} \quad (\text{A.11})$$

or that the largest eigenvalue Λ_1 is bounded by the maximum of the sums of the elements of the rows of \mathbf{M} and that the modulus of the remaining eigenvalues are bounded by Λ_1 . This ensures that the modulus of the eigenvalues of $\lambda \overline{M}$ are bounded by $\lambda \rho$ and that the inverse $[I - \lambda \overline{M}]^{-1}$ exists. \square

APPENDIX B. PROOF OF PROPOSITION 2

The transition matrix M is lower triangular, all elements are non-negative and the sum of the elements in each row sum to ρ .

Proof. The transition matrix for the hierarchy $\theta_{t|t}^{(0:\infty)}$ is given by

$$M = \begin{bmatrix} \rho & \mathbf{0}_{1 \times \infty} \\ (K_1 - \delta \gamma \lambda K_2) \rho & (I - KL) M + K_2 \mathbf{a}_{2:\infty} M_{1:\infty} \end{bmatrix} \quad (\text{B.1})$$

We thus need to show that the lower right hand block $(I - KL) M + K_2 \mathbf{a}_{1:\infty} M_{1:\infty}$ is lower triangular. Using that

$$L = \begin{bmatrix} 1 & \mathbf{0} & 0 \\ a_1 & \cdots & a_\infty \end{bmatrix} \quad (\text{B.2})$$

and some manipulation yields

$$\begin{aligned} & (I - KL) M + K_2 \mathbf{a}_{1:\infty} M_{1:\infty} = \\ & \left(I - \begin{bmatrix} k_{11} & \mathbf{0} \\ \vdots & \mathbf{0} \\ k_{\infty 1} & \mathbf{0} \end{bmatrix} \right) M + \\ & \begin{bmatrix} \sum_{s=2}^{\infty} k_{12} a_s m_{s1} & \sum_{s=2}^{\infty} k_{12} a_s m_{s2} & \cdots & \sum_{s=2}^{\infty} k_{12} a_s m_{s\infty} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{s=2}^{\infty} k_{\infty 2} a_s m_{s1} & \sum_{s=2}^{\infty} k_{\infty 2} a_s m_{s2} & \cdots & \sum_{s=2}^{\infty} k_{\infty 2} a_s m_{s\infty} \end{bmatrix} - \\ & \begin{bmatrix} \sum_{s=1}^{\infty} k_{12} a_s m_{s1} & \sum_{s=1}^{\infty} k_{12} a_s m_{s2} & \cdots & \sum_{s=1}^{\infty} k_{12} a_s m_{s\infty} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{s=1}^{\infty} k_{\infty 2} a_s m_{s1} & \sum_{s=1}^{\infty} k_{\infty 2} a_s m_{s2} & \cdots & \sum_{s=1}^{\infty} k_{\infty 2} a_s m_{s\infty} \end{bmatrix} \end{aligned} \quad (\text{B.3})$$

where m_{ij} denotes the element in the i th row and the j th column of M . By exogeneity of θ_t , we know that $m_{1,j} = 0 : j > 1$ which implies that $\sum_{s=2}^{\infty} k_{12} a_s m_{s2} = \sum_{s=1}^{\infty} k_{12} a_s m_{s2}$ so the terms appearing in the second row and off the diagonal in (B.3) cancel out so that

$m_{2,j} = 0 : 1 < j$. We have now established that all elements in row 1 and 2 of M above the diagonal are zero. Induction yields $m_{ij} = 0 \forall i, j : i < j$ and $i = 1, 2, \dots, \infty$, i.e. the matrix M is lower triangular. That the rows sum to ρ is a direct implication of the Assumption 1. Common knowledge of rationality implies that if all orders of contemporaneous expectations coincide, then so must all expectations about future values of the θ_t which will only be the case if the rows sum to ρ . An alternative proof is by direct computation of the elements in (B.1). This shows that the sum of row i of M equals the sum of row $i - 1$. We know from the exogeneity of θ_t that $\Sigma m_{1,j} = \rho$, and the proof then follows from induction. \square

APPENDIX C. COMPUTING THE PRICE FUNCTION BY EULER EQUATION ITERATION

An alternative (and computationally faster) way to find the vector \mathbf{a} is by exploiting the Euler-type structure of (5.2) and the method of undetermined coefficients. Substitute (5.3) into the price function (5.2) to get

$$\begin{aligned} & \mathbf{a}\theta_{t|t}^{(0:\infty)} + \frac{\lambda\psi}{1-\lambda\psi}c_t - \delta\gamma\lambda\epsilon_t \\ &= \lambda \left(\int E \left[\left(\mathbf{a}\theta_{t+1|t+1}^{(0:\infty)} + \frac{\lambda\psi}{1-\lambda\psi}c_{t+1} - \delta\gamma\lambda\epsilon_{t+1} \right) \mid I_t(j) \right] dj \right) + \lambda\psi c_t - \delta\gamma\lambda[\theta_t + \epsilon_t] \end{aligned} \quad (\text{C.1})$$

The fact that

$$\frac{\lambda\psi}{1-\lambda\psi}c_t - \delta\gamma\lambda\epsilon_t = \lambda E \left[\frac{\lambda\psi}{1-\lambda\psi}c_{t+1} - \delta\epsilon_{t+1} \right] + \lambda\psi c_t - \delta\gamma\lambda\epsilon_t \quad (\text{C.2})$$

leaves us with

$$\mathbf{a}\theta_{t|t}^{(0:\infty)} = \lambda \left(\int E \left[\mathbf{a}\theta_{t+1|t+1}^{(0:\infty)} \mid I_t(j) \right] dj \right) - \delta\gamma\lambda\theta_t \quad (\text{C.3})$$

Common knowledge of rationality implies that

$$\int E \left[\theta_{t+1|t+1}^{(0:\infty)} \mid I_t(j) \right] dj = M\theta_{t|t}^{(1:\infty)} \quad (\text{C.4})$$

Substituting (C.4) back into (C.3) yields

$$\mathbf{a}\theta_{t|t}^{(0:\infty)} = \lambda\mathbf{a}M\theta_{t|t}^{(1:\infty)} - \delta\gamma\lambda\theta_t \quad (\text{C.5})$$

Equating coefficients imply that

$$\mathbf{a}_{(1:\infty)} = \lambda\mathbf{a}_{(0:\infty)}M \quad (\text{C.6})$$

$$a_0 = -\delta\gamma\lambda \quad (\text{C.7})$$

We can compute the elements of \mathbf{a} for a given M and δ , by iterating on (C.6) starting from $a_0 = -\delta\gamma\lambda$. This iterative process is faster than using the formula (A.5) since it does not involve any matrix inversions.

APPENDIX D. COMPUTING THE CONDITIONAL VARIANCE

The conditional variance of $(c_{t+1} + p_{t+1})$, δ , is the variance of investors' forecast error of the sum $c_{t+1} + p_{t+1}$ based on their information sets in period t and is given by

$$\delta = E \left[\left(1 + \frac{\lambda\psi}{1 - \lambda\psi} \right) u_t + \mathbf{a}\theta_{t|t}^{(0:\infty)} - \mathbf{a}M\theta_{t-1|t-1}^{(1:\infty)} - \delta\gamma\lambda\epsilon_t \right]^2 \quad (\text{D.1})$$

which can be rearranged to

$$\begin{aligned} \delta &= \left[1 + 2\frac{\lambda\psi}{1 - \lambda\psi} + \left(\frac{\lambda\psi}{1 - \lambda\psi} \right)^2 \right] \sigma_u^2 \\ &\quad + \mathbf{a}P\mathbf{a}' + (\delta\gamma\lambda)^2 \sigma_\epsilon^2 - 2E \left[\left(\mathbf{a}\theta_{t|t}^{(0:\infty)} - \mathbf{a}M\theta_{t-1|t-1}^{(1:\infty)} \right) \delta\gamma\lambda\epsilon_t \right] \end{aligned} \quad (\text{D.2})$$

The expression on the second line of (D.2) can be computed by putting the hierarchy of contemporaneous expectations into state space form together with the transitory supply shock ϵ_t

$$\begin{bmatrix} \theta_{t|t}^{(0:\infty)} \\ \epsilon_t \end{bmatrix} = \begin{bmatrix} M & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \theta_{t-1|t-1}^{(0:\infty)} \\ \epsilon_{t-1} \end{bmatrix} + \begin{bmatrix} N_1 & N_2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v_t \\ \epsilon_t \end{bmatrix} \quad (\text{D.3})$$

$$\begin{bmatrix} s_t(j) \\ p_t - \frac{\lambda\psi}{1 - \lambda\psi} c_t \end{bmatrix} = \begin{bmatrix} e'_1 & 0 \\ \mathbf{a} & -\delta\gamma\lambda \end{bmatrix} \begin{bmatrix} \theta_{t|t}^{(0:\infty)} \\ \epsilon_t \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} \eta_t(j) \quad (\text{D.4})$$

Define

$$X_t \equiv \begin{bmatrix} \theta_{t|t}^{(0:\infty)} \\ \epsilon_t \end{bmatrix} \quad (\text{D.5})$$

$$\hat{P} \equiv E (X_t - X_{t|t-1}) (X_t - X_{t|t-1})' \quad (\text{D.6})$$

$$\hat{\mathbf{a}} \equiv \begin{bmatrix} \mathbf{a} & -\delta\gamma\lambda \end{bmatrix} \quad (\text{D.7})$$

then

$$\hat{\mathbf{a}}\hat{P}\hat{\mathbf{a}}' = \mathbf{a}P\mathbf{a}' + (\delta\gamma\lambda)^2 \sigma_\epsilon^2 - 2E \left[\left(\mathbf{a}\theta_{t|t}^{(0:\infty)} - \mathbf{a}M\theta_{t-1|t-1}^{(1:\infty)} \right) \delta\gamma\lambda\epsilon_t \right] \quad (\text{D.8})$$

where \hat{P} is the one period ahead forecast error covariance matrix associated with the state space system (D.3)-(D.4). The conditional variance of the sum of the coupon payment and the price is then given by

$$\delta = \hat{\mathbf{a}}\hat{P}\hat{\mathbf{a}}' + \left[1 + 2\frac{\lambda\psi}{1 - \lambda\psi} + \left(\frac{\lambda\psi}{1 - \lambda\psi} \right)^2 \right] \sigma_u^2. \quad (\text{D.9})$$