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 derives properties of higher order expectations in a linear-Gaussian setting where it is common knowledge that agents form expectations rationally. The usefulness of the approach is demonstrated by solving Singleton's (1987) dynamic asset pricing model with disparately informed traders but without assuming that shocks can be observed perfectly with a lag. Under the same parameter restrictions that guarantee that a solution exists under full information, the impact of expectations can be shown to decrease with the order of expectation. This allows for a finite dimensional equilibrium representation that can be made arbitrarily accurate. The method is generally applicable to linear-Gaussian rational expectations models with private information.
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: July 8, 2010. The author thanks Jesus Fernandez-Villaverde, Mirko Wiederholt, 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, Australian Workshop for Macro Dynamics, Institute for International Economic Studies at Stockholm University, University of Amsterdam, and the 2009 Econometric Society meeting in San Francisco for useful comments and suggestions. Financial support from Ministerio de Ciencia e Innovacion (ECO2008-01665), Generalitat de Catalunya (2009SGR1157), Barcelona GSE Research Network and the Government of Catalonia is gratefully acknowledged.
Address: CREI, Universitat Pompeu Fabra, Ramon Trias Fargas 25-27, Barcelona 08005.
e-mail: knimark@crei.cat *web page*: www.kris-nimark.net .

[1] Townsend (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 modeling 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 the paper, it is assumed that it is common knowledge that agents form expectations rationally. That is, all agents know that all agents know, and so on that, all agents form optimal expectations given their information sets. Common knowledge of rationality gives enough structure 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 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 approximation error introduced if one considered only a finite number of orders of expectations converges to zero as the maximum order of expectation considered increases. The dynamics of an infinite horizon model with private information and with agents that optimize intertemporally can thus be approximated to an arbitrary accuracy by a finite state representation.

The results described above are derived without relying on the common strategy of making additional assumptions to ensure that private information is short lived. 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,

[2]Some 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).

[3]A 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.

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

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 generalized 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.

The frequency domain methods have recently been extended by Rondina and Walker (2010) who show that for carefully chosen process for fundamentals, endogenous variables can display waves of optimism and pessimism. Their method is particularly suited to models where there is a finite order MA component and the number of signals observed by a=gents is the same as the number of shocks in the fundamental process. In contrast, the method proposed here is suitable for the more common filtering problem with more shocks than observables, so that non-invertibility of the equilibrium process is guaranteed. That is, there are no parameter configurations that make the state an invertible function of observables, as long as enough shocks have strictly positive variance.

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 (2007) shows that when the information set of the traders in Singleton's model is enlarged to include one of the shocks that are latent in Singleton's original model, equilibrium prices reveal the other latent state perfectly. In general, private information is preserved in Singleton's model, but we show that a special case, when one of the fundamental shocks have zero variance, replicates the result of Walker, so that all order of expectations collapses to a single value coinciding with the true value of the fundamental.

The next section defines the mathematical space the analysis will take place in as well as defines notation for dynamic higher order expectations. This is followed by a brief presentation of Singleton's model, that will serve as a vehicle for the rest of the paper. Section 4 derives properties of higher order expectations that must hold in any equilibrium. Particularly, it here that we show that the variance of higher order expectation are bounded by the variance of the true process. Section 5 introduces an average expectations operator and shows how it can be used to compute the higher order expectation that determine the equilibrium price of the asset. Section 6 contains the main results of the paper. It is here that the approximation results are presented, demonstrating that a finite number of orders of expectations are sufficient for an arbitrarily accurate representation of equilibrium. Section 7 presents and algorithm to find an equilibrium and proves that an equilibrium exists, under general conditions. Section 8 presents properties of the solved model and shows that in practise, only a low number of orders of expectations are necessary as equilibrium dynamics converge rapidly as the maximum order of expectation is increased. Section 9 concludes, partly by discussing the generality of the results.

2. Preliminaries

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

2.1. The inner product space L^2.

In the model presented in the next section, the signals that traders observe and their expectations of fundamentals and endogenous variables are elements of the inner product space L^2, which we now define.

Definition 1. *(The inner-product space L^2.) The inner product space L^2 is the collection of all random variables X with finite variance*

$$EX^2 < \infty \tag{2.1}$$

and with inner-product

$$\langle X, Y \rangle \equiv E(XY) : X, Y \in L^2 \tag{2.2}$$

Definition 2. *Let Ω be a subspace of L^2. An orthogonal projection of X onto Ω , denoted $\mathcal{P}_\Omega X$, is the unique element in L^2 satisfying*

$$\langle X - \mathcal{P}_\Omega X, \omega \rangle = 0 \tag{2.3}$$

for any $\omega \in \Omega$.

In a linear model with Gaussian shocks, conditional expectations are equivalent to orthogonal projections. The equality

$$E\left(X \mid \Omega\right) = \mathcal{P}_{\Omega}X \tag{2.4}$$

thus implies that the conditional expectations in the model share the properties of orthogonal projections in L^2.

2.2. **Defining higher order expectations.** There is a continuum of agents are indexed by $j \in (0,1)$. Agent j's first order expectation of a variable $\theta_t \in L^2$ conditional on his period t information set $\Omega_t(j)$ is denoted as

$$\theta_t^{(1)}(j) \equiv E\left[\theta_t \mid \Omega_t(j)\right] \tag{2.5}$$

The average first order expectation $\theta_t^{(1)}$ is obtained by taking averages of (2.5) across agents

$$\theta_t^{(1)} \equiv \int E\left[\theta_t \mid \Omega_t(j)\right] \quad dj \tag{2.6}$$

The average second order expectation is obtained by taking the average of agents' expectations of (2.6)

$$\theta_t^{(2)} \equiv \int E\left[\theta_t^{(1)} \mid \Omega_t(j)\right] \; dj \tag{2.7}$$

and so on so that the k^{th} order expectation of θ_t is given by

$$\theta_t^{(k)} \equiv \int E\left[\theta_t^{(k-1)} \mid \Omega_t(j)\right] \quad dj \tag{2.8}$$

It is sometimes useful to define the zero order expectation of θ_t as the actual value of the variable

$$\theta_t^{(0)} \equiv \theta_t \tag{2.9}$$

Full information rational expectations implies that the variable θ_t is common knowledge so that $\theta_t = \theta_t^{(k)} : k = 1, 2, ...$ for all periods t. We call a sequence of 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

$$\theta_t^{(0:k)} = \left[\begin{array}{cccc} \theta_t^{(0)} & \theta_t^{(1)} & ... & \theta_t^{(k)} \end{array}\right]' \tag{2.10}$$

2.2.1. *Expectations about future expectations.* In later sections, it will prove useful to also have a notation for average expectation held in period t of the average expectation held in period $t+1$ of the value of a variable in period $t+2$, and so on. For that purpose, we define the following notation. The first order expectation in period t of θ_{t+1} is defined as

$$\theta_{t+1|t}^{(1)} \equiv \int E\left[\theta_{t+1} \mid \Omega_t(j)\right] \quad dj \tag{2.11}$$

Similarly, the average expectation in period t of the average expectation in period $t+1$ of the θ_{t+2} is defined as

$$\theta_{t+2|t+1|t}^{(2)} \equiv \int E\left[\theta_{t+2|t+1}^{(1)} \mid \Omega_t(j)\right] \quad dj \tag{2.12}$$

Generalizing this notation

$$\theta^{(k)}_{t+k|\cdots|t} \equiv \int E\left[\theta^{(k-1)}_{t+k|\cdots|t+1} \mid \Omega_t(j)\right] \; dj \tag{2.13}$$

3. The Singleton Asset Pricing Model

This section presents 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. 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 this 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) process are perfectly and publicly observed with a two period lag. This allows him to derive a 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.

3.1. **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 $\overline{r}$. The wealth of trader j then evolves according to

$$w_{t+1}(j) = z_t(j)\left[p_{t+1} + c_{t+1}\right] - \left[z_t(j)p_t - w_t(j)\right](1+\overline{r}) \tag{3.1}$$

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

$$E\left[-e^{-\gamma w_{t+1}(j)} \mid I_t(j)\right] \tag{3.2}$$

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

$$c_t = \overline{c} + \psi c_{t-1} + u_t : u_t \sim N\left(0, \sigma_u^2\right) \tag{3.3}$$

Maximizing (3.2) subject to (3.1) yields agent j's optimal demand for the risky asset

$$z_t^d(j) = \frac{\left(E\left[p_{t+1} \mid I_t(j)\right] - (1+\overline{r})\, p_t\right) + (\overline{c} + \psi c_t)}{\gamma\delta} \tag{3.4}$$

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

$$z_t^s = \xi p_t + \theta_t + \epsilon_t : \epsilon_t \sim N\left(0, \sigma_\epsilon^2\right) \tag{3.5}$$

$$\theta_t = \rho\theta_{t-1} + v_t : v_t \sim N\left(0, \sigma_v^2\right) \tag{3.6}$$

Equating net demand and supply

$$\int z_t^d(j) = z_t^s \tag{3.7}$$

yields the equilibrium price

$$p_t = \lambda \left(\int E\left[p_{t+1} \mid I_t(j)\right] \quad dj \right) + \lambda \psi c_t - \delta\gamma\lambda \left[\theta_t + \epsilon_t\right] \tag{3.8}$$

where

$$\lambda \equiv \frac{1}{\xi\gamma\delta + (1+\overline{r})}. \tag{3.9}$$

For later reference, note that $0 < \lambda < (1+\overline{r})^{-1}$.

3.2. **Traders' Information Sets.** The basic structure of the model described above is identical to Model 8-12 in Singleton (1987). Where this paper differ from Singleton's 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

$$I_t^S = \{s_{t-T}(j), p_{t-T}, c_{t-T} : T \geq 0; v_{t-T}, \epsilon_{t-T} : T \geq 2\} \tag{3.10}$$

where

$$s_t(j) = \theta_t + \eta_t(j) : \eta_t(j) \sim N\left(0, \sigma_\eta^2\right) \ \forall \ j \tag{3.11}$$

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. The transitory supply shock ϵ_t thus serves the same purpose here as the noise traders do in Admati (1985). 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

$$I_t(j) = \{s_{t-T}(j), p_{t-T}, c_{t-T} : T \geq 0\} \tag{3.12}$$

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 (3.2) subject to the structural equations (3.3) - (3.6).

3.3. **The full information solution.** To solve the model we need to integrate out the average expectations term $\int E\left[p_{t+1} \mid I_t(j)\right] \quad dj$ from equation (3.8). Under full information, this could be done by iterating (3.8) forward

$$p_t = \sum_{k=1}^{\infty} \lambda^k E\left(c_{t+k} \mid c_t\right) - \delta\gamma\lambda \sum_{k=0}^{\infty} \lambda^k E\left(\theta_{t+k} \mid \theta_t\right) - \delta\gamma\lambda\epsilon_t. \tag{3.13}$$

Using the law of iterated expectations, (3.13) then simplifies to

$$p_t = \frac{\lambda\psi}{1-\lambda\psi}c_t - \frac{\delta\gamma\lambda}{1-\lambda\rho}\theta_t - \delta\gamma\lambda\epsilon_t \tag{3.14}$$

if $|\lambda\psi| < 1$ and $|\lambda\rho| < 1$.

3.4. **A complication.** With privately informed traders, we can still use forward substitution of the Euler equation (3.8). This yields the equilibrium price as a function of higher order expectations of future values of the persistent supply process θ_t

$$p_t = \frac{\lambda\psi}{1-\lambda\psi}c_t - \delta\gamma\lambda\epsilon_t - \delta\gamma\lambda\sum_{k=0}^{\infty}\lambda^k\theta^{(k)}_{t+k|\cdots|t} \tag{3.15}$$

where we used the notation for higher order expectations of future values of θ_t defined in Section 2.2.1. The current price of the asset thus depends on the average expectation in period t of θ_{t+1}, the average expectation in period t of the average expectation in period $t+1$ of θ_{t+2} and so on. As has been noted before, e.g. Allen, Morris and Shin (2006), average higher order expectations, i.e. expectations about other agent's expectations generally differ from average first order expectations and we cannot use the law of iterated expectations to integrate out the higher order expectations. To see why, note that the law of iterated expectations can loosely speaking be attributed to that agents do not believe that they have 'incorrect' expectations so that they do not expect to revise their own expectations in a particular direction. That is, first order expectations are martingales. The same is not true about expectations about other agents' 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. It is also this fact that makes the dynamics of models with private information interesting.

3.5. **The strategy.** The rest of the paper is devoted to finding a finite dimensional representation of the equilibrium price (3.15) of the form

$$p_{\overline{k},t} = \mathbf{a}_{\overline{k}}\theta_t^{(0:\overline{k})} - \delta\gamma\lambda\epsilon_t + \frac{\lambda\psi}{1-\lambda\psi}c_t \tag{3.16}$$

that is arbitrarily close to the solution to the equilibrium price (3.15) and where $\mathbf{a}'_{\overline{k}}$ and $\theta_t^{(0:\overline{k})}$ are finite dimensional vectors. We thus need to show that the discounted sum of higher order expectations of all future values of θ_t in (3.15) can be approximated by a linear function of a finite number of orders of expectations of the current value of θ_t so that

$$\mathbf{a}_{\overline{k}}\theta_t^{(0:\overline{k})} = \delta\gamma\lambda\sum_{k=0}^{\infty}\lambda^k\theta^{(k)}_{t+k|\cdots|t} \tag{3.17}$$

approximately holds. To do so, we will conjecture (and later verify) that there exists a law of motion for the hierarchy of higher order expectations of the current value of θ_t of the form

$$\theta_t^{(0:\bar{k})} = M\theta_{t-1}^{(0:\bar{k})} + Nw_t : w_t \sim N(0, I) \tag{3.18}$$

The solution will then consist of the equilibrium price (3.16) and the law of motion for the state (3.18).

The plan from here on is the following. First we will derive some properties of higher order expectations that must hold in any equilibrium. We then show how the price of the asset can be expressed as a function of the conjectured law of motion (3.18). This will give the problem enough of a structure so that it will be straight forward to show that there exists representation with a finite number of orders of expectations that can be made arbitrarily accurate. This results are quite general, in that they will hold under the same conditions that guarantee that a stable solution exists under full information, i.e. that $|\lambda\rho| < 1$.

4. Equilibrium properties of higher order expectations

It is possible to characterize some of the properties of higher order expectations using only that it is common knowledge that agents are rational, in the sense that the model is common knowledge. In this section we take the existence of an equilibrium as given, and present some properties of higher order expectations that must hold in any equilibrium.

4.1. **First order expectations.** We start by establishing some properties of first order expectations. This may seem pedantic, since properties of first order expectations are well known. However, this will lay the groundwork for recursively deriving similar, but more interesting, properties of higher order expectations. We start by defining a useful subspace of L^2.

Definition 3. *The (closed) subspace $\Omega_t(j) \equiv \overline{sp}\left\{s_{t-T}(j), p_{t-T}, c_{t-T} : T \geq 0\right\}$ is the space spanned by the history of variables observed by trader j at period t. Projections onto $\Omega_t(j)$ are denoted $\mathcal{P}_{t,j}$.*

From the projection theorem (e.g. Brockwell and Davis (2006)) we then know that there exist an element $\theta_t^{(1)}(j) \in L^2$ such that

$$\left\langle \theta_t - \theta_t^{(1)}(j), \omega_j \right\rangle = 0 \forall \omega_j \in \Omega_t(j) \tag{4.1}$$

that is, there exists a minimum variance expectation of θ_t conditional on trader j's information set. Given the linear structure of the model, past realizations of v_t, ϵ_t and $\eta_t(j)$ form an orthogonal basis for the subspace $\Omega_t(j)$. The conditional expectation $E\left[\theta_t \mid \Omega_t(j)\right]$ thus has a representation of the form

$$\theta_t^{(1)}(j) = A(L)v_t + B(L)\epsilon_t + C(L)\eta_t(j) \tag{4.2}$$

where by the symmetry of traders, the (potentially infinite order) lag polynomials $A(L)$, $B(L)$ and $C(L)$ are common across traders. Expectations will differ across traders only because of different realizations of the idiosyncratic noise shocks $\eta_t(j)$.

4.2. **The variance of first order expectations.** Here, the orthogonality property (4.1) and the representation (4.2) will be used to prove that the variance of average higher order expectations are bounded by the variance of lower order expectations. This result will later be used for the approximation results in Section 6 as well as for the existence results in Section 7. We start by showing that the variance of trader j's first order expectations of θ_t is bounded by the variance of the actual process θ_t.

Lemma 1. *The variance of trader j's expectation of θ_t is bounded by the variance of θ_t, i.e.*

$$E\left[\theta_t\right]^2 \geq E\left[\theta_t^{(1)}(j)\right]^2 \tag{4.3}$$

Proof. Define trader j's first order expectation error $\varepsilon_t^{(1)}(j)$ as

$$\theta_t - \theta_t^{(1)}(j) \equiv \varepsilon_t^{(1)}(j) \tag{4.4}$$

and rearrange

$$\theta_t \equiv \theta_t^{(1)}(j) + \varepsilon_t^{(1)}(j) \tag{4.5}$$

The variance of the l.h.s. is $E\left[\theta_t\right]^2$. By (4.1), the error $\varepsilon_t^{(1)}(j)$ is orthogonal to $\theta_t^{(1)}(j) \in \Omega_t(j)$ so the variance of the r.h.s. is simply the sum of the variances of the individual terms, which gives the equality

$$E\left[\theta_t\right]^2 = E\left[\theta_t^{(1)}(j)\right]^2 + E\left[\varepsilon_t^{(1)}(j)\right]^2 \tag{4.6}$$

The proof then follows from the fact that variances are non-negative

$$0 \leq E\left[\varepsilon_t^{(1)}(j)\right]^2 \tag{4.7}$$

so that

$$E\left[\theta_t\right]^2 \leq E\left[\theta_t^{(1)}(j)\right]^2 \tag{4.8}$$

□

According to the representation (4.2), trader j's expectation has both a common and idiosyncratic component. The fact that the idiosyncratic component is orthogonal to the common component allows us to prove our next result.

Lemma 2. *The variance of the average expectation of θ_t is bounded by the variance of θ_t, i.e.*

$$E\left[\theta_t\right]^2 \geq E\left[\theta_t^{(1)}\right]^2 \tag{4.9}$$

Proof. The representation (4.2) implies that the variance of trader j's first order expectations is the sum of the variances of the terms in the MA representation

$$E\left[\theta_t^{(1)}(j)\right]^2 = E\left[A(L)v_t\right]^2 + E\left[B(L)\epsilon_t\right]^2 + E\left[C(L)\eta_t(j)\right]^2 \tag{4.10}$$

Since $\int \eta_t(j)dj = 0 \forall t$ the average first order expectation is simply

$$\theta_t^{(1)} = A(L)v_t + B(L)\epsilon_t + \int C(L)\eta_t(j)dj \tag{4.11}$$

$$= A(L)v_t + B(L)\epsilon_t \tag{4.12}$$

with variance

$$E\left[\theta_t^{(1)}\right]^2 = E\left[A(L)v_t\right]^2 + E\left[B(L)\epsilon_t\right]^2 \tag{4.13}$$

Combining (4.8) and (4.10)

$$\begin{aligned} E\left[\theta_t\right]^2 &\geq E\left[\theta_t^{(1)}(j)\right]^2 && (4.14)\\ &= E\left[A(L)v_t\right]^2 + E\left[B(L)\epsilon_t\right]^2 + E\left[C(L)\eta_t(j)\right]^2 && (4.15)\\ &\geq E\left[A(L)v_t\right]^2 + E\left[B(L)\epsilon_t\right]^2 && (4.16)\\ &= E\left[\theta_t^{(1)}\right]^2 && (4.17) \end{aligned}$$

gives the desired result where the third line follows from the fact that $0 \leq E\left[C(L)\eta_t(j)\right]^2$ and the last equality is from (4.13) □

4.3. **Variance bounds for higher order expectations.**

Proposition 1. *The variance of higher order expectations of θ_t are bounded by the variance of lower order expectations, i.e.*

$$E\left[\theta_t^{(k)}\right]^2 \geq E\left[\theta_t^{(k+1)}\right]^2$$

Proof. To prove the proposition, replace the definition of trader j's first order expectations error $\varepsilon_t^{(1)}(j)$ in the proof of Lemma 1 with the definition of the k order expectation error

$$\theta_t^{(k-1)} - \theta_t^{(k)}(j) \equiv \varepsilon_t^{(k)}(j)$$

Noting that the k order error $\varepsilon_t^{(k)}(j)$ is orthogonal to $\theta_t^{(k)}(j) \in \Omega_t(j)$ allows for recursively establishing the proposition for $k = 2, 3, ...$by following the same steps as in the proofs of Lemma 1 and 2. □

It is straightforward to extend this result to higher order expectations of future values of θ_t.

Proposition 2. *The variance of higher order expectations of future expectations of θ_t are bounded by the variance of lower order expectations, i.e.*

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

Proof. To prove the proposition, replace the definition of trader j's first order expectations error $\varepsilon_t^{(1)}(j)$ in the proof of Lemma 1 with the definition of the k order future expectation error

$$\theta_{t+k|\cdots|t}^{(k-1)} - \theta_{t+k+1|\cdots|t}^{(k)}(j) \equiv \varepsilon_{t+k+1|\cdots|t}^{(k+1)}(j)$$

Again, since the k order error $\varepsilon_{t+k+1|\cdots|t}^{(k+1)}(j)$ is orthogonal to $\theta_{t+k+1|\cdots|t}^{(k)}(j) \in \Omega_t(j)$, the same recursive procedure as in Proposition 1 can be applied to establish the desired result for $k = 1, 2, 3, ...$ □

As noted above, these results will prove useful for the purpose of finding an accurate solution to the model, but they also illustrates well how the assumption of common knowledge of rational (i.e. model consistent) expectations allow us to derive properties of higher order expectations. It is the fact that first order expectations are formed rationally and that this is common knowledge that allow us to derive the variance bounds above. In the absence of common knowledge of model consistent expectations, we would have to make alternative assumptions about how traders in the model believe that other traders form expectations in order to determine how traders form second order expectations. Whether the variance bounds derived above would hold or not, would then depend on the properties of the second order beliefs about how other traders form expectations.

4.4. **Properties of the law of motion for higher order expectations.** Here, some useful equilibrium properties of the transition matrix M in the law of motion

$$\theta_t^{(0:\overline{k})} = M\theta_{t-1}^{(0:\overline{k})} + Nw_t : w_t \sim N(0, I) \tag{4.18}$$

are derived.

Proposition 3. *Common knowledge of rationality and $|\rho| < 1$ implies that max $|eig\,(M)| <$ 1*

Proof. The proof is by contradiction. Consider the case if $max\,|eigM| = 1$. This implies that at least one order $k \neq 0$ of expectation has a unit root and that at least one k' order expectations error $\theta_t^{(k')} - \theta_t^{(k'+1)}$ is increasing in t. Then for a large enough t we have that

$$E\left(\theta_t^{(k')} - \theta_t^{(k'+1)}\right)^2 > E\left[\theta_t^{(k')}\right]^2 \tag{4.19}$$

But then $\theta_t^{(k'+1)}$ cannot be the minimum variance estimate of $\theta_t^{(k')}$ since there then exist an element $\omega_0 \in \cap_{j\in(0,1)}\Omega_t(j)$ s.t.

$$E\left(\theta_t^{(k')} - \omega_0\right)^2 \leq E\left[\theta_t^{(k')}\right]^2 \tag{4.20}$$

e.g. the unconditional expectation $\omega_0 = E\left[\theta_t^{(k')}\right] = 0.$ $\square$

Proposition 3 simply says that if the true state is a stable process, then the law of motion for higher order expectations must also be a stable process. Proposition 3 could also be proved as a direct corollary of Proposition 1 since the fact that higher order expectation has a bounded variance also implies that $max\,|eig\,(M)| < 1$.

Proposition 4. *Each row of the matrix M in the law of motion (3.18) sum to ρ.*

Proof. First, note that common knowledge of rationality implies that if all order of expectations coincide so that

$$\theta_t = \theta_t^{(k)} : k = 1, 2, ... \tag{4.21}$$

then so must all orders of expectations about future values of θ_t

$$M^s \begin{bmatrix} \theta_t \\ \theta_t \\ \vdots \end{bmatrix} = \begin{bmatrix} \rho^s\theta_t \\ \rho^s\theta_t \\ \vdots \end{bmatrix} \tag{4.22}$$

That is, if all agents are believed to agree about the current state of the world, common knowledge of rationality then implies that there should also agreement about future expected states of the world. Equating elements in (4.22) for $s = 1$ immediately gives the desired result

$$\sum_{j=1}^{\infty} m_{i,j} = \rho : i = 1, 2, ... \tag{4.23}$$

where $m_{i,j}$ is the element in the i^{th} row and j^{th} column of M. □

In this section, we have derived some properties of higher order expectations that must hold in any equilibrium when it is common knowledge that traders form expectations rationally. Specifically, we showed that the variance of higher order expectations are bounded using orthogonality properties of expectation errors. While based on a simple insight, these results will later turn out to be very useful for both deriving a accurate solution as well as demonstrating that such a solution exists.

5. The equilibrium price

This section demonstrates how a simple matrix operator can be used to compute the equilibrium price for a given law of motion of the hierarchy of higher order expectations. The law of motion for the hierarchy of expectations is derived in the Section 7.

5.1. **An average higher order expectations operator.** To compute the higher order expectations we will use the linear operator $H : \mathbb{R}^{\infty} \to \mathbb{R}^{\infty}$ defined so that

$$\theta_t^{(k+1:\infty)} = H\theta_t^{(k:\infty)} \tag{5.1}$$

That is, H applied to a hierarchy of expectations move the hierarchy one step up in order of expectations. If the state of the economy is given by $\theta_t^{(0:\infty)}$ then the average expectations of the true state is given by $H\theta_t^{(0:\infty)}$ and the operator H thus annihilates the first element of a vector of higher order expectations. The operator is given by the matrix

$$H \equiv \left[\begin{array}{cc} \mathbf{0} & I_{\infty} \end{array} \right] \tag{5.2}$$

where I_{∞} is the identity matrix.[5]

5.2. **Equilibrium asset prices.** We can now derive an explicit expression for the equilibrium price of the asset. Given the conjectured law of motion (3.18) and the higher order expectations operator we can now compute the higher order expectations of the future values of θ_t in the forward iteration (3.15) of the price Euler equation (3.8).

The one step ahead average expectation of θ_t is simply given by first applying H to the complete hierarchy of expectation to get the average expectation of the state and then apply

[5]Allen, Morris and Shin (2006) defines an *average belief operator* $\overline{E} : \mathbb{R}^2 \to \mathbb{R}^2$. 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 (3.18) could be generated by a similar operator if θ_t was a non-persistent process.

M to form the average expectation of the state in the next period. The average expectation in period t of the value of persistent supply shock θ_t is then given by

$$\int E\left[\theta_{t+1} \mid I_t(j)\right] \quad dj = e_1' M H \theta_t^{(0:\infty)} \tag{5.3}$$

where e_1 is a vector with 1 in the first element and zeros elsewhere. Using similar reasoning, the expectation in period t of the average expectation in period $t+1$ of θ_{t+2} is then given by

$$\theta_{t+2|t+1|t}^{(2)} = e_1' \left(MH\right)^2 \theta_t^{(0:\infty)} \tag{5.4}$$

More generally, the k order expectation of θ_{t+k} is given by

$$\theta_{t+k|\cdots|t}^{(k)} = e_1' \left(MH\right)^k \theta_t^{(0:\infty)} \tag{5.5}$$

Substituting (5.5) into (3.15) gives the equilibrium price p_t as a function of the period t hierarchy of higher order expectations of θ_t as

$$p_t = -\delta\gamma\lambda \sum_{k=0}^{\infty} e_1' \left(\lambda MH\right)^k \theta_t^{(0:\infty)} - \delta\gamma\lambda\epsilon_t + \frac{\lambda\psi}{1-\lambda\psi} c_t \tag{5.6}$$

where we used that $\theta_t = e_1' \theta_t^{(0:\infty)}$.

6. A Finite Dimensional Approximation

In the previous sections, several properties of higher order expectations were derived that must hold in equilibrium. Though some of these properties may be interesting per se, here we show how they can be used to prove a practical result: The equilibrium characterized by an infinite number of orders of expectations can be approximated to an arbitrary accuracy by a finite dimensional system. That is, for practical purposes, we do not need to consider the complete hierarchy of expectation, but instead we can find a maximum (and finite) order of expectation that we need to consider, for any desired degree of accuracy. We denote this maximum order of expectation $\overline{k}$.

Two results are proved formally here. First, we show that the weight on higher order expectations tend to zero as the order of expectation increases. Secondly, we show that the approximation error variance tend to zero as we increase the maximum number of orders of expectations $\overline{k}$. These results are all derived using a similar technique. First, we define an infinite series indexed by the number of orders of expectations. We then show that the series converges. Since convergence of an infinite series implies that the individual elements in the sequence tend to zero (while the converse is not generally true), convergence of a series indexed by the maximum order of expectation considered implies that the accumulative effect of terms depending on orders of expectations higher than $\overline{k}$ tend to zero.

6.1. **The diminishing impact of higher order expectations.**

The solved model will deliver an expression for the equilibrium price p_t as a function of the current hierarchy of expectations about θ_t of the form

$$p_{\overline{k},t} = \mathbf{a}_{\overline{k}} \theta_t^{(0:\overline{k})} - \delta\gamma\lambda\epsilon_t + \frac{\lambda\psi}{1-\lambda\psi} c_t \tag{6.1}$$

where $\mathbf{a}_{\overline{k}}$ is a row vector with elements defined as

$$\mathbf{a}_{\overline{k}} \equiv \left[\begin{array}{cccc} a_0 & a_1 & \cdots & a_{\overline{k}} \end{array} \right]$$

From the price equation (3.8) we already know that $a_0 = -\delta\gamma\lambda$. The next proposition establishes that as $\overline{k} \to \infty$ the coefficient $a_{\overline{k}}$ tends to zero.

Proposition 5. *For $0 \leq |\lambda\rho| < 1$,there exists a finite number $\overline{k}$ such that*

$$|a_k + a_{k+1} + ... + a_{k+n-1} + a_{k+n}| < \varepsilon \tag{6.2}$$

for any $\varepsilon > 0$ and and $k > \overline{k}$.

Proof. The result is an immediate implication of the fact that $\{\Sigma a_k\}_{k=0}^{\infty}$ is a convergent series, which we first establish. To do so, we will use the fact that in the special case when $\theta_t = \theta_t^{(k)} = 1 \forall k$ the equilibrium price equals the sum of the elements in the row vector $\mathbf{a}$.

By Proposition 4 the rows of M sums to ρ and since $MH = \left[\begin{array}{cc} \mathbf{0} & M \end{array} \right]$ the rows of MH also sum to ρ so that

$$e_1' (MH)^k \times \mathbf{1}_\infty = \rho^k \times 1 : k = 0, 1, 2... \tag{6.3}$$

where $\mathbf{1}_\infty$ is an infinite dimensional vector of ones. Substituting (6.3) into the price equation (5.6) then gives

$$p_t = -\delta\gamma\lambda \sum_{k=0}^{\infty} (\lambda\rho)^k - \delta\gamma\lambda\epsilon_t + \frac{\lambda\psi}{1 - \lambda\psi} c_t \tag{6.4}$$

or that

$$\begin{aligned} -\delta\gamma\lambda \sum_{k=0}^{\infty} e_1' (\lambda MH)^k \times \mathbf{1}_\infty &= -\delta\gamma\lambda \sum_{k=0}^{\infty} (\lambda\rho)^k && (6.5) \\ &= -\frac{\delta\gamma\lambda}{1 - \lambda\rho} && (6.6) \end{aligned}$$

Intuitively, if all orders of expectations coincide then the price must equal the full information price (adjusted for the appropriate private information values of λ and δ which depend on the conditional variance). To see why this implies that the sum $\sum_{k=0}^{\infty} a_k$ converges, note that by definition, the l.h.s. of (6.5) equals the infinite sum of the elements of the row vector $\mathbf{a}_\infty$, i..e.

$$\begin{aligned} -\delta\gamma\lambda \sum_{k=0}^{\infty} e_1' (\lambda MH)^k \times \mathbf{1}_\infty &\equiv \mathbf{a}_\infty \times \mathbf{1}_\infty && (6.7) \\ &= \sum_{k=0}^{\infty} a_k && (6.8) \end{aligned}$$

Combining (6.6) and (6.8) establishes the limit of $\{\Sigma a_k\}_{k=0}^{\infty}$ as

$$\sum_{k=0}^{\infty} a_k = -\frac{\delta\gamma\lambda}{1 - \lambda\rho} \tag{6.9}$$

which is finite. Since the infinite series (6.9) converges, there exists a number $\overline{k}$ such that

$$|a_k + a_{k+1} + ... + a_{k+n-1} + a_{k+n}| < \varepsilon \text{ for all } n \geq 1 \text{ and } k > \overline{k} \tag{6.10}$$

(see for instance Ok 2007). □

Proposition 6 thus establishes that the coefficients a_k that multiply the k order expectation in the conjectured solution tend to zero as the order of expectation increases, and that this will hold under the same conditions that guarantee that a stable solution to the full information model exists, i.e. that $|\lambda\rho| < 1$.

6.2. **The variance of the approximation error.** Above, we demonstrated that the impact of expectations on the price tend to zero as the order of expectations increases. Combined with the fact that the variance of higher order expectations are bounded, one might conjecture that the variance of the contribution of the higher order expectation to the price also tend to zero as the order of expectation increases. Here, we will now demonstrate that this is the case but using a more direct approach that does not involve using the result of Proposition 6 above. Instead, we will define a particular convergent series (again indexed by $\overline{k}$) so that the remainder of the sum corresponds to the variance of the approximation error. Since the series converges, the remainder can be made arbitrarily small for large enough $\overline{k}$.To prove this result, we will need the following lemma.

Lemma 3. *The variance of the price p_t is finite.*

Proof. The proof uses that the higher order expectations about future expectations of future values of θ_t in the price equation (3.15) are discounted by $|\rho| < 1$ and have finite variances. The complete proof can be found in the Appendix. □

Definition 4. *The approximation error $\Delta_{\overline{k}}$ associated with considering only $\overline{k}$ orders of expectations is defined as*

$$\Delta_{\overline{k}} \equiv p_t - p_{\overline{k},t} \tag{6.11}$$

where

$$p_t = \mathbf{a}_{\overline{k}}\theta_t^{(0:\infty)} - \delta\gamma\lambda\epsilon_t + \frac{\lambda\psi}{1-\lambda\psi}c_t \tag{6.12}$$

and

$$p_{\overline{k},t} = \mathbf{a}_{\overline{k}}\theta_t^{(0:\overline{k})} - \delta\gamma\lambda\epsilon_t + \frac{\lambda\psi}{1-\lambda\psi}c_t \tag{6.13}$$

so that

$$\Delta_{\overline{k}} = \left(\mathbf{a} - \begin{bmatrix} \mathbf{a}_{\overline{k}} & 0 \end{bmatrix}\right)\theta_t^{(0:\infty)} \tag{6.14}$$

Proposition 6. *The variance of the approximation error $\Delta_{\overline{k}}$ tends to zero as $\overline{k}$ tends to infinity.*

Proof. First, define the sequence$\{z_{\overline{k}}\}$

$$\{z_{\overline{k}}\} = \left\{ \sum_{j=0}^{\overline{k}} \sum_{i=0}^{\overline{k}} a_i a_j cov\left[\theta_t^{(i)}, \theta_t^{(j)}\right] \right\} \tag{6.15}$$

and denote its limit z_∞

$$z_\infty \equiv \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} a_i a_j cov\left[\theta_t^{(i)}, \theta_t^{(j)}\right] \tag{6.16}$$

If the limit exist we know that there exists a $\overline{k}$ such that

$$|z_{\infty} - z_{\overline{k}}| < \varepsilon : \varepsilon > 0 \tag{6.17}$$

From (6.14) the variance of the approximation error is given by

$$\begin{aligned} E\left(\Delta_{\overline{k}}\right)^2 &= \left(\mathbf{a} - \left[\begin{array}{cc} \mathbf{a}_{\overline{k}} & 0 \end{array}\right]\right) E\left(\theta_t^{(0:\infty)}\theta_t^{\prime(0:\infty)}\right)\left(\mathbf{a} - \left[\begin{array}{cc} \mathbf{a}_{\overline{k}} & 0 \end{array}\right]\right)' \qquad (6.18) \\ &= \sum_{j=\overline{k}+1}^{\infty}\sum_{i=\overline{k}+1}^{\infty} a_i a_j cov\left[\theta_t^{(i)}, \theta_t^{(j)}\right] \qquad (6.19) \end{aligned}$$

which equals the difference between $z_{\overline{k}}$ and its limit so that

$$z_{\infty} - z_{\overline{k}} = \sum_{j=\overline{k}+1}^{\infty}\sum_{i=\overline{k}+1}^{\infty} a_i a_j cov\left[\theta_t^{(i)}, \theta_t^{(j)}\right] \tag{6.20}$$

To prove that the approximation error tend to zero as $\overline{k}$ increases it is thus sufficient to show that the sequence $\{z_{\overline{k}}\}$ converges, i.e. that the limit z_{∞} exists and is finite. Taking variances of both sides of (6.1) as $\overline{k} \to \infty$, we find that the variance of the equilibrium price is given by

$$\begin{aligned} E\left[p_t\right]^2 &= \sum_{j=0}^{\infty}\sum_{i=0}^{\infty} a_i a_j cov\left[\theta_t^{(i)}, \theta_t^{(j)}\right] \qquad (6.21) \\ &\quad + 2\delta\gamma\lambda \sum_{j=0}^{\infty} a_j cov\left[\theta_t^{(j)}, \epsilon_t\right] \\ &\quad + \left(\delta\gamma\lambda\right)^2 \sigma_{\epsilon}^2 + \left(\frac{\lambda\psi}{1-\lambda\psi}\right)^2 \sigma_c^2 \end{aligned}$$

By Lemma 3 we know that the variance of the price is finite so that $E\left[p_t\right]^2 < \infty$. This in turn implies that each term on the right hand side of (6.21) also must be finite. Since the first term on the right hand side of (6.21) equals the limit z_{∞} we know that $\{z_{\overline{k}}\}$ converges. There thus exists a finite $\overline{k}$ such that

$$\begin{aligned} E\left(\Delta_{\overline{k}}\right)^2 &= |z_{\infty} - z_{\overline{k}}| \qquad (6.22) \\ &< \varepsilon : 0 < \varepsilon \qquad (6.23) \end{aligned}$$

where the first line follows from (6.18) - (6.20) and the second line follows from (6.17) and the fact that $\{z_{\overline{k}}\}$ converges. □

In this and previous sections we have derived properties of higher order expectations that must hold in any equilibrium. We now turn to how such an equilibrium can be found in practice.

7. A Solution Algorithm

Solving the model implies finding the matrices M and N in the law of motion for the hierarchy of expectations, the row vector $\mathbf{a}$ in the price equation and the conditional variance δ. This section presents an iterative solution algorithm and shows that a solution exists.

There are three basics steps in each iteration (indexed by s) of the algorithm (i) For given M_s and δ_s, find the row vector $\mathbf{a}_{\overline{k},s}$ that maps the hierarchy of expectations into the price. (ii) For given values of $M_s, N_s, \mathbf{a}_{\overline{k},s}$ and δ_s find the new matrices M_{s+1} and N_{s+1} of the law of motion of the hierarchy. (iii) For given M_{s+1}, $N_{s+1}, \mathbf{a}_{\overline{k},s}$ and δ_s find the new conditional variance δ_{s+1}. These steps are described in detail below. This is followed by a proposition that uses Brouwer's fixed point theorem to prove that a fixed point of the iteration on (i) - (iii) exists under general conditions. From now on, all derivations pertain to a finite dimensional approximation of the equilibrium. To simplify notation, subscripts indicating that vectors and matrices are finite dimensional have been suppressed at instances where this will not cause confusion.

7.1. **Step 1: Computing the price.**

The first step is to find the price of the asset as a function of the contemporaneous expectation hierarchy of the supply disturbance $\theta_t^{(0:\overline{k})}$ for a given law of motion of the hierarchy and a given conditional variance of $(p_t + c_t)$, δ. That is, we want to find $\mathbf{a}_{\overline{k}}$ in

$$p_t = \mathbf{a}_{\overline{k}}\theta_t^{(0:\overline{k})} - \delta\gamma\lambda\epsilon_t + \frac{\lambda\psi}{1-\lambda\psi}c_t \tag{7.1}$$

as a function of M and δ. This simply entails computing the infinite sum in (5.6) so that

$$\mathbf{a}_{\overline{k}}\theta_t^{(0:\overline{k})} = -\delta\gamma\lambda\sum_{k=0}^{\infty} e_1'\left(\lambda MH\right)^k\theta_t^{(0:\overline{k})} \tag{7.2}$$

For a finite $\overline{k}$ and given M_s and δ_s implies that $\mathbf{a}_{\overline{k},s}$ is given by

$$\mathbf{a}_{\overline{k},s} = -\delta_s\gamma\lambda_s e_1'\left(I - \lambda_s M_s H\right)^{-1} \tag{7.3}$$

where we know that $\lambda_s M_s H$ is stable matrix (since the variance of p_t is finite). The price function thus resembles a standard discounted expected sum of future fundamentals, but where the coefficient matrix M from the true law of motion is replaced with MH. The relationship between M and MH is thus similar to that of physical and risk neutral dynamics in the finance literature. Of course, the interpretation is different: Here, the asset is priced as if the hierarchy was observed perfectly but followed a law of motion with transition dynamics determined by MH.

7.2. **Step 2: The dynamics of the expectation hierarchy.**

This section shows how to find the law of motion for the hierarchy of expectations (3.18) for a given vector $\mathbf{a}_{\overline{k}}$ 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, so to compute the $s+1$ iteration M_{s+1} and N_{s+1}, "old" values M_s and N_s are also needed.

The state consists of the actual supply disturbance θ_t and the hierarchy of expectations of the supply disturbance $\theta_t^{(0:\overline{k})}$, so the law of motion of the state is determined by the actual supply process (3.6) and the law of motion of the higher order estimates. The Kalman filter thus plays a dual role: it both determines traders' estimate of the state as well as the law of motion of the very same state that the traders are estimating. To find the updated law of motion M_{s+1} and N_{s+1} we compute first find the recursive updating equation for trader j's estimate of the hierarchy of higher order expectations conditional on the the previous law of motion M_s and N_s. We then take averages of this recursive updating equation to find the new law of motion for the hierarchy of average expectations M_{s+1} and N_{s+1}.

7.2.1. *Trader j's estimate of the hierarchy.* For given values M_s and N_s, the conjectured law of motion for the hierarchy (3.18) and trader j's information set (3.12) can be written as a state space system of the form

$$\begin{aligned} \theta_t^{(0:\infty)} &= M_s\theta_{t-1}^{(0:\overline{k})} + N_s w_t && (7.4) \\ S_t(j) &= L_s\theta_t^{(0:\overline{k})} + Q_s c_t + \begin{bmatrix} R_{1s} & R_2 \end{bmatrix} \begin{bmatrix} w_t \\ w_t(j) \end{bmatrix} : \begin{bmatrix} w_t \\ w_t(j) \end{bmatrix} \sim N(0,I) && (7.5) \end{aligned}$$

where w_t is a vector of aggregate shocks and $w_t(j)$ is the idiosyncratic shock to trader j's private signal of θ_t. The following definitions were also used

$$\begin{aligned} S_t(j) &= \begin{bmatrix} s_t(j) \\ p_t \end{bmatrix}, \ L_s = \begin{bmatrix} e_1' \\ \mathbf{a}_{\overline{k},s} \end{bmatrix}, \ Q = \begin{bmatrix} 0 \\ \frac{\lambda_s\psi}{1-\lambda_s\psi} \end{bmatrix} \\ R_{1s} &= \begin{bmatrix} 0 & 0 \\ 0 & -\delta_s\gamma\lambda_s\sigma_\epsilon \end{bmatrix}, \ R_2 = \begin{bmatrix} \sigma_\eta \\ 0 \end{bmatrix} \end{aligned}$$

where subscript s indicates that a matrix may be changing at each iterative step in the algorithm. Trader j estimates the hierarchy of contemporaneous expectations recursively, using the Kalman filter updating equation

$$\theta_t^{(1:\overline{k})}(j) = M_s\theta_{t-1}^{(1:\overline{k})}(j) + K_s\left[S_t(j) - L_sM_s\theta_{t-1}^{(1:\overline{k})}(j) - Qc_t\right] \tag{7.6}$$

The Kalman gain K_s is given by

$$\begin{aligned} K_s &= \left(P_sL_s' + N_sR_{s1}'\right)\left(L_sP_sL_s' + R_sR_s'\right)^{-1} && (7.7) \\ P_s &= M_s\left(P_s - \left(P_sL_s' + N_sR_{1s}'\right)\left(L_sP_sL_s' + R_sR_s'\right)^{-1}\left(P_sL_s' + N_sR_{1s}'\right)'\right)M_s' && (7.8) \\ &\quad + N_sN_s' \end{aligned}$$

and where $R_s = \begin{bmatrix} R_{1s} & R_2 \end{bmatrix}$. Note that the s subscript on K and P denotes the step in the solution algorithm, not the time period. For given M_s and N_s, we compute the time invariant, or steady state, Kalman gain.

7.2.2. *The average expectation hierarchy.* We want to find the conjectured vector AR(1) law of motion (3.18) 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 (7.6) across traders and express all remaining terms as functions of the lagged expectation hierarchy $\theta_{t-1}^{(0:\overline{k})}$ and the aggregate shocks w_t. Use the definition of the private signal $s_t(j)$

(3.11), the price equation (3.15) and that $\int R_2 w_t(j)\ dj = \mathbf{0}$ to write the average signal $S_t \equiv \int S(j)\ dj$ as

$$S_t = L_s M_s \theta_{t-1}^{(0:\overline{k})} + L_s N_s w_t + R_{1s} w_t \tag{7.9}$$

Substituting the average signal (7.9) into the updating equation (7.6) gives the law of motion of the average of traders' estimate of the state

$$\theta_t^{(1:\overline{k})} = \left(I - K_s L_s\right) M_s \theta_{t-1}^{(1:\overline{k})} + K_s L_s M_s \theta_{t-1}^{(0:\overline{k})} + \left(K_s L_s N_s + K_s R_{1s}\right) w_t \tag{7.10}$$

The final step to get the conjectured form (3.18) is to collect terms and append the actual supply disturbance process

$$\theta_t = \rho \theta_{t-1} + v_t \tag{7.11}$$

to the updating equation (7.10) to get

$$\begin{aligned} \begin{bmatrix} \theta_t \\ \theta_t^{(1:\overline{k})} \end{bmatrix} &= \begin{bmatrix} \rho & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \theta_t \\ \theta_t^{(1:\overline{k})} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ K_s L_s M_s \end{bmatrix} \begin{bmatrix} \theta_t \\ \theta_t^{(1:\overline{k})} \end{bmatrix} \\ &\quad + \begin{bmatrix} 0 & \mathbf{0} \\ \mathbf{0} & \left(I - K_s L_s\right) M_s \end{bmatrix} \begin{bmatrix} \theta_t \\ \theta_t^{(1:\overline{k})} \end{bmatrix} + \begin{bmatrix} \sigma_v e_1' \\ \left(K_s L_s N_s + K R_1\right) \end{bmatrix} w_t \end{aligned} \tag{7.12}$$

where the last row and/or columns of the matrices have been cropped to make the matrices conformable (i.e. implementing the approximation that expectations of order $k > \overline{k}$ are redundant, and therefore setting $\theta_t^{(k)} = 0 : k > \overline{k}$). Equating coefficients in (7.12) and (3.18) then gives the updated matrices M_{s+1} and N_{s+1}

$$M_{s+1} = \begin{bmatrix} \rho & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ K_s L_s M_s \end{bmatrix} + \begin{bmatrix} 0 & \mathbf{0} \\ \mathbf{0} & \left(I - K_s L_s\right) M_s \end{bmatrix} \tag{7.13}$$

$$N_{s+1} = \begin{bmatrix} \sigma_v e_1' \\ \left(K_s L_s N_s + K_s R_{1s}\right) \end{bmatrix} \tag{7.14}$$

7.3. **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

$$\delta_{s+1} = \widehat{\mathbf{a}}_{\overline{k},s} \widehat{P}_s \widehat{\mathbf{a}}_{\overline{k},s}' + \left[1 + 2\frac{\lambda_s \psi}{1 - \lambda_s \psi} + \left(\frac{\lambda_s \psi}{1 - \lambda_s \psi}\right)^2\right] \sigma_u^2 \tag{7.15}$$

$$\widehat{\mathbf{a}}_{\overline{k},s} = \begin{bmatrix} \mathbf{a}_{\overline{k},s} & -\delta_s \gamma \lambda_s \end{bmatrix} \tag{7.16}$$

where $\widehat{P}_s$ 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 $\widehat{P}_s$ are given in the Appendix.

7.4. **The existence of a fixed point.** Solving the model implies finding a fixed point of equations (7.3),(7.7),(7.8),(7.13),(7.14) and (7.15). We now prove the existence of such a fixed point using Brouwer's fixed point theorem, which we first restate.[6]

[6] The relevant version of Brouwer's fixed point theorem is for compact subsets of $\mathbb{R}^n$ and is also known as *Kakutanis' fixed point theorem.*

Lemma 4. *(Brouwer fixed point theorem) Every continuous map from a convex compact set into itself has a fixed point.*

We thus need to show that iterating on Step 1 - 3 above is indeed a map from a convex compact set into itself. In order to do so, we will redefine the mapping $\{M_s, N_s, a_{\overline{k}}, \delta_s, K_s, L_s\} \rightarrow \left\{M_{s+1}, N_{s+1}, a_{\overline{k},s+1}, \delta_{s+1}, K_{s+1}, L_{s+1}\right\}$ described by Step 1 - 3 above in two ways.

First, note that for given M_{s+1}, N_{s+1} and δ_{s+1} we can find $a_{\overline{k},s+1}, K_{s+1}$ and L_{s+1} that do not depend on $a_{\overline{k},s}, K_s$ and L_s. It is thus sufficient to find a fixed point of the mapping $\{M_s, N_s, \delta_s\} \rightarrow \{M_{s+1}, N_{s+1}, \delta_{s+1}\}$.

Secondly, we will redefine the matrix M to an equivalent function of two covariance matrices with known properties, i.e. matrices that belong to the convex compact set $\mathcal{S}$, which we now define.

Definition 5. *The set $\mathcal{S}$ is the set of $\left(\overline{k}+1\right) \times \left(\overline{k}+1\right)$ matrices Σ matrix with i^{th}row, j^{th} column element $\sigma_{i,j}$ such that $|\sigma_{i,j}| \leq E\left(\theta_t^2\right) : i, j = 1, 2, ..., \overline{k}+1$*

Lemma 5. *The matrix M_s can equivalently be expressed as a function of the matrices Σ_s and $\Sigma_{+1,s}$ defined as*

$$\Sigma_s \equiv cov_s\left(\theta_t^{(0:\overline{k})}, \theta_t^{(0:\overline{k})}\right) \tag{7.17}$$

$$\Sigma_{+1,s} = cov_s\left(\theta_{t+1}^{(0:\overline{k})}, \theta_t^{(0:\overline{k})}\right) \tag{7.18}$$

where cov_s denotes the covariance conditional on the law of motion described by M_s and N_s.

Proof. From the projection theorem we know that

$$E\left[\theta_{t+1}^{(0:\overline{k})} \mid \theta_t^{(0:\overline{k})}\right] = \Sigma_{+1,s}\Sigma_s^{-1}\theta_t^{(0:\overline{k})} \tag{7.19}$$

i.e. M_s is given by $\Sigma_{+1,s}\Sigma_s^{-1}$. □

Lemma 6. *The covariance matrices Σ_s and Σ_s^{+1} belong to $\mathcal{S}$, that is, that all elements of Σ_s and Σ_s^{+1} lie in the closed interval $\left[-E\left(\theta_t^2\right), E\left(\theta_t^2\right)\right]$.*

Proof. The mapping $\{\Sigma_s^{-1}, \Sigma_{+1,s}, N_s, \delta_s\} \rightarrow \left\{\Sigma_{s+1}^{-1}, \Sigma_{+1,s+1}, N_{s+1}, \delta_{s+1}\right\}$ defines a new law of motion for the hierarchy $\theta_{t,s+1}^{(1:\overline{k})}$ that is the optimal estimate of the hierarchy $\theta_{t,s}^{(0:\overline{k})}$ if $\theta_{t,s}^{(0:\overline{k})}$ is governed by the law of motion $\{M_s, N_s\}$. We know that the variance of an optimal estimate cannot be larger than the variance of the object being estimated, so the inequality

$$E\left(\theta_{t,s}^{(k)}\right)^2 \geq E\left(\theta_{t,s+1}^{(k+1)}\right)^2 \tag{7.20}$$

must hold for each iteration s. Starting from an initial guess of M_0 and a N_0 (for instance the M and N implied by the full information solution) such that

$$E\left(\theta_t^2\right) \geq E\left(\theta_{t,0}^{(k)}\right)^2 : k = 1, 2, ... \tag{7.21}$$

ensures that

$$\Sigma_s \in \mathcal{S}, \Sigma_{+1,s} \in \mathcal{S} : s = 0, 1, 2, ... \tag{7.22}$$

The last result follows from the Cauchy-Schwarz inequality (in L^2 with the square root norm)

$$|E(XY)| \leq \sqrt{E\left(X\right)^2}\sqrt{E\left(Y\right)^2} \tag{7.23}$$

so that

$$\begin{aligned} \left|cov(\theta_{t+s}^{(k)}, \theta_t^{(k+l)})\right| &\leq \max\left\{E\left(\theta_t^{(k)}\right)^2, E\left(\theta_{t+s}^{(k+l)}\right)^2\right\} : k,l,s = 0,1,2... &(7.24)\\ &\leq E\left(\theta_t^2\right) &(7.25)\end{aligned}$$

i.e. all elements of Σ_s and Σ_s^{+1} must lie in the closed interval $[-E\left(\theta_t^2\right), E\left(\theta_t^2\right)]$ which proves the lemma. □

Definition 6. *The set $\mathcal{N}$ is the set such that if $N \in \mathcal{N}$ then N is $\left(\overline{k}+1\right) \times 2$ matrix with elements $|n_{i,j}| \leq \sqrt{E\left(\theta_t^2\right)} : i = 1,2,...,\overline{k}+1$ and $j = 1,2$.*

Lemma 7. *The matrices $N_s : s = 0,1,2...$ in the iteration described by belong to $\mathcal{N}$.*

Proof. $M_s \Sigma_s M_s'$ in the Riccati equation for Σ_s

$$\Sigma_s = M_s \Sigma_s M_s' + N_s N_s' \tag{7.26}$$

is a positive semi-definite matrix. Since $\Sigma_s \in \mathcal{S}$ for each iteration s, each element $n_{s,ij}$ of $N_s$ust lie in the interval $[-\sqrt{E\left(\theta_t^2\right)}, \sqrt{E\left(\theta_t^2\right)}]$ since the i^{th} element on the diagonal of $N_s N_s'$ is given by

$$\left(N_s N_s'\right)_{ii} = \sum_{j=1}^{\overline{k}+1} n_{ij} n_{ij} \tag{7.27}$$

The results then follows immediately the fact that the diagonal elements of Σ_s are non-negative for positive semi-definite matrices. □

Definition 7. *The set $\mathcal{D}$ is the closed interval $[0, \overline{\sigma}_{pc}^2]$ on $\mathbb{R}$ where $\overline{\sigma}_{pc}^2$ is the upper bound of the unconditional variance of $p_t + c_t$.*

It follows immediately that $\delta_s \in \mathcal{D}$ since the conditional variance is bounded by the unconditional variance

$$E\left(p_t + c_t\right)^2 \leq E\left(p_t\right)^2 + E\left(c_t\right)^2 + 2\max\left\{E\left(p_t\right)^2, E\left(c_t\right)^2\right\} \tag{7.28}$$

where the inequality follows from the Cauchy-Schwarz inequality and that both the price and coupon payments have finite variances.

Proposition 7. *The set $\mathcal{Z} \equiv \mathcal{S} \times \mathcal{S} \times \mathcal{N} \times \mathcal{D}$ is convex and compact and a fixed point described the iteration $\{\Sigma_s^{-1}, \Sigma_{+1,s}, N_s, \delta_s\} \to \left\{\Sigma_{s+1}^{-1}, \Sigma_{+1,s+1}, N_{s+1}, \delta_{s+1}\right\}$ exists.*

Proof. For finite dimensional sets, compactness is equivalent to a set being closed and bounded, so compactness follows directly from the definitions of $\mathcal{S}$,$\mathcal{N}$.and $\mathcal{D}$. Convexity follows from that if $|x| \leq c$ and $|z| \leq c$ then $\alpha|x| + (1-\alpha)|z| \leq c$. The existence of a fixed point follows from Lemma 4 - 7. □

In this section we have shown how a solution to the model can be found for a finite $\overline{k}$. In practice, we need to choose a maximum order of expectations to include in the representation of the model. In the next section shows how this can be done by ensuring that the impulse responses of prices changes have converged and thus remian unchanged as the maximum order $\overline{k}$ is increased further.

8. Properties of the Solved Model

In this section, the properties of the model is explored in more detail. First, we give an example of how the private information changes the price responses to supply shocks in contrast to when the model is solved under full or imperfect but common information. Secondly, we demonstrate how the representation of the equilibrium dynamics of the model can be used to compute two different types of dispersion of expectations: (i) Dispersion of expected returns across traders, and dispersion across orders of expectations. Both of these types of dispersion may be of interest to quantify and it is straight forward to compute either for a given parameterization of the model.

8.1. **Dynamics.** One question of interest is how private information affect the responses of the asset's price to innovations to the supply of assets. In the row of Figure 1 below, we have plotted the impulse response function of the price of the asset to an innovation to the persistent component of supply (left column) and a to a transitory shock (right column). For comparison, we have also plotted the impulse response to the same innovation under the alternative assumptions of full information, i.e. the state is observed perfectly by all traders, and imperfect but common information, i.e. it is common knowledge that all traders observe the same noisy signal about θ_t. The parameters $\left\{\overline{k}, \gamma, \xi, \psi, \rho, \overline{r}, \sigma_u^2, \sigma_v^2, \sigma_\epsilon^2, \sigma_{\eta i}^2, \right\}$ was set to $\{15, 1, 1.5, 0.5, 0.9, 0.01, 0.01, .1, 0.001, 1\}$. For the imperfect but common information case we set the variance of the noise in the common signal to the same as the idiosyncratic noise variance in the private information case.

The impulse of this parameterization is displayed in Figure 1 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, with the trough appearing later with private information than with imperfect but common information. 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]

That private information can be a strong force of inertia in endogenous variables has been noted before, e.g. Woodford (2002), Nimark (2008), Graham and Wright (2010) and Angeletos and L'ao (2009). As first pointed out by Woodford (2002) in a setting where agents faced a dynamic filtering problem (but with static choices), it is the fact that higher order expectations respond much more sluggishly to a shock than lower order expectations

[7] Indeed, Singleton found that what mattered most in his model was that agents had *imperfect* information, rather than *private* information per se. An earlier version of this paper demonstrated that this was due to the large variances of the innovations in the supply process in Singleton's calibration. Since the discount factor λ depends on the conditional variance of returns δ, absolute (and not only relative) variances of shocks matter. Larger variances imply faster discounting of the higher order expectation in ().

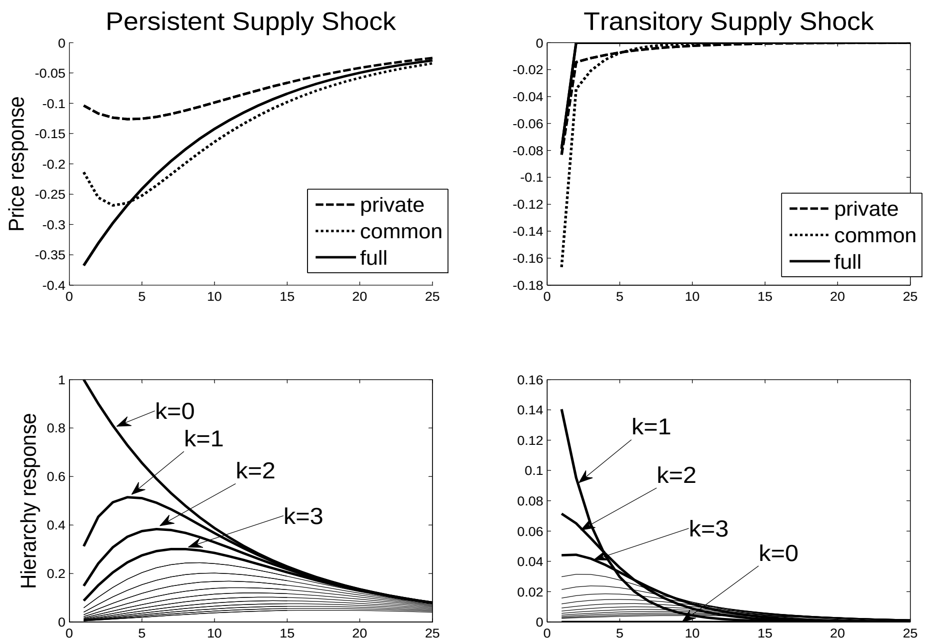

FIGURE 1. Impulse responses of p_t (top row) and $\theta_t^{(0:50)}$ (bottom row) to innovation to persistent (left column) and transitory (right column) component of supply.

that causes the inertial response of the endogenous variable. This is illustrated in the bottom row of Figure 1 the responses of the hierarchy of expectations to the two shocks are plotted. Average first order expectations ($k = 1$) respond stronger than higher order expectations in the impact period to both persistent and transitory shocks. That higher order expectation respond less than lower order expectations is intuitive. First order expectation respond less than the true shock on impact since some of the actual supply shock will be attributed to the transitory shock. Since traders know that first order expectations on average respond less than the actual shock, second order expectations must respond less than first order shock. This argument can then be applied recursively to understand why a $k+1$ order expectation responds less than a k order expectation in the impact period.

After a transitory shock ϵ_t and for $k \geq 1$, lower order expectations of θ_t also respond more strongly on impact. However, lower order expectations respond quicker to the higher than expectated asset prices that follows the impact period and converge faster towards the true shock (zero) than lower order expectations. The fact that this convergence of the (higher order) expectations about θ_t towards zero is not immediate that introduces some persistence of the price response even to a transitory shock.

8.2. **Cross-sectional dispersion of expectations.** Survey evidence suggest that market participants may have dispersed expectations about future economic outcomes, e.g. Swanson (2006) and Mankiw, Reis and Wolfers (2003). Private information is one way of introducing such dispersion in a model and there are at least two reason why this may be of interest. First, we may want to use quantitative information from for instance surveys to calibrate the parameters of a model to match the measured dispersion of expectations. Secondly, and as in Nimark (2010), computing the implied dispersion for a model with parameters estimated using only aggregate variables, one can gauge the plausibility of the model by judging whether the dispersion of expectations implied by the parameters that generate the best fit to aggregate variables is realistic. In the framework presented here, it is straight forward to compute the cross-sectional dispersion of expectations.

The idiosyncratic noise shocks $\eta_t(j)$ are white noise processes that are orthogonal to the aggregate shocks v_t and ϵ_t. This implies that the cross-sectional variance of expectations is equal to the unconditional variance of trader j's expectations that is due to idiosyncratic shocks, which can be computed by finding the variance of the estimates in trader j's updating equation (7.6), but with the aggregate shocks v_t and ϵ_t "switched off". The covariance Σ_j of trader j's expectations due to idiosyncratic shocks is then given by

$$\Sigma_j = (I - KL)\, M\Sigma_j M' \,(I - KL)' + KR_2R_2'K'. \tag{8.1}$$

where Σ_j is defined as

$$\Sigma_j = E\left(\theta_t^{(1:\overline{k})}(j) - \int \theta_t^{(1:\overline{k})}(j')dj'\right)\left(\theta_t^{(1:\overline{k})}(j) - \int \theta_t^{(1:\overline{k})}(j')dj'\right)' \tag{8.2}$$

The variance of price expectations s periods ahead is then given by

$$E\left(E\left[p_{t+s} \mid \Omega_t(j)\right] - \int E\left[p_{t+s} \mid \Omega_t(j')\right]dj'\right)^2 = \mathbf{a}_{\overline{k}} M^s \Sigma_j \left(\mathbf{a}_{\overline{k}} M^s\right)' \tag{8.3}$$

since

$$E\left[p_{t+s} \mid \Omega_t(j)\right] = \mathbf{a}_{\overline{k}} M^s E\left[\theta_t^{(0:\overline{k})} \mid \Omega_t(j)\right] \tag{8.4}$$

The cross sectional variance of expectations will generally depend on all the parameters of the model, but some have a more direct influence on the dispersion than others. For instance, Figure 2 illustrates how the cross sectional variance of one period ahead price expectations depends on the variance σ_η^2 of the idiosyncratic noise shock $\eta_t(j)$ (left panel) and the variance σ_ϵ^2 of the transitory demand shock ϵ_t (right panel). Both graphs start at the origin, i.e. if either the variance of the idiosyncratic noise shocks or the transitory supply shocks are zero, there is no dispersion of expectations. Of course, if there are no idiosyncratic noise shocks, there is no private information since all traders observe θ_t directly and without error. Similarly, if there are no transitory supply shocks, traders can infer θ_t perfectly from observing the price p_t and again, there is no private information in equilibrium. This result is reminiscent of the result in Walker (2007) who uses a version of Singleton's model to show that if one of the supply shocks is observed directly, equilibrium prices reveal the other shock perfectly and there is then no role for private information.

While the limit case of zeros variance is similar for the two shocks in the figure, the change in dispersion as the variance is increased is quite different. If the variance of the idiosyncratic

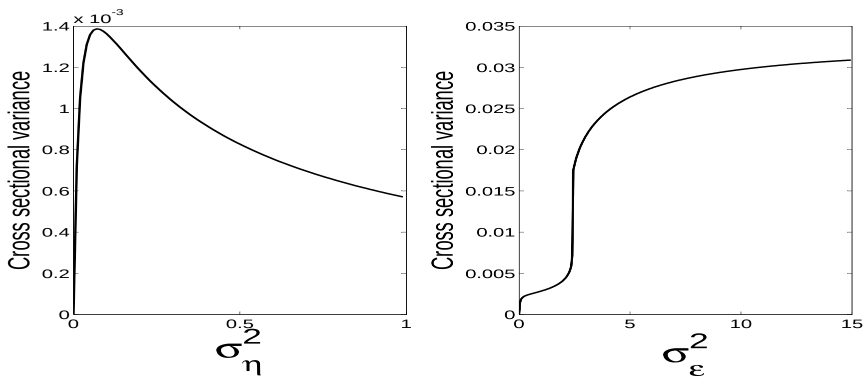

Figure 2. Cross sectional variance of one period ahead price expectation.

shock is increased, dispersion of expectations first increase as traders observe private signals with increased dispersion. At some point though, the variance of the idiosyncratic noise shocks become large enough so that the weight on the private signal decreases faster than the variance of the noise increases. This explains the hump shape dependence of cross-sectional dispersion on the variance σ_η^2.

We do not see the hump shape in the right panel. The reason is that when the variance of the transitory shock is increased, prices become more noise as signals about θ_t and traders tend to put more weight on their private signal $s_t(j)$. Where the graph flattens out, the price is so noisy that traders do not put any weight on it when estimating θ_t.

8.3. **Dispersion across orders of expectations.** The framework presented here can also be used to compute a different type of dispersion of expectations, that is, when different orders of expectations do not coincide. Unlike the cross-sectional dispersion, dispersion across orders of expectations vary over time and gives rise to new dynamics. Indeed, it is the fact that there is a divergence between orders of expectations that makes models with private information to display different dynamics, as the full information rational expectations equilibrium can be thought of as a special case when all orders of expectations coincide in every period so that $\theta_t = \theta_t^{(k)} : k = 1, 2, ...$for all t. As with cross-sections dispersion, the amount of dispersion across orders of expectations depends generally on all the parameters in the model but the variance of the transitory supply shock and the variance of the idiosyncratic noise shock again play a more direct role. Figure 3 illustrates how the response of the hierarchy of expectations of θ_t from order zero to 50 to a unit innovation in θ_t depend on the variance of the transitory supply shock ϵ_t. (Apart from σ_ϵ^2, the parameterization is the same as that used for Figure 1.) The thick solid line is the response of the actual shock, or $\theta_t^{(0)}$, the dashed line immediately beneath it is the first order expectations, the dotted line next is the second order expectation and so on. 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 prices will be more informative about θ_t and other traders' (higher order) expectations of

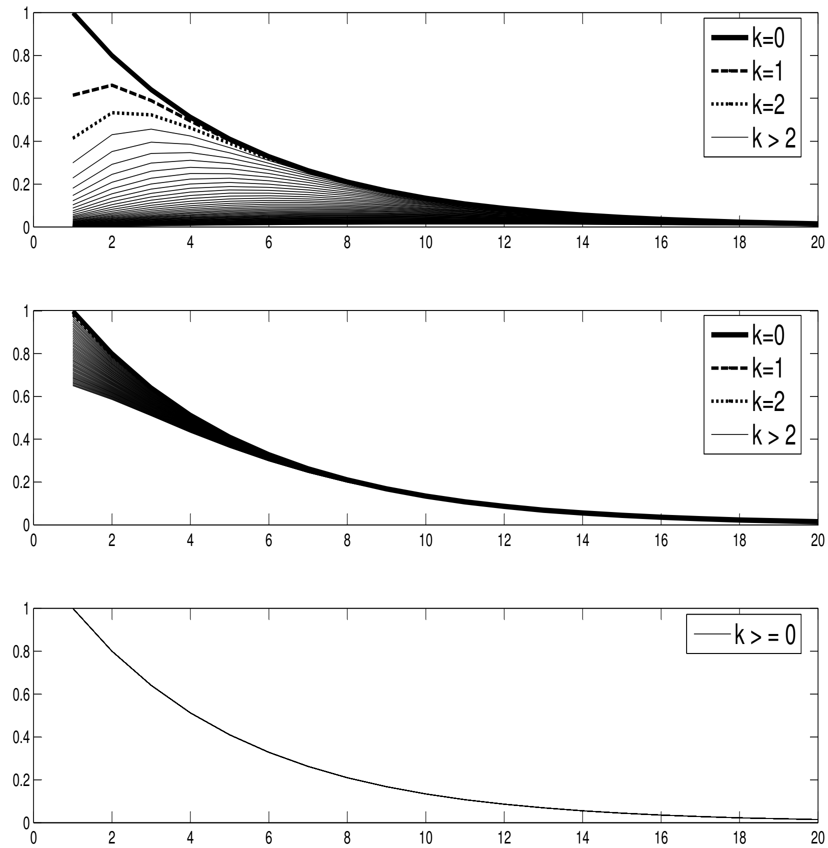

FIGURE 3. Impulse responses of $\theta_{t|t}^{(0:50)}$ to a unit innovation v_t

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 $\theta_t^{(0:50)}$. 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 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.

8.4. **Accuracy.** In the previous section, we demonstrated that a finite number of orders of expectations are sufficient to accurately represent the equilibrium dynamics of the model. In practice, a maximum order $\overline{k}$ need to be chosen such that we are confident that including a larger number of orders of expectations would not change the dynamics of the model. The number of orders of expectations required to accurately represent the equilibrium depends on the parameters of the model and a decision of how many orders of expectations to include have to be made on a case by case basis. Here, we illustrate that for both the row vector $\mathbf{a}_{\overline{k}}$ and the impulse response functions to the aggregate shocks to converge, relatively few orders of expectations are needed.

In Figure 4, the row vector $\mathbf{a}_{\overline{k}}$ is plotted for $\overline{k} = 1, 3, ...10$. We can see that the vector converges quite rapidly. When 6 or 7 orders of expectations are included, adding higher order expectations beyond that does not further alter the elements of $\mathbf{a}_{\overline{k}}$. We Can also see that the elements of $\mathbf{a}_{\overline{k}}$converges quite rapidly towards zero, so that the Proposition 6, which stated that the series $\{\Sigma a_k\}_{k=0}^{\infty}$ converges , seems "bite" already for relatively low values of k.

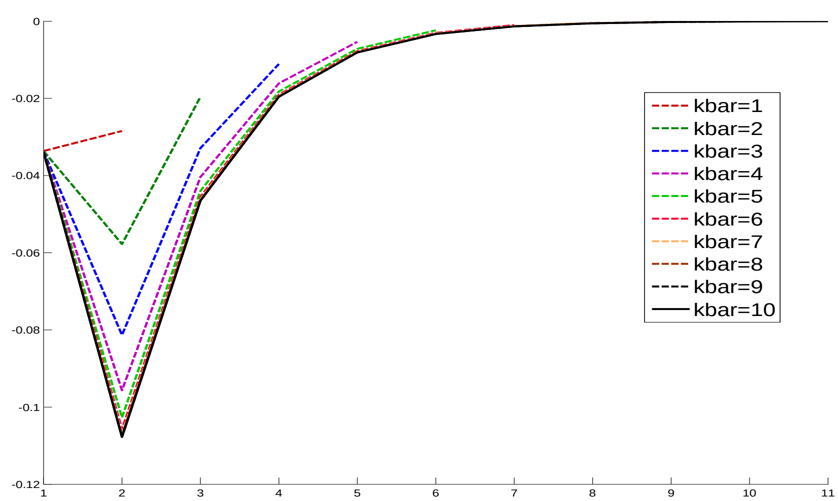

FIGURE 4. Equilibrium $\mathbf{a}_{\overline{k}}$ for $\overline{k} = 1, 2, ..., 10$

In order to have a satisfactory approximation to the infinite dimensional dynamics, we would also like the responses of the endogenous price to aggregate shock to converge. In Figure 5, the impulse response functions to a persistent (left panel) and transitory (right panel) supply shocks are plotted for $\overline{k} = 1, 3, ...15$. We can see that also also the impulse response converges rapidly and that they do not change much when additional orders beyond the first 6 or 7 are added.

Since the impulse response functions to the two aggregate shocks completely describes the dynamics of the price, convergence of one implies convergence of the other. 6 or seven orders of expectations thus appear sufficient to accurately represent the equilibrium dynamics of the

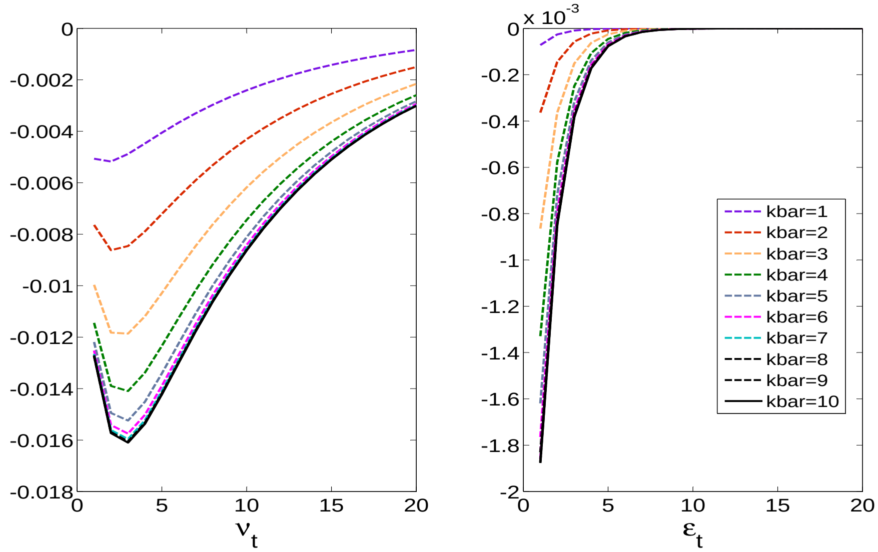

FIGURE 5. Impulse responses of p_t for $\overline{k} = 1, 2, ..., 10$

price. Of course, this depends on the parameterization of the model. In general, the more persistent the supply shocks is, i.e. the closer ρ is to unity, the more orders of expectations are necessary. However, it is difficult to find parameterizations for which more than 10 orders of expectations are needed.

9. 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 form expectations rationally. 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 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 future fundamentals. 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 it self 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.

The equilibrium representation derived here can be taken as a literal description of agents' behavior, i.e. as representing agents who explicitly form expectations about other agents' expectations. The convergence results derived here can then be comforting for readers who find it implausible on cognitive limitations grounds that traders form an infinite number of higher order expectations. Indeed, what has been shown here is that forming only a finite and even low number of orders of expectations may in some settings be sufficient.

An alternative interpretation is to view the equilibrium representation simply as a convenient functional form to model agents who have access to to private information and condition on the entire history of observables. The main advantage of the method is then to deliver a finite dimensional and time invariant representation of equilibrium dynamics.

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 modeling 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. This paper helps 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.

While the model here had scalar process as the latent fundamental that traders formed expectations about, none of the proofs rely on this fact. The method works for a general vector valued latent process and have been applied both to calibrated macro models, as in Nimark (2008) Graham and Wright (2010) and to an estimated term structure model in Nimark (2010).

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*, Queen Mary, University of London 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] Graham, L. and S. Wright, 2010, "Information, heterogeneity and market incompleteness", *Journal of Monetary Economics*, vol. 57(2), pp64-174.

[6] Kasa, Kenneth, 2000, "Forecasting the forecast of others in the frequency domain", *Review of Economic Dynamics,* 3, pp726-756.

[7] Kasa, K., T. Walker, and C. Whiteman, 2007, "Asset Prices in a Time Series Model with Perpetually Disparately Informed, Competitive Traders", mimeo, Indiana University.

[8] Lorenzoni, Guido, 2005, "A Theory of Demand Shocks" NBER working paper.

[9] Luenberger, David G., 1969, *Optimization by Vector Space Methods*, Wiley-Interscience, New York.

[10] Mankiw, G. R. Reis and J. Wolfers, 2003, "Disagreement about Inflation Expectations", NBER Macroeconomics Annual, 18, 209-248.

[11] Morris, S. and H.S. Shin, 2002, "The social value of public information", *American Economic Review* 92, pp1521-1534.

[12] Pearlman, J.G. and T.J. Sargent, 2005, "Knowing the forecasts of others", *Review of Economic Dynamics*, Volume 8, pp480-497.

[13] Rondina, G. and T. Walker, 2010, "Information Equilibria in Dynamic Economies", *mimeo*, UCSD and University of Indiana.

[14] Sargent, Thomas J., 1991, "Equilibrium with Signal Extraction from Endogenous Variables", *Journal of Economic Dynamics and Control* 15, pp245-273.

[15] 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.

[16] Strang, Gilbert, 1988, "Linear Algebra and Its Applications", 3rd ed., Saunders.

[17] Townsend, Robert M., 1983, Forecasting the Forecasts of Others, *Journal of Political Economy*, vol 91, pp546-588.

[18] Walker, Todd, 2007, "How Equilibrium Prices Reveal Information in Time Series Models with Disparately Informed, Competitive Traders", forthcoming, *Journal of Economic Theory.*

[19] Weinstein, J. and M. Yildiz, forthcoming, "Impact of Higher-Order Uncertainty", *Games and Economic Behavior.*

[20] 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. 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

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

which can be rearranged to

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

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

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

$$\begin{bmatrix}s_t(j)\\ p_t-\frac{\lambda\psi}{1-\lambda\psi}c_t\end{bmatrix} = \begin{bmatrix}e_1' & 0\\ \mathbf{a} & -\delta\gamma\lambda\end{bmatrix}\begin{bmatrix}\theta_{t|t}^{(0:\infty)}\\ \epsilon_t\end{bmatrix} + \begin{bmatrix}1\\ 0\end{bmatrix}\eta_t(j) \tag{A.4}$$

Define

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

$$\widehat{P} \equiv E\left(X_t - X_{t|t-1}\right)\left(X_t - X_{t|t-1}\right)' \tag{A.6}$$

$$\widehat{\mathbf{a}} \equiv \begin{bmatrix}\mathbf{a} & -\delta\gamma\lambda\end{bmatrix} \tag{A.7}$$

then

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

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

$$\delta = \widehat{\mathbf{a}}\widehat{P}\widehat{\mathbf{a}}' + \left[1+2\frac{\lambda\psi}{1-\lambda\psi}+\left(\frac{\lambda\psi}{1-\lambda\psi}\right)^2\right]\sigma_u^2. \tag{A.9}$$

Appendix B. Proof of Lemma 4

Lemma 8. *The variance of the price p_t is finite*

Proof. We want to show that $E(p_t)^2 < \infty$. Taking variances of both sides of the expression for the equilibrium price (3.15) we get

$$E(p_t)^2 = (\delta\gamma\lambda)^2 \sum_{j=0}^{\infty}\sum_{i=0}^{\infty} \lambda^{(i+j)} cov\left[\theta_{t+i|\cdots|t}, \theta_{t+j|\cdots|t}\right] \tag{B.1}$$

$$+2\delta\gamma\lambda \sum_{j=0}^{\infty} \lambda^j cov\left[\theta_{t+i|\cdots|t}, \epsilon_t\right] \tag{B.2}$$

$$+(\delta\gamma\lambda)^2 \sigma_\epsilon^2 + \left(\frac{\lambda\psi}{1-\lambda\psi}\right)^2 \sigma_c^2 \tag{B.3}$$

The two terms on the last line are finite and given exogenously. We thus need to show that the infinite sums on the first and second line converge. We will do this by demonstrating that the absolute values of the covariance term is bounded by the variance of the true supply process, i.e

$$\left|cov\left[\theta_{t+i|\cdots|t}, \theta_{t+j|\cdots|t}\right]\right| \leq E(\theta_t)^2 \tag{B.4}$$

By the Cauchy-Schwartz inequality we know that

$$\left|cov\left[\theta_{t+i|\cdots|t}, \theta_{t+j|\cdots|t}\right]\right| \leq \max\left\{E\left(\theta_{t+i|\cdots|t}\right)^2, E\left(\theta_{t+j|\cdots|t}\right)^2\right\} \tag{B.5}$$

and from Proposition 4 we know that

$$E\left(\theta_{t+i|\cdots|t}\right)^2 \leq E(\theta_t)^2 \tag{B.6}$$

i.e. that the variance of higher order expectations are bounded by the variance of the true process. Applying these results to the first infinite series in (B.1) we have that

$$(\delta\gamma\lambda)^2 \sum_{j=0}^{\infty}\sum_{i=0}^{\infty} \left|\lambda^{(i+j)} cov\left[\theta^{(i)}_{t+i|\cdots|t}, \theta^{(j)}_{t+j|\cdots|t}\right]\right| \leq (\delta\gamma\lambda)^2 \sum_{j=0}^{\infty}\sum_{i=0}^{\infty} \left|\lambda^{(i+j)} E(\theta_t)^2\right| \tag{B.7}$$

$$= \frac{(\delta\gamma\lambda)^2}{(1-\lambda)^2} E(\theta_t)^2 \tag{B.8}$$

$$< \infty \tag{B.9}$$

Similarly, for the second infinite series, we have that

$$2\delta\gamma\lambda \sum_{j=0}^{\infty} \lambda^j \left|cov\left[\theta_{t+i|\cdots|t}, \epsilon_t\right]\right| \leq \max\left\{2\delta\gamma\lambda \sum_{j=0}^{\infty} \lambda^j E(\theta_t)^2, 2\delta\gamma\lambda \sum_{j=0}^{\infty} \lambda^j E(\epsilon_t)^2\right\} \tag{B.10}$$

$$< \infty \tag{B.11}$$

□