

SIMON FRASER UNIVERSITY

Department of Economics

Working Papers

12-06

**“Heterogeneous Beliefs
and Tests of Present
Value Models”**

Ken Kasa, Todd Walker and
Charles Whiteman

November, 2011



HETEROGENEOUS BELIEFS AND TESTS OF PRESENT VALUE MODELS

KENNETH KASA, TODD B. WALKER, AND CHARLES H. WHITEMAN

ABSTRACT. This paper develops a dynamic asset pricing model with persistent heterogeneous beliefs. The model features competitive traders who receive idiosyncratic signals about an underlying fundamentals process. We adapt Futia's (1981) frequency domain methods to derive conditions on the fundamentals that guarantee noninvertibility of the mapping between observed market data and the underlying shocks to agents' information sets. When these conditions are satisfied, agents must 'forecast the forecasts of others'. The additional dynamics of the heterogeneous beliefs equilibrium can account for observed violations of variance bounds, predictability of excess returns, and rejections of cross-equation restrictions.

JEL Classification Numbers: G12, D82

Date: November, 2011.

Kasa, Department of Economics, Simon Fraser University, 8888 University Drive, Burnaby, BC V5A 1S6 CANADA, kkasa@sfu.ca. Walker, Department of Economics, Indiana University, 100 S. Woodlawn, Bloomington, IN 47405-7104, walkertb@indiana.edu. Whiteman, Department of Economics, University of Iowa, Iowa City, IA 52242, whiteman@uiowa.edu. We would like to thank Tim Cogley, Martin Evans, Ron Michener, Bart Taub, Pierre-Olivier Weill and seminar participants at NYU, the Federal Reserve Board, UC Berkeley, UC Davis, UC Santa Barbara, UC Santa Cruz, Indiana University, VPI, UNC, and the Atlanta Fed for helpful discussions. Support from the National Science Foundation under grant SES85-10505 and SES-0962221 is gratefully acknowledged. This paper subsumes and extends Walker and Whiteman's *Equilibrium Volatility and Heterogeneity in a Simple Asset Pricing Model* and Kasa's *Asset Pricing with Heterogeneous Beliefs: A Frequency Domain Approach*.

1. INTRODUCTION

Standard present value models have a difficult time explaining several features of observed asset prices. Prices seem to be excessively volatile, excess returns seem to be predictable, and the model's cross-equation restrictions are typically rejected as well. As a result, linear present value models have all but disappeared from serious academic research on asset pricing.¹ Instead, attention has shifted to models with time-varying risk premia. Unfortunately, these models offer little improvement empirically, although Constantinides and Duffie (1996) achieve some success by introducing heterogeneity, in the form of nondiversifiable labor income risk.

Our paper returns to the linear constant discount rate setting, and argues that a different form of heterogeneity, an *informational* heterogeneity, can account for many of the model's apparent empirical shortcomings. In particular, we make two simple changes to the standard present value model: (1) We assume some fundamentals are unobserved by *all* speculative traders. Although the variance of these unobserved fundamentals can be arbitrarily small, this additional noise breaks the no trade theorems of Milgrom and Stokey (1982) and Tirole (1982). (2) More importantly, we assume speculative traders are heterogeneously informed about the observable fundamentals. Specifically, we assume observed fundamentals consist of a sum of orthogonal components, and that in addition to observing the sum, each trader observes realizations of one of the underlying components. We think of this as a natural information structure. All traders no doubt observe current earnings or dividends, but at the same time they are likely to have heterogeneous information about their underlying determinants.²

Following Futia (1981), we derive conditions under which traders are unable to infer the realizations of the other components of fundamentals. Instead, they are only able to infer a weighted average of them. The weighted averages encode each trader's forecast of other traders' forecasts (Townsend (1983)).

The two main contributions of the paper are the following: [i] Our paper provides an explicit analytical characterization of the resulting higher-order beliefs dynamics. [ii] We show how these additional dynamics play an important role in observed asset prices. More specifically, we show that excess volatility, predictability, and the rejection of cross-equation

¹Cochrane (2001) discusses the empirical failings of constant discount rate models. He argues that many of these apparently distinct anomalies are manifestations of the same underlying problem; namely, misspecification of the discount rate. He also points out that the same problems show up in all asset markets, e.g., stocks, bonds, foreign exchange, real estate, etc.

²Recent papers documenting an important role of heterogeneous beliefs in asset markets include Anderson, Ghysels, and Juergens (2005) for the case of equity markets, and Piazzesi and Schneider (2009) for the case of housing markets. However, this previous work does not formally link heterogeneous beliefs to heterogeneous information within a dynamic equilibrium model.

restrictions can all be reconciled with theory when asset price dynamics follow from the persistent heterogeneous beliefs equilibrium.

Of course, this is not the first paper to study asymmetric information in asset markets.³ However, our paper *is* the first to combine several key ingredients. First, our model is *dynamic*. It features *persistent* heterogeneous beliefs, and a *stationary* equilibrium. Following Grossman and Stiglitz (1980), most work on asset pricing with asymmetric information is confined to static, or nonstationary finite-horizon models. Although this is a useful abstraction for some theoretical questions, it is obviously problematic for empirical applications. There has been some work devoted to dynamic extensions of the Grossman-Stiglitz framework (see, e.g., Wang (1993), He and Wang (1995), Foster and Viswanathan (1996), Albagli, Hellwig, and Tsyvinski (2011)). Most of this literature uses clever modeling assumptions (e.g., hierarchical information structures, truncation solution strategies) to avoid the forecasting the forecasts of others problem first highlighted by Townsend (1983). We analytically derive the component of the asset price that is due to higher-order beliefs.

Second, our approach features *signal extraction from endogenous prices*. This distinguishes our work from most work on global games and imperfect common knowledge (Morris and Shin (2003)). Although this literature has made important contributions to our understanding of higher-order beliefs, it is not directly applicable to asset pricing, since it abstracts from asset markets. As Atkeson (2000) notes, prices play an important role in aggregating information, and it remains to be seen how robust the work on global games is to the inclusion of asset markets.⁴

Third, our approach delivers an *analytical solution*, with explicit closed-form expressions for the role of higher-order beliefs. Although this may seem like a minor contribution given the power of computation, analytical solutions are extremely useful in models featuring a potential infinite regress of higher-order beliefs. Numerical methods in this setting are fraught with dangers. In particular, they require prior knowledge of the relevant state. As first noted by Townsend (1983), it is not at all clear what the state is when agents forecast the forecasts of others. Townsend argued that the logic of infinite regress produces an infinite-dimensional state. Townsend short-circuited the infinite regress and obtained a tractable numerical solution by assuming that information becomes common knowledge after a (small) number of periods. This truncation strategy has been refined by a number of subsequent researchers (see, e.g., Singleton (1987), Bacchetta and van Wincoop (2006, 2008), and Nimark (2007)). However, recent work by Pearlman and Sargent (2005) and Walker (2007) demonstrates that numerical approaches can be misleading.

³See Brunnermeier (2001) for a review.

⁴Angeletos and Werning (2006) and Hellwig, Mukherji, and Tsyvinski (2006) incorporate signal extraction from prices into the Morris-Shin framework. However, their models focus on the issue of equilibrium uniqueness, and are essentially static.

Our approach adapts and extends the frequency domain methods of Futia (1981). These methods exploit the power of the Riesz-Fischer Theorem, which allows us to transform a difficult time-domain/sequence-space signal extraction problem into a much easier function space problem.⁵ Rather than guess a state vector and then solve a Kalman filter’s Riccati equation, a frequency domain approach leads to the construction of so-called Blaschke factors. Finding these Blaschke factors is the key to solving an agent’s signal extraction problem. Our model’s solution takes the form of a nonfundamental (i.e., noninvertible) moving-average representation, mapping the underlying shocks to agents’ information sets to observed prices and fundamentals. Blaschke factors convert this to a Wold representation, which delivers the endogenous information set of the agents. The (statistical) innovations of the Wold representation turn out to be complicated moving averages of the entire histories of the underlying (economic) shocks. These moving averages encode the model’s higher-order belief dynamics. By solving the model in the frequency domain, we are able to isolate the component of the equilibrium due to higher-order beliefs and derive conditions under which heterogeneous beliefs are preserved in equilibrium.⁶

A key contribution of our paper is that this equilibrium representation can be taken to the data in a direct, *quantitative* way. This allows us to revisit past empirical failures of linear present value models. In particular, we ask the following question - Suppose asset markets feature heterogeneous beliefs, but an econometrician mistakenly assumes agents have homogeneous beliefs, what will he conclude?

One might think, based on the conditioning down arguments of Hansen and Sargent (1991a) and Campbell and Shiller (1987), that this would not create problems. Interestingly, this is not the case because conditioning down does not work here. The arguments of Hansen-Sargent and Campbell-Shiller apply to settings where agents and *econometricians* have different information sets. They do not apply in general to settings where there is informational heterogeneity among the agents themselves. This is because the law of iterated expectations does not apply to the average beliefs operator (Allen, Morris, and Shin (2006), Morris and Shin (2003)).

Using updated data from Shiller (1989) on the U.S. stock market, we show that many of the empirical shortcomings documented by Shiller can be accounted for by higher-order belief dynamics, as opposed to fads or ‘market psychology’. We show that present value

⁵Kasa (2000) uses frequency domain methods to solve Townsend’s model. He comes to the same conclusion as Pearlman and Sargent (2005). Appendix B provides a review of the Riesz-Fischer Theorem.

⁶Makarov and Rytchkov (2008), Bernhardt, Seiler, and Taub (2010) and Rondina and Walker (2011) also use frequency domain techniques to solve dynamic models with heterogeneously informed agents. Makarov and Rytchkov (2008) argue that a finite-state equilibrium does not exist. However, their fundamentals specification does not satisfy our existence condition, which could explain the nonexistence. Bernhardt, Seiler, and Taub (2010) examine an asset pricing model with strategic use of information when traders are influential. This additional complication calls for a numerical solution procedure. Rondina and Walker (2011) extend (Futia 1981) and the results of this paper to the case of dispersed information.

models with heterogeneous beliefs generate predictable excess returns, produce violations of variance bound inequalities, and rejections of cross-equation restrictions. In fact, we argue that *rational* heterogeneous belief dynamics could well be mistaken for fads or irrational expectations.

Hence, our paper sounds a note of caution when interpreting previous rejections of present value models. Perhaps it is not the constant discount rate that is the problem, but rather the (implicit) assumption of homogeneous beliefs, or equivalently, a fully revealing equilibrium.

2. THE MODEL

Consider the following linear present value model

$$p_t = \beta \int_0^1 E_t^i p_{t+1} di + f_t + u_t \quad (2.1)$$

where p_t represents the price of an asset (e.g., an equity price or an exchange rate), f_t represents a (commonly) observed fundamental (e.g., dividends or the money supply), and u_t represents the influence of unobserved fundamentals (e.g., noise or liquidity traders). The parameter, $\beta < 1$, represents a constant discount factor. The model in (2.1) is entirely standard, with two key exceptions.⁷ The first is the presence of the noise term, u_t . Since ultimately we are going to focus on nonrevealing, heterogeneous beliefs equilibria, it is important that some noise be present to sustain trade. (Milgrom and Stokey (1982), Tirole (1982))⁸. The second key difference is that expectations in (2.1) are indexed by traders, to acknowledge the possibility that beliefs may differ in equilibrium.

What gives rise to these heterogeneous beliefs? One possibility is heterogeneous priors (Harrison and Kreps (1978)). Besides posing awkward questions about the source of this heterogeneity, another problem with this approach is that it generates nonstationary equilibria, in which belief heterogeneity dissipates over time (Morris (1996)). In response, we instead suppose that belief heterogeneity arises from, and is sustained by, an exogenous ongoing process of heterogeneous information. The idea is that each period investors acquire information about some aspect of an asset's underlying observed fundamentals. For simplicity, we suppose there are just two types of traders, Type 1 and Type 2, and that observed fundamentals are driven by the exogenous process:

$$f_t = a_1(L)\varepsilon_{1t} + a_2(L)\varepsilon_{2t} \quad (2.2)$$

⁷Appendix C develops a microfounded model that delivers (2.1).

⁸Although noisy rational expectations models have a long history in finance and macroeconomics, prior applications assume *homogeneous* beliefs. Whiteman (1983) uses frequency domain methods to characterize the solutions of equation (2.1) in the case of homogeneous expectations. He refers to models of the form (2.1) as 'perturbed equations'. Hansen and Sargent (1991a) refer to them as 'inexact' rational expectations models.

where $a_1(L)$ and $a_2(L)$ are square-summable polynomials in the lag operator L . The innovations, ε_{1t} and ε_{2t} , are zero mean Gaussian random variables, and are assumed to be uncorrelated. Each period both traders observe p_t and f_t . However, in addition, Type 1 traders observe realizations of ε_{1t} , while Type 2 traders observe realizations of ε_{2t} . Both the model and the information structure are common knowledge.

Specification of the information structure also requires assumptions about the noise process, u_t . Although neither trader observes u_t , both its existence and its law of motion are common knowledge.⁹ Again to keep things simple, we suppose u_t is i.i.d., and is driven by:

$$u_t = b_1\varepsilon_{1t} + b_2\varepsilon_{2t} + v_t \quad (2.3)$$

where v_t is a zero mean, i.i.d. Gaussian random variable that is uncorrelated with both ε_{1t} and ε_{2t} . Notice that in general the noise process, u_t , can be correlated with the observable component of fundamentals, f_t . This assumption simply serves to simplify the algebra that follows and is not necessary for our results. Moreover as we emphasize below, the excess volatility delivered by the model with heterogeneous beliefs is *not* due to the additional volatility generated by noise traders. In the results below, we show that excess volatility holds as we drive the variance of the noise trader term to zero.

It is important to keep in mind that both traders behave in a competitive, price-taking manner. By assumption, their only task each period is to forecast next period's price. We assume they do this in a statistically optimal way, given their information. In contrast to the global games literature, there is no *explicit* effort here to infer other agents' forecasts. In our Walrasian environment, there is no need to, since nothing you do can influence the expectations of others. However, and this is the crucial point, since traders use the *endogenously* determined history of prices as a basis for their own individual forecasts, and these prices depend on other agents' forecasts, there is a sense in which each trader's optimal forecast does embody a forecast of other traders' forecasts, but these forecasts are simply a by-product of each agent's own atomistic efforts to forecast prices.

A key aspect of the environment here is that it is both stationary and linear. As a result, we can employ the tools of Wiener-Kolmogorov prediction theory to solve each trader's forecasting problem. The first step in doing this is to derive the mapping between what he observes and the underlying shocks driving the system. The symmetry between the agents, along with the orthogonality between ε_{1t} and ε_{2t} , implies that we can focus on the problem of a single trader, say Type 1. Given the solution to Type 1's problem, we can infer the

⁹Engel and West (2005) argue that noise, or unobserved fundamentals, appear to be necessary to reconcile present value models with observed exchange rates. Hamilton and Whiteman (1985) argue that the mere possibility of unobserved fundamentals vitiates standard bubbles tests. Our results suggest that it is the *interaction* between unobserved fundamentals and heterogeneous information about observed fundamentals that is critical to the success of present value models.

solution to Type 2's via symmetry. For Type 1 traders, the mapping between observables and the underlying shocks takes the following form,

$$\begin{bmatrix} \varepsilon_{1t} \\ f_t \\ p_t \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ a_1(L) & a_2(L) & 0 \\ \pi_1(L) & \pi_2(L) & \pi_3(L) \end{bmatrix} \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ v_t \end{bmatrix} \quad (2.4)$$

where the $\pi_i(L)$ polynomials are equilibrium pricing functions. Each trader knows these functions when forecasting next period's price. Of course, these pricing functions depend on the forecasts via the equilibrium condition (2.1), so we have a fixed point problem. Traditionally, this fixed point problem is resolved in one of two ways. First, one may pursue a 'guess and verify' strategy. That is, posit functional forms for the $\pi_i(L)$ functions, use them to solve the agents' prediction problems, plug the predictions into the equilibrium condition, and then match coefficients. This approach works well when the relevant state is clear (and low dimensional). Unfortunately, in dynamic settings with potentially heterogeneous beliefs, it is not at all clear what the state is, or equivalently, what forms the $\pi_i(L)$ functions should take. What we are really searching for is an unknown *function* which, absent prior information, lies in an infinite dimensional space. Kasa (2000) shows that this infinite dimensional fixed point problem can be solved using frequency domain methods, and that will be the approach we pursue here.

Before doing that, however, it might be worth considering the second strategy that is commonly employed when solving this kind of fixed point problem. Rather than guessing an unknown pricing function, and using it to forecast next period's price, an alternative strategy is to iterate equation (2.1) forward. With homogeneous beliefs, this strategy is quite powerful as it produces an expression for the current price as a single conditional expectation of the discounted sum of future fundamentals, which can then be solved, for example, using the Hansen-Sargent prediction formula. No guessing and verifying is required. What makes this work is the law of iterated expectations. Unfortunately, as noted earlier, the law of iterated expectations does not apply when there are heterogeneous expectations. Still, one could in principle approach the problem via iteration.¹⁰ To do this, define

$$\bar{\mathbb{E}}_t^0 f_{t+1} = \int_0^1 E[f_{t+1} | \Omega_t^i] di$$

where Ω_t^i denotes agent- i 's information set, and then analogously define

$$\bar{\mathbb{E}}_t^k f_{t+k+1} = \bar{\mathbb{E}}_t \bar{\mathbb{E}}_{t+1} \cdots \bar{\mathbb{E}}_{t+k} f_{t+k+1}$$

as the k -fold iteration of these averaged expectations. Using this notation we can then write the equilibrium condition in (2.1) as

¹⁰Bacchetta and van Wincoop (2008) adopt this approach.

$$p_t = f_t + u_t + \beta \sum_{k=0}^{\infty} \beta^k \bar{\mathbb{E}}_t^k (f_{t+k+1} + u_{t+k+1}) \quad (2.5)$$

The problem here is that $\bar{\mathbb{E}}_t^k$ depends on the information conveyed by p_t , but p_t in turn depends on the entire infinite sequence of $\bar{\mathbb{E}}_t^k$. Hence, we are back to an infinite dimensional fixed point problem. Existing approaches either approximate the solution of this infinite dimensional problem (e.g., Nimark (2007), Bernhardt, Seiler, and Taub (2010)) or effectively truncate it by supposing that all relevant information becomes common knowledge after a certain lag (e.g., Townsend (1983), Singleton (1987), Bacchetta and van Wincoop (2006, 2008)). In the next section we show how to tackle this problem head on by using frequency domain methods.

3. CONSTRUCTING A NONREVEALING EQUILIBRIUM

A frequency domain approach is useful here for two reasons. First, as noted above, we have to solve an infinite dimensional fixed point problem. Without prior knowledge of the functional forms of the equilibrium prices, we must be prepared to match an infinite number of unknown coefficients. By transforming the problem to the frequency domain we can convert this to the problem of finding a single analytic *function*, which via the Riesz-Fischer Theorem, is equivalent to the unknown coefficient sequence.¹¹ Second, and related to this, the underlying source of our infinite dimensional fixed point problem is that we are attempting to calculate an *informational* fixed point. In particular, we must somehow guarantee that traders are unable to infer the private information of other traders via the *infinite history* of observed market data. This is a difficult problem to even formulate in the time domain. In contrast, handling this problem in the frequency domain is straightforward, as the information revealing properties of analytic functions are completely characterized by the locations of their zeros. Zeros inside the unit circle correspond to noninvertible moving average representations and unobservable shocks.¹²

3.1. The Signal Extraction Problem. In our model each trader observes a *vector* of new information each period, as summarized by (2.4), so noninvertibility relates to the zeros of the determinant of a matrix. In particular, write (2.4) as $\mathbf{x}_{1t} = \mathbf{M}_1(L)\boldsymbol{\varepsilon}_{1t}$. If $\det \mathbf{M}_1(L)$ has all its zeros outside the unit circle, then $\mathbf{M}_1(L)$ possesses a one-sided inverse in positive powers of L . In this case, the observed history of \mathbf{x}_t would reveal $\boldsymbol{\varepsilon}_{1t}$ to Type 1 traders, and therefore, they would be able to infer the private information of Type 2 traders, $\boldsymbol{\varepsilon}_{2t}$.

Letting $\mathbf{M}_1(z)$ denote the z -transform of $\mathbf{M}_1(L)$, one can readily verify from (2.4) that

¹¹The Appendix provides a brief discussion of this theorem and its implications. See Whiteman (1983) for a more detailed discussion.

¹²This point has been emphasized in particular in the work of Bart Taub. See, e.g., Taub (1990).

$$\det \mathbf{M}_1(z) = a_2(z)\pi_3(z)$$

Analogously, one can easily verify that the observer system for Type 2 traders implies $\det \mathbf{M}_2(z) = a_1(z)\pi_3(z)$. Hence, a sufficient condition for neither trader to be able to infer the other trader's private information is if $\pi_3(z)$ turns out to have a zero inside the unit circle (while at the same time neither $a_i(z)$ possesses a corresponding pole). It turns out, however, to be more natural and convenient to suppose that noninvertibility stems from noninvertible roots in the $a_i(z)$ functions. At the same time, it proves to be convenient to suppose that these roots are identical, and that they are inherited by the equilibrium pricing functions. To summarize, we assume the following,

Assumption 3.1. *The analytic functions $a_1(z)$ and $a_2(z)$ have a single, identical root inside the unit circle. This common root is shared by the equilibrium pricing functions, $\pi_1(z)$ and $\pi_2(z)$. At the same time, the pricing component, $\pi_3(z)$, has no zeros inside the unit circle.*

These assumptions are more restrictive than necessary. Their main purpose is to simplify the mathematics. All we really need is for $a_1(\cdot)$ and $a_2(\cdot)$ to have coincident or non-coincident zeros inside the unit circle for all the paper's substantive conclusions to go through. This is sufficient to ensure $\det \mathbf{M}_1(L)$ and $\det \mathbf{M}_2(L)$ remain noninvertible.

Ensuring that Assumption 3.1 is satisfied imposes restrictions on the observed fundamentals, f_t . Constructing a nonrevealing equilibrium involves deriving these restrictions. Fortunately, as we shall see, these restrictions can be interpreted as merely imposing a scaling, or relative variance factor.

Assumption 3.1 allows us to easily derive the Wold representation of each trader's observer system, which yields the trader's information set. First write $a_i(z)$ and $\pi_i(z)$ as $(z - \lambda)\tilde{a}_i(z)$ and $(z - \lambda)\tilde{\pi}_i(z)$, where $\tilde{a}_i(z)$ and $\tilde{\pi}_i(z)$ are analytic functions with all roots outside the unit circle, and $|\lambda| < 1$ represents the common noninvertible root of the pricing and fundamentals processes. Applying the methods of Rozanov (1967) and Hansen and Sargent (1991b) then delivers the following Wold representation for Type 1 traders

$$\begin{bmatrix} \varepsilon_{1t} \\ f_t \\ p_t \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ (L - \lambda)\tilde{a}_1(L) & (1 - \lambda L)\tilde{a}_2(L) & 0 \\ (L - \lambda)\tilde{\pi}_1(L) & (1 - \lambda L)\tilde{\pi}_2(L) & \pi_3(L) \end{bmatrix} \begin{bmatrix} \varepsilon_{1t} \\ e_{2t} \\ v_t \end{bmatrix} \quad (3.1)$$

where $e_{2t} = [(L - \lambda)/(1 - \lambda L)]\varepsilon_{2t}$. Write this system as $\mathbf{x}_{1t} = \mathbf{M}_1^*(L)\boldsymbol{\varepsilon}_{1t}^*$. Note that $\det \mathbf{M}_1^*(L) = (1 - \lambda L)\tilde{a}_2(L)\pi_3(L)$. Hence, we have effectively 'flipped' the root inside the unit circle, λ , to λ^{-1} , which is outside the unit circle. The key point here is that Type 1 traders are unable to use the observed history of \mathbf{x}_{1t} to infer realizations of Type 2 traders' private information, ε_{2t} . The best they can do is estimate the moving average defined by e_{2t} . The lag polynomial $(L - \lambda)/(1 - \lambda L)$, which relates e_{2t} to ε_{2t} is an example of a 'Blaschke factor'.

It compactly summarizes Type 1's best efforts to use Type 2's information to forecast future prices and fundamentals. Although it may appear as if e_{2t} is autocorrelated, one can easily show that it is in fact an i.i.d. innovation sequence.

Given (3.1), Type 1's optimal forecast of $\mathbf{x}_{1,t+1}$ is a straightforward application of the Wiener-Kolmogorov prediction formula

$$\begin{aligned} E_t^1 \mathbf{x}_{1,t+1} &= \left[\frac{\mathbf{M}_1^*(L)}{L} \right]_+ \boldsymbol{\varepsilon}_{1t}^* \\ &= L^{-1} [\mathbf{M}_1^*(L) - \mathbf{M}_1(0)^*] \boldsymbol{\varepsilon}_{1t}^* \end{aligned} \quad (3.2)$$

A completely symmetrical expression characterizes Type 2's forecast, with the crucial difference that the unobservable shock is now ε_{1t} . Using (3.2) and its Type 2 analog, yields the following expressions for the optimal price forecasts

$$E_t^1 p_{t+1} = L^{-1} \{ [(L - \lambda)\tilde{\pi}_1(L) + \lambda\tilde{\pi}_1(0)]\varepsilon_{1t} + [(1 - \lambda L)\tilde{\pi}_2(L) - \tilde{\pi}_2(0)]e_{2t} + [\pi_3(L) - \pi_3(0)]\varepsilon_{3t} \} \quad (3.3)$$

$$E_t^2 p_{t+1} = L^{-1} \{ [(1 - \lambda L)\tilde{\pi}_1(L) - \tilde{\pi}_1(0)]e_{1t} + [(L - \lambda)\tilde{\pi}_2(L) + \lambda\tilde{\pi}_2(0)]\varepsilon_{2t} + [\pi_3(L) - \pi_3(0)]\varepsilon_{3t} \} \quad (3.4)$$

As noted earlier, symmetry between the traders and the orthogonality between ε_{1t} and ε_{2t} allows us to focus on just one of the components. In what follows, let E_t^{i,ε_1} denote Type- i 's projection onto the time- t history of ε_1 . Without loss of generality, we can further parameterize the unknown pricing functions as

$$\tilde{\pi}_i(L) = \rho_i + Lg_i(L) \quad (3.5)$$

where $g_i(z)$ is assumed to be an invertible function. We then have

$$E_t^{1,\varepsilon_1} p_{t+1} = [\rho_1 + (L - \lambda)g_1(L)]\varepsilon_{1t} \quad (3.6)$$

$$E_t^{2,\varepsilon_1} p_{t+1} = \left[\rho_1 + (L - \lambda)g_1(L) - \frac{\rho_1(1 - \lambda^2)}{1 - \lambda L} \right] \varepsilon_{1t} \quad (3.7)$$

where we have used Type 2's Blaschke factor, $e_{1t} = [(L - \lambda)/(1 - \lambda L)]\varepsilon_{1t}$, to substitute out e_{1t} in Type 2's forecast function.

Equations (3.6) and (3.7) show explicitly how the forecasts of the two traders differ as a function of the entire history of past realizations of ε_{1t} . By combining these with the analogous expressions for the ε_{2t} projections, we obtain the following key result.

Proposition 3.2. *If ρ_1 and ρ_2 are negative, then traders respond more aggressively to realizations of other traders' (unobserved) signals.*

Proof. This can be seen by differencing (3.6) and (3.7) and the analogous expressions in the ε_2 shock,

$$E_t^1 p_{t+1} - E_t^2 p_{t+1} = \frac{1 - \lambda^2}{1 - \lambda L} (\rho_1 \varepsilon_{1t} - \rho_2 \varepsilon_{2t}). \quad (3.8)$$

□

At each point in time, forecasts of the two traders differ as a function of the infinite histories of their observed signals. If the equilibrium ρ_i parameters are negative, then the trader who has received higher signals (on average) will forecast *lower* prices. We shall see that ρ_i will be negative when the b_i coefficients in (2.3) are positive, i.e., when noise is positively correlated with observable fundamentals. The intuition here is simple - when noise and observable fundamentals affect prices in the same way, price increases due to noise, for example, are partly attributed to other traders having good news about observable fundamentals. Thus, traders ‘overreact’ to realizations of other traders’ (unobserved) signals. In other words, since information about other traders’ signals is only obtainable by observing prices and commonly observed fundamentals, Proposition 3.2 yields the well known result that agents overreact to public signals (Allen, Morris, and Shin (2006)). We shall see that this overreaction to public signals can generate apparent violations of variance bounds.

3.2. The Fixed Point Problem. From a technical standpoint, our paper makes two main contributions. The first is the use of Blaschke factors to solve the traders’ otherwise difficult signal extraction problems, as summarized by equations (3.3) and (3.4) in the previous subsection. The second is to adopt a backsolving strategy to simplify the construction of a nonrevealing Rational Expectations equilibrium. It is to this task that we now turn.

The sense in which we are working backwards is that we *presume* the equilibrium pricing functions, $\pi_i(L)$, inherit a noninvertible root, λ , from the unobserved components of the fundamentals process, f_t . Clearly, this will not be the case for arbitrary specifications of f_t . Instead, specification of λ imposes a restriction on the fundamentals. In particular, we must make the following assumption¹³

Assumption 3.3. *Let $\kappa_i =$ the share of Type- i traders. Then for any given value of $|\lambda| < 1$, the components of the observable fundamentals satisfy the restrictions*

$$\beta \tilde{a}_i(\beta) + b_i \frac{\kappa_i^{-1} - \lambda\beta}{1 - \lambda\beta} = 0 \quad i = 1, 2$$

Although it will not be clear where this assumption comes from until we write down the fixed point problem, it should be clear at this point that Assumption 3.3 merely imposes a scaling, or relative variance, factor on the f_t process. For example, suppose the functions $\tilde{a}_i(L)$ are identical, and that $\tilde{a}_i(L) = \tilde{a}_i(0)/(1 - \gamma L)$, implying that f_t is ARMA(1,1). Then Assumption 3.3 requires the scaling factors, $\tilde{a}_i(0)$, be equal to $-b_i(1 - \gamma\beta)(\kappa_i^{-1} - \lambda\beta)/[\beta(1 - \lambda\beta)]$.

¹³As noted earlier, this existence condition stems from our assumption that prices and fundamentals share the common noninvertible root, λ . It is not a *necessary* condition. All that is necessary is that \mathbf{M}_1 and \mathbf{M}_2 remain noninvertible. It is not even necessary that the noninvertible roots of the fundamentals components be the same.

From here on we can follow entirely standard procedures to compute a Rational Expectations equilibrium. Assuming the share of each trader type is constant, we have the following equilibrium condition

$$p_t = \beta [\kappa_1 E_t^1 p_{t+1} + \kappa_2 E_t^2 p_{t+1}] + f_t + u_t \quad (3.9)$$

Combining this with the solutions to the signal extraction problems in equations (3.3) and (3.4) produces the following central result

Proposition 3.4. *Under Assumptions 3.1 and 3.3, for any $|\lambda| < 1$ there exists a heterogeneous beliefs Rational Expectations equilibrium for the model described by equations (3.9), (2.2) and (2.3). The z -transforms of the equilibrium pricing functions are given by:*

$$\pi_1(z) = (z - \lambda) \left\{ \tilde{a}_1(\beta) + \frac{\kappa_2 b_1}{\kappa_1} \frac{\lambda}{1 - \lambda\beta} + \frac{z}{z - \beta} \left[\tilde{a}_1(z) - \tilde{a}_1(\beta) + \frac{\kappa_2 b_1}{\kappa_1} \lambda \left(\frac{1}{1 - \lambda z} - \frac{1}{1 - \lambda\beta} \right) \right] \right\} \quad (3.10)$$

$$\pi_2(z) = (z - \lambda) \left\{ \tilde{a}_2(\beta) + \frac{\kappa_1 b_2}{\kappa_2} \frac{\lambda}{1 - \lambda\beta} + \frac{z}{z - \beta} \left[\tilde{a}_2(z) - \tilde{a}_2(\beta) + \frac{\kappa_1 b_2}{\kappa_2} \lambda \left(\frac{1}{1 - \lambda z} - \frac{1}{1 - \lambda\beta} \right) \right] \right\} \quad (3.11)$$

The proof is by construction. Notice that it is sufficient to verify the result for $\pi_1(z)$, due to symmetry. Substituting the conditional expectations (3.3) and (3.4) into the equilibrium condition (3.9), and using the notation in (3.5), gives

$$(L - \lambda)(\rho_1 + Lg_1(L))\varepsilon_{1t} = \beta \left[\rho_1 + (L - \lambda)g_1(L) - \kappa_2 \frac{\rho_1(1 - \lambda^2)}{1 - \lambda L} \right] \varepsilon_{1t} + (L - \lambda)\tilde{a}_1(L)\varepsilon_{1t} + b_1\varepsilon_{1t}$$

The requirement that this hold for *all* realizations of ε_{1t} implies the z -transforms of the two sides must be identical as analytic functions inside the unit circle.¹⁴

$$(z - \lambda)(\rho_1 + zg_1(z)) = \beta \left[\rho_1 + (z - \lambda)g_1(z) - \kappa_2 \frac{\rho_1(1 - \lambda^2)}{1 - \lambda z} \right] + (z - \lambda)\tilde{a}_1(z) + b_1 \quad (3.12)$$

Since $(\rho_1 + zg_1(z))$ is presumed analytic, the right-hand side of (3.12) must be zero when evaluated at $z = \lambda$. (This ‘removes’ the singularity at $z = \lambda$). Setting the right side to zero at $z = \lambda$ determines the unknown constant, $\rho_1 = -b_1/\beta\kappa_1$. If we then substitute this back in, collect terms in $g_1(z)$, and divide by $z - \lambda$, we get

$$(z - \beta)g_1(z) = \tilde{a}_1(z) + \frac{b_1}{\beta\kappa_1} + \frac{\kappa_2}{\kappa_1} b_1 \frac{\frac{1 - \lambda^2}{1 - \lambda z} - 1}{z - \lambda} \quad (3.13)$$

Notice the right-hand side is analytic by construction. Since $g_1(z)$ is also assumed to be analytic, the right-hand side of (3.13) must be zero when evaluated at $z = \beta$. Evaluating

¹⁴The Appendix provides more detail concerning this solution method. See also Whiteman (1983).

the right-hand side at $z = \beta$ and setting it to zero yields the restriction

$$\beta \tilde{a}_1(\beta) + \frac{b_1}{\kappa_1} + \beta \frac{\kappa_2}{\kappa_1} b_1 \frac{\frac{1-\lambda^2}{1-\lambda\beta} - 1}{\beta - \lambda} = 0 \quad (3.14)$$

which can be simplified to obtain the existence condition in Assumption 3.3. Finally, given $g_1(z)$ and ρ_1 , the expression for $\pi_1(z)$ given by (3.10) follows from plugging into $\pi_1(z) = (z - \lambda)(\rho_1 + zg_1(z))$.

A few points are worth making about this result. First, the economics behind the existence condition can now be clarified. The existence condition places restrictions on the exogenous process that guarantee Type 1 agents do not learn fully about ε_{2t} and vice versa. Type 1 agents will use common knowledge to back out as much information as possible from the endogenous variable, p_t . This implies Type 1 agents will form expectations according to the information in

$$p_t - \beta \kappa_1 E_t^1 p_{t+1} - f_t = \beta \kappa_2 E_t^2 p_{t+1} + u_t \quad (3.15)$$

That is, the left-hand side of (3.15) contains all variables in Type 1's information set at date t , and the right-hand side of (3.15) are the objects that Type 1 agents do not observe directly but try to infer from endogenous variables. Assumption (3.3) guarantees that the right-hand side of (3.15) is noninvertible (with root λ inside the unit circle).

Second, notice that no mention of $\pi_3(z)$ is made in Proposition 3.4. Since v_t is known to be i.i.d., it's clear that in equilibrium $\pi_3(z) = 1$.

Third, the expressions in (3.10) and (3.11) are less formidable than they appear. They look complicated simply because they apply for *any* stationary specification of the observable fundamentals (subject to the constraints described in Assumptions 3.1 and 3.3). This kind of generality is a key virtue of a frequency domain approach. However, for simple specifications of the $\tilde{a}_i(z)$ functions, equations (3.10) and (3.11) produce simple expressions for equilibrium prices. For example, suppose both components are ARMA(1,1), with common AR coefficient γ and common MA root λ (i.e. $\tilde{a}_i(z) = \tilde{a}_i(0)/(1 - \gamma z)$). Substituting into (3.10) and (3.11), imposing the existence condition, and then simplifying, produces the following ARMA(2,2) process for prices¹⁵

$$p_t = \left(\frac{L - \lambda}{\beta(1 - \lambda\beta)(1 - \gamma L)(1 - \lambda L)} \right) \left\{ \sum_{i=1,2} b_i [\beta \lambda \kappa_i^{-1} (1 - \kappa_i)(1 - \gamma L) - (\kappa_i^{-1} - \lambda\beta)(1 - \lambda L)] \varepsilon_{it} \right\}$$

Finally, note that no claim of uniqueness was made in Proposition 3.4. Uniqueness would require an additional 'regularity' condition on the traders' Wold representations, which would impose analyticity conditions on the individual elements of $\mathbf{M}_i^*(z)$, in addition to conditions on the roots of the determinant. We have already imposed these conditions on the $\pi_i(z)$

¹⁵Ignoring, for simplicity, the v_t term. One can readily verify that adding this term still produces an ARMA(2,2).

functions, so if we impose them on the individual $a_i(z)$ functions we obtain uniqueness as well.¹⁶

3.3. Heterogenous Beliefs Dynamics. To interpret the heterogeneous beliefs equilibrium given by equations (3.10) and (3.11), it is useful to consider the benchmark case of homogeneous beliefs. Suppose each trader observes *both* ε_{1t} and ε_{2t} . In this case, we can solve for the equilibrium as usual by applying the Hansen-Sargent prediction formula. This yields,

$$\pi_i^s(z) = \frac{z(z - \lambda)\tilde{a}_i(z) - \beta(\beta - \lambda)\tilde{a}_i(\beta)}{z - \beta} + b_i \quad i = 1, 2 \quad (3.16)$$

where the superscript s in the $\pi_i^s(z)$ functions emphasizes the fact that they pertain the a *symmetric* information, homogeneous beliefs, equilibrium. Using these in (3.10) and (3.11) we obtain the the following result

Proposition 3.5. *For any specification of observable fundamentals that satisfy Assumptions 3.1 and 3.3, the heterogeneous beliefs pricing functions can be decomposed as the following sum of the homogeneous beliefs pricing functions and an autoregressive component,*

$$\pi_i(z) = \pi_i^s(z) + \frac{b_i(\kappa_i^{-1} - 1)}{1 - \lambda\beta} \left(\frac{1 - \lambda^2}{1 - \lambda z} \right) \quad i = 1, 2 \quad (3.17)$$

where the $\pi_i^s(z)$ are given by (3.16).

This is a key result. The first term on the right-hand side of (3.17) shows how prices respond to commonly observed shocks to fundamentals. The second term then exhibits the additional dynamics induced when shocks to fundamentals are heterogeneously observed, and traders must ‘forecast the forecasts of others’. Thus, the second term captures in a clear and precise way the additional dynamics associated higher-order beliefs. These dynamics follow an AR(1) specification, independently of any autoregressive components in the observed fundamentals. The persistence is solely determined by the (noninvertible) *moving average* components of fundamentals. Interestingly, Woodford (2003) obtains a qualitatively similar result in a quite different setup.

Equation (3.17) makes clear how higher-order beliefs can generate additional price volatility. One manifestation of this is the following,

Corollary 3.6. *If the b_i coefficients are positive, heterogeneous beliefs amplify the initial response of asset prices to innovations in fundamentals.*

Proof. Evaluate (3.17) at $z = 0$. This yields

$$\pi_i(0) = \pi_i^s(0) + \frac{b_i(\kappa_i^{-1} - 1)(1 - \lambda^2)}{1 - \lambda\beta} > \pi_i^s(0)$$

¹⁶See Whiteman (1983) for a discussion of the role of regularity in delivering unique solutions of multivariate Rational Expectations models.

□

4. EMPIRICAL IMPLICATIONS

As noted in the Introduction, there are by now many papers that discuss theoretical aspects of heterogeneous belief dynamics in asset pricing. However, ours is the first to embed these dynamics within a conventional, *econometric*, asset pricing model, which allows us to explore the *quantitative* significance of heterogeneous beliefs in real world asset markets.¹⁷ Obviously, if conventional homogeneous beliefs versions of these models were successful, this would not be an interesting exercise. However, in light of the well documented failures of this model, it is of some interest to revisit these failures accounting for the possibility that heterogeneous belief dynamics are present. Thus we now ask what kind of inferential errors could result if the world is described by a heterogeneous expectations equilibrium, but an outside econometrician interprets the data as if they were generated from a homogeneous expectations equilibrium. We focus on three empirical results that have been common in the asset pricing literature: (1) violations of variance bounds, (2) predictability of excess returns, and (3) rejections of cross-equation restrictions.

Although it seems likely that heterogeneous expectations are present in *all* asset markets, we focus our attention on the U.S. stock market, since this has been the most widely studied case. Figure 1 displays annual data on real stock prices and dividends for the period 1871-2006, downloaded from Shiller’s website.

Following the original work of Shiller (1981), a common exponential trend was removed from both series. Of course, Shiller’s work unleashed a deluge of responses, many of which pointed to biases in his methods. Removing a common deterministic trend is one of them. However, we are not going to be concerned with this subsequent literature, for a couple of reasons. First, as documented in Shiller (1989), the basic message from Shiller’s original work survives these subsequent criticisms.¹⁸ Second, since the subsequent literature argued that Shiller’s methods tended to produce false rejections, if we can explain his results with heterogeneous beliefs, that only strengthens our argument. In other words, having to prove violations of a bound that are biased toward rejection makes our job harder.

4.1. Wold Representation. As noted above, in this section we are putting ourselves in the shoes of an outside econometrician. By ‘outside’ we mean the econometrician is not an active participant in the market. In particular, he does not observe any of the underlying

¹⁷Recent work by Nimark (2010) also studies the empirical significance of heterogeneous beliefs. However, his work focuses on the term structure of interest rates, and is based on a numerical approximation of the equilibrium.

¹⁸In particular, Campbell and Shiller (1987) and West (1988) develop tests that are robust to the presence of unit roots in prices and fundamentals, and continue to find evidence of excess volatility. Our methods and conclusions apply equally to their tests.

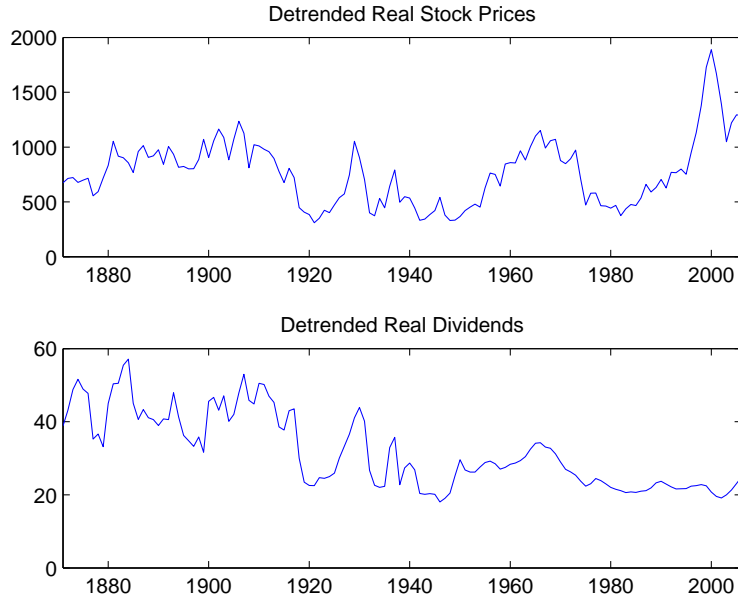


Figure 1: Data Plots

shocks driving observable fundamentals. Instead, he just witnesses the realizations of prices and the fundamentals themselves. Using this information, he wants to see whether a linear present value model can explain the data he observes, under the *assumption* that market participants have homogeneous beliefs.

One mistake he does not make is to assume the data-generating process is linear. Thus, like any econometrician his starting point is to fit VARs, or perhaps VARMA, to the data. His mistake will arise when using and interpreting these VARs. Specifically, he will make two mistakes: (1) He will misinterpret the residuals as representing innovations to the information sets of *traders*, when in fact they are not, and (2) He will incorrectly apply the law of iterated expectations when constructing estimates of traders' multiperiod forecasts.

The first step in our analysis, therefore, is to derive the Wold representation that the econometrician will estimate. Estimates of this Wold representation are the best the econometrician can do, given his (false) model and his information. Note first that deriving the econometrician's Wold representation requires a spectral factorization, since there are three underlying shocks, yet the econometrician only observes the realization of two random processes. In general, this is a messy problem. However, we can greatly simplify it by setting $\sigma_v^2 \approx 0$. This is reasonable for our purposes, since we are interested in the effects of higher-order beliefs rather than unobserved fundamentals per se. Given this, the econometrician's

Wold representation becomes¹⁹

$$\begin{bmatrix} f_t \\ p_t \end{bmatrix} = \begin{bmatrix} (1 - \lambda L)\tilde{a}_1(L) & (1 - \lambda L)\tilde{a}_2(L) \\ (1 - \lambda L)\tilde{\pi}_1(L) & (1 - \lambda L)\tilde{\pi}_2(L) \end{bmatrix} \begin{bmatrix} e_{1,t} \\ e_{2,t} \end{bmatrix} \quad (4.1)$$

where

$$e_{i,t} = \left(\frac{L - \lambda}{1 - \lambda L} \right) \varepsilon_{i,t} \quad (4.2)$$

It is natural to ask at this point what U.S. stock market data suggest about this Wold representation. Note that if the fundamentals components are purely autoregressive and $\tilde{a}_1(L) = \tilde{a}_2(L)$, then the crucial parameter, λ , can be identified from the estimated MA root of dividends, which play the role of observed fundamentals, f_t , in this case. Another possibility is that one shock dominates the other, in the sense that its variance is much larger. This occurs when one component's scale parameter $a_i(0)$ is much larger than the other's, which given the existence condition in Assumption 3.3, translates into a restriction on the relative magnitudes of the free parameters b_i . This is the identification strategy we pursue in the following analysis, since it allows us to retain a general distribution of traders across the underlying shocks. We shall see that many of the apparent failings of present value models can be explained if a large fraction of traders observe a relatively high variance shock.

Table 1 displays estimates of univariate ARMA(1,1) models for stock prices and dividends, which is what the model would predict if $\tilde{a}(L) = \tilde{a}(0)/(1 - \gamma L)$. (Higher order terms are insignificant).

TABLE 1
ARMA(1,1) ESTIMATES: ANNUAL DATA (1871-2006)

	AR(1)	MA(1)
d_t	.889 (.043)	-.197 (.092)
p_t	.865 (.050)	-.260 (.092)

Notes: (1) Estimates pertain to the model, $x_t = \frac{1-\lambda L}{1-\gamma L} e_t$
(2) Asymptotic standard errors are in parentheses.

These estimates suggest that $\lambda = -.20$ would be a reasonable value. One can see from our earlier example that the estimated price process is also predicted to be ARMA(1,1), although it is ARMA(2,2) when expressed in terms of the underlying shocks. However, the MA root will be a complicated function of λ and the other parameters, so it is easier to infer λ from dividends. The estimates in Table 1 also suggest that one cannot statistically reject

¹⁹See Hansen and Sargent (1991b) or Rozanov (1967) for details.

the model’s prediction that the AR coefficients of prices and dividends are identical. Hence, to summarize, our subsequent analysis assumes that the variance of ε_2 is vanishingly small, and then we shall impose the restriction that $\tilde{a}_1(L) = \tilde{a}_1(0)/(1 - \gamma L)$, where $\tilde{a}_1(0)$ is set to satisfy the model’s existence condition in Assumption 3.3, given an exogenously specified value of b_1 . As a benchmark, we set $\lambda = -.20$ and $\gamma = .90$.

4.2. Variance Bounds. Corollary 3.6 suggests that higher-order belief dynamics might make asset prices appear to be ‘too volatile’ relative to their fundamentals. A key contribution of our paper is the ability to quantify the degree of excess volatility associated with higher-order belief dynamics. One way of doing this is to show that heterogeneous beliefs equilibria can violate standard variance bounds inequalities. Violations of these bounds are a robust empirical finding.

Variance bounds are based on the idea that observed asset prices should be less volatile than their perfect foresight counterparts (i.e., the subsequent realization of discounted future fundamentals). Since prices represent *expectations* of discounted future fundamentals, it makes sense that they should be smoother than the realizations of discounted future fundamentals. However, the logic behind this rests on two key assumptions. First, there are no missing fundamentals.²⁰ Note, however, that the logic of variance bounds can still be applied when agents use *unobserved information* to forecast *observed fundamentals*.²¹ Since our model embodies unobserved fundamentals, in the form of forecasts of other agents’ forecasts, we will not emphasize this distinction, although when reporting estimates of cross-equation restrictions, we report variance bounds that are robust to inside information. Second, and more importantly, the logic of variance bounds is premised on the assumption that forecast errors are orthogonal to forecasts. In models with heterogeneous beliefs, this simply is not the case. Although each trader’s own individual forecast errors are of course orthogonal to his own information set, which includes prices, there is no guarantee that the *econometrician’s* forecast errors are uncorrelated with prices. Standard conditioning down arguments do not apply here, for the simple reason that with heterogeneous beliefs, the law of iterated expectations breaks down. It breaks down because traders are not just forecasting future fundamentals, they are forecasting other traders’ beliefs. These forecasts play the role of ‘missing fundamentals’.

To show that heterogeneous beliefs equilibria can violate variance bounds, it therefore suffices to show that the variance of observed prices can exceed the variance of perfect foresight prices. Although one could demonstrate this even without imposing the restriction $\text{var}(\varepsilon_2) \rightarrow 0$, the analysis becomes especially transparent when this is the case. Also, since in practice variance bounds tests are based solely on observed price and dividend data, it is

²⁰Hamilton and Whiteman (1985) emphasize this point.

²¹See, e.g., West (1988) and Campbell and Shiller (1987).

important that one derive the result using the econometrician’s Wold representation, using the observable shocks, $e_{i,t}$.

Proposition 4.1. *As $\text{var}(\varepsilon_2) \rightarrow 0$, a necessary and sufficient condition for asset prices to violate the standard variance bound is that $\beta < 2\lambda\kappa_1/(1 + \lambda^2\kappa_1^2)$.*

The proof is algebraically messy, so it is relegated to the Appendix. Note that bound violations are more likely to occur when a dominant (i.e., high variance) shock is observed by a large fraction of traders (in this case, as $\kappa_1 \rightarrow 1$). Although this is perhaps a rather special result, it is important nonetheless. Previous criticisms of Shiller’s work have been based on *statistical* problems. By way of contrast, this result is based on *economic considerations*, namely, the assumption that prices fully reveal private information, so that in equilibrium everyone makes the same forecasts. Without this assumption, the orthogonality that drives the result goes out the window.

Of course, the issue here is whether the bound is breached for plausible parameter values. Notice that violations cannot occur when $\lambda < 0$. Hence, we know our benchmark specification for dividends will satisfy the Shiller bound. Moreover, Shiller’s work suggests the bound is violated by a significant margin, by a factor of 2-5 depending on statistical assumptions. To investigate whether significant bound violations occur for reasonable parameter values we must resort to numerical simulations.²²

Figure 2 reports plots of $\text{var}(p)/\text{var}(p^{pf})$ for alternative parameter values. Since Proposition 4.1 presumes $\text{var}(\varepsilon_2) \rightarrow 0$, Figure 2 imposes the restriction $\sigma_{\varepsilon_2}^2 = 0$. Also, it is apparent that b_1 merely scales up both variances by the same amount, so for these plots the value of b_1 is irrelevant. We set it at $b_1 = 0.2$. What is important are the values of κ_1 and κ_2 . One can easily see from the necessary and sufficient condition in Proposition 4.1 that for violations to occur for empirically realistic values of β and λ , it must be the case that κ_1 exceed 0.5 by a significant margin, about $\kappa_1 > 0.8$. That is, most traders must observe the high variance shock. Figure 2 is based on the value $\kappa_1 = .99$. Finally, given annual data, we set $\beta = .90$. This is close to the original discount rates used by Shiller.

Figure 2 plots the variance ratio as a function of γ , for two different values of λ . The top panel displays the results when $\lambda = -0.20$, the value suggested by the univariate ARMA estimates for dividends. One can see that the bound is comfortably satisfied for values of γ close to the point estimates in Table 1.

In retrospect, it is not too surprising that small values of λ generate only a small amount of additional volatility. One can see why by inspecting the price function decomposition in Proposition 3.5. Notice that the *persistence* of heterogeneous beliefs is dictated by λ .

²²Note, bound violations also occur without our simplifying Assumption 3.1. Relaxing the common root restriction leads to the sufficient conditions: $\kappa_1 \approx \kappa_2$, $\beta > \lambda$, and $(1 + \lambda^2)^2 > (1 + \beta)(1 - \lambda)^2(1 + \lambda)[2(1 + \lambda^2) + (1 - \lambda)/(1 - \beta)]$. Again, these are just *sufficient* conditions.

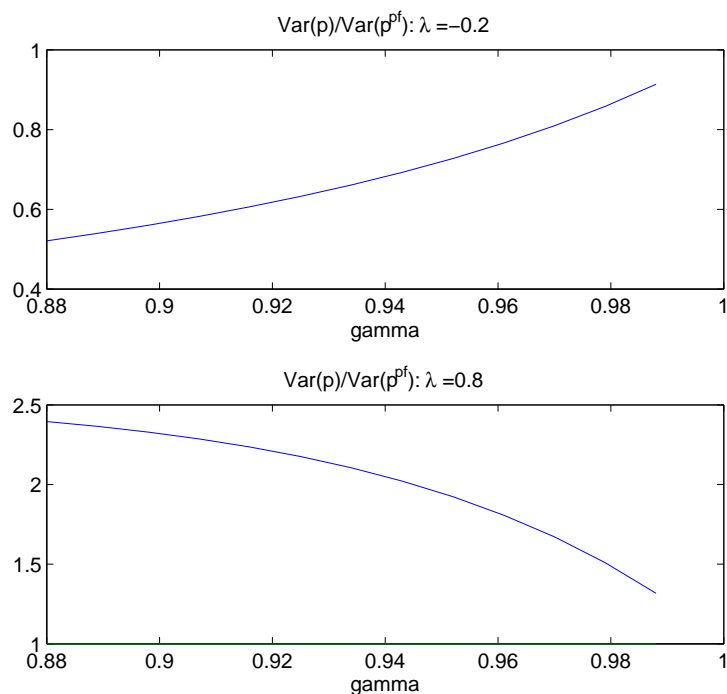


Figure 2: Shiller Bound

When λ is close to one, heterogeneous beliefs will be very persistent. Conversely, when λ is close to zero, as in the top panel of Figure 2, heterogeneous beliefs just add a small amount high frequency noise to prices. In particular, since the variance of heterogeneous beliefs is proportional to $(1 - \lambda^2)/(1 - \lambda\beta)^2$, one can see that it is an increasing function of λ for plausible values of λ and β .

This intuition is confirmed in the bottom panel of Figure 2, which displays the variance ratio when $\lambda = 0.8$. Now we see significant violations of the bound, of the same order of magnitude that Shiller found. For example, when $\gamma = .90$, its benchmark value, the variance of observed prices is more than double its hypothetical upper bound! Note this is not an artifact of biased statistical procedures, since we are comparing *population* moments. Unfortunately, when $\gamma = .9$ and $\lambda = .8$, the implied persistence of dividends is far too low. One can see, however, that significant bound violations still occur as $\gamma \rightarrow 1$. For example, when $\gamma = .98$ the variance of observed prices exceeds the bound by about 50%, and in this case dividends are fairly persistent, with a first-order autocorrelation above .55, which then damps out slowly. Although we can statistically reject this specification, it does not produce wildly implausible sample paths.

This point is reinforced in Figure 3, which compares the sample path of observed stock prices with their perfect foresight counterparts. The top panel updates Shiller's (1981) original plot.

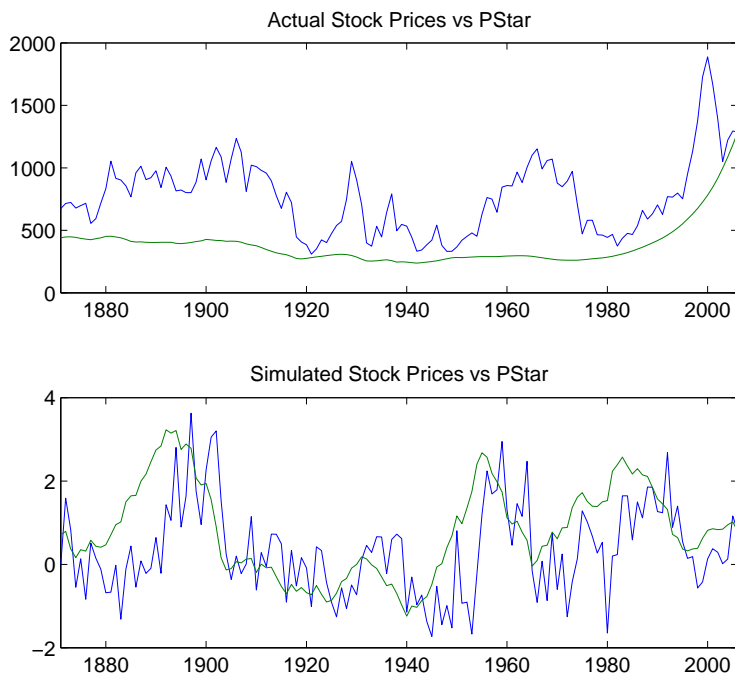


Figure 3: Stock Prices vs Present Value of Future Dividends

This graph, more than anything else, is what struck a chord with the profession (and its potential to be misleading is what motivated subsequent critics). The only difference is that we've used the observed terminal price as the end-of-sample estimate of discounted future dividends, rather than Shiller's original strategy of using the sample average price. It is now well known that using average prices produces a bias toward rejection, whereas use of the terminal price is unbiased. (See, e.g., Mankiw, Romer, and Shapiro (1985)). The bottom panel of Figure 3 follows the same procedure using data generated by our nonrevealing Rational Expectations model. All the parameters are the same as before, except now we've set $\lambda = .8$ and set $\gamma = .98$. Although it is not as striking as in the data, the plot still gives the distinct impression that prices are too volatile relative to their fundamentals.²³

4.3. Return Predictability. Another widely documented failure of linear present value models is the ability to predict excess returns, which in the case of a constant discount rate, just means the ability to predict returns themselves. Initially, excess volatility and return predictability were thought to be distinct puzzles. However, it is now well known that they are two sides of the same coin.²⁴ In fact, a finding of excess volatility can be interpreted as long (i.e., infinite) horizon return predictability. Both puzzles are driven by the violation of

²³Our conjecture is that by relaxing Assumption 3.1 we could improve the fit of the model along this dimension.

²⁴Both Cochrane (2001) and Shiller (1989) emphasize this point.

the model's implied orthogonality conditions. Still, it is useful to show how and why this occurs even in the case of one-period returns.

Of course, by construction the model's orthogonality condition in (2.1) is satisfied. The equilibrium pricing functions were computed by imposing this condition. However, this is not the condition our econometrician is testing. He is falsely assuming that everyone has the same expectation. Although *average* expectations of returns are indeed zero, our econometrician does not observe the underlying shocks that generate these expectations, so he cannot test this prediction of the model. Instead, he uses the the Wold representation in (4.1) to construct what he (falsely) believes is the 'market's' expectation of next period's price. The results are summarized as follows

Proposition 4.2. *Define the time- $t + 1$ excess return as $R_{t+1} = \beta p_{t+1} + d_t - p_t$. Then as $\text{var}(\varepsilon_2) \rightarrow 0$ the Wold representation in (4.1) generates the following projection of R_{t+1} onto the (econometrician's) time- t information set*

$$R_{t+1} = b_1 \lambda \left(\frac{\tilde{a}_1^{-1}(L)}{1 - \lambda L} \right) d_t.$$

Again, the proof is by construction. As $\text{var}(\varepsilon_2) \rightarrow 0$ and after several simplifications to (3.10), we can write the equilibrium price as follows

$$p_t = (L - \lambda) \left[\frac{L\tilde{a}_1(L) - \beta\tilde{a}_1(\beta)}{L - \beta} + \frac{\kappa_2}{\kappa_1} \frac{b_1 \lambda}{1 - \lambda\beta} \frac{1}{1 - \lambda L} \right] \varepsilon_{1t}$$

Expressed in terms of the Wold representation we have

$$\begin{aligned} p_t &= (1 - \lambda L) \left[\frac{L\tilde{a}_1(L) - \beta\tilde{a}_1(\beta)}{L - \beta} \right] e_{1,t} + \frac{\kappa_2}{\kappa_1} \frac{b_1 \lambda}{1 - \lambda\beta} e_{1,t} \\ &\equiv A(L)e_{1,t} + \frac{\kappa_2}{\kappa_1} \frac{b_1 \lambda}{1 - \lambda\beta} e_{1,t} \end{aligned}$$

where $e_{1,t} = [(L - \lambda)/(1 - \lambda L)]\varepsilon_{1t}$. Therefore,

$$\begin{aligned} \beta E_t p_{t+1} + d_t - p_t &= \left\{ \beta L^{-1}[A(L) - A(0)] + (1 - \lambda L)\tilde{a}_1(L) - [A(L) + \kappa_2 b_1 / \kappa_1 (1 - \lambda\beta)] \right\} e_{1,t} \\ &= \lambda b_1 e_{1,t} \end{aligned}$$

The result now follows from the fact that $d_t \rightarrow (1 - \lambda L)\tilde{a}_1(L)e_{1,t}$ as $\text{var}(\varepsilon_2) \rightarrow 0$.

Proposition 4.2 suggests that if we regress R_t onto lagged information we should find statistically significant coefficients. (The result is stated in terms of lagged dividends, but it could just as easily have been stated in terms of lagged prices or returns, given their equilibrium relationship). Table 2 contains a small set of results from these kinds of regressions, using Shiller's annual data. (Since in Shiller's data prices are sampled in January and dividends accrue throughout the year, we do not assume time- t dividends are in the time- t information set).

TABLE 2
RETURN PREDICTABILITY

	R_{t-1}	R_{t-2}	D_{t-1}	D_{t-2}	\bar{R}^2
<u>Data</u>					
$R_t =$.344 (.195)				.111
$R_t =$.398 (.166)	-.160 (.141)			.127
$R_t =$			3.19 (1.29)		.034
$R_t =$			12.9 (4.58)	-10.2 (5.10)	.046
<u>Model</u>					
$R_t =$	-.504 (.075)				.246
$R_t =$	-.681 (.081)	-.354 (.082)			.338
$R_t =$			2.81 (.459)		.214
$R_t =$			-3.72 (.493)	1.90 (.495)	.291

Notes: (1) All regressions include a constant.

(2) Heteroskedasticity-robust standard errors in parentheses.

The results here are quite consistent with previous results. There is some modest evidence in favor of predictability, but it is not overwhelming. Evidence of predictability is somewhat stronger using model-simulated data. There are of course a plethora of statistical pitfalls associated with these regressions, but again, our results suggest that even if one had access to the *population* second moments, we should still observe predictability. Finding that returns are predictable is not a puzzle if investors have heterogeneous expectations of returns.

4.4. Cross-Equation Restrictions. The Rational Expectations revolution ushered in many methodological changes. One of the most important concerned the way econometricians identify their models. Instead of producing zero restrictions, the Rational Expectations Hypothesis produces cross-equation restrictions. Specifically, parameters describing the laws of motion of exogenous forcing processes enter the laws of motion of endogenous decision processes. In fact, in an oft-repeated phrase, Sargent dubbed these restrictions the ‘hallmark of Rational Expectations’. Hansen and Sargent (1991b) and Campbell and Shiller (1987)

proposed useful procedures for testing these restrictions. When these tests are applied to present value asset pricing models, they are almost without exception rejected, and in a resounding way. There have been many responses to these rejections. Some interpret them as evidence in favor of stochastic discount factors. Others interpret them as evidence against the Rational Expectations Hypothesis. We offer a different response. We show that rejections of cross-equation restrictions may simply reflect an informational misspecification, one that presumes a revealing equilibrium and homogeneous beliefs when in fact markets are characterized by heterogeneous beliefs.

For reference purposes we repeat the Wold Representation given by equation (4.1)

$$\begin{bmatrix} f_t \\ p_t \end{bmatrix} = \begin{bmatrix} (1 - \lambda L)\tilde{a}_1(L) & (1 - \lambda L)\tilde{a}_2(L) \\ K(\tilde{a}_1(L)) & K(\tilde{a}_2(L)) \end{bmatrix} \begin{bmatrix} e_{1,t} \\ e_{2,t} \end{bmatrix} \quad (4.3)$$

where we now write the pricing function $(1 - \lambda L)\tilde{\pi}(L)$ as $K(\tilde{a}(L))$ to emphasize the fact that it is the output of a linear operator, $K(\cdot)$, defined by equations (3.10) and (3.11). Note that the heterogeneous beliefs pricing operator actually consists of the sum of *two* linear operators, $K = K^s + K^h$, where K^s denotes the conventional symmetric information operator, given by the Hansen-Sargent formula, and K^h denotes the heterogeneous beliefs operator defined in Proposition 3.5. This gives rise to the following result,

Proposition 4.3. *Standard cross-equation restriction tests, which falsely presume a common information set, can produce spurious rejections.*

Proof. When there are heterogeneous beliefs, $\pi_i(L) = K(a_i(L))$, where $K = K^s + K^h$. Cross-equation restriction tests based on the false assumption of homogeneous beliefs amount to dropping the K^h component of the pricing operator. This can produce strong rejections when $K^h(a_i(L))$ is ‘big’ (in the operator sense). \square

Again, the real issue here is the quantitative significance of this result. Although one could perhaps investigate this analytically, it is simpler to just perform a simulation. Table 3 contains two sets of results. The top panel replicates the VAR testing strategy of Campbell and Shiller (1987) using updated data from Shiller’s website.²⁵ Given the annual frequency, a VAR(1) appears adequate. The final three columns report the outcomes of various tests and diagnostics. The $\chi^2(2)$ column reports the Wald statistic for the model’s two cross-equation restrictions. As many others have found, these restrictions are strongly rejected. The $\text{var}(P)/\text{var}(\hat{P})$ column reports the ratio

²⁵In contrast to Campbell and Shiller (1987), we do not assume unit roots and cointegration. To maintain consistency with our previous results we use detrended data.

TABLE 3
CROSS-EQUATION RESTRICTIONS: ANNUAL DATA (1871-2006)

<u>Data</u>	D_{t-1}	P_{t-1}	\bar{R}^2	DW	$\chi^2(2)$	$\text{var}(P)/\text{var}(\hat{P})$	$\text{corr}(P, \hat{P})$
$D_t =$.912 (.036)	.001 (.001)	.842	1.72	11.8	38.8	.47
$P_t =$	-1.28 (1.16)	.927 (.040)	.814	1.72			
<u>Simulation</u>							
$D_t =$.425 (.100)	.004 (.006)	.211	2.31	91.4	451.4	.72
$P_t =$	3.85 (1.65)	.128 (.107)	.106	2.14			

Notes: (1) $\chi^2(2)$ is the Wald statistic for the cross-equation restrictions $(0, 1)[I - \beta\psi] = (1, 0)\beta\psi$, where ψ is the VAR(1) coefficient matrix.

(2) $\hat{P} \equiv$ Expected present discounted value of dividends with $\beta = .90$.

(3) Asymptotic standard errors in parentheses.

between the variance of observed prices and the variance of predicted prices, using the VAR to construct the present discounted value of future dividends. Under the null, this ratio should be one. The point estimate suggests even a stronger rejection than the earlier variance bound results, which is somewhat surprising in light of the fact that this estimate is robust to presence of inside information, which would tend to make observed prices more volatile than expected. The final column reports the sample correlation coefficient between actual and predicted prices. As Campbell and Shiller (1987) emphasized, even though the model is strongly rejected statistically, it does have some ability to track observed prices.

The bottom panel reports the results from following the exact same procedures using data generated by the model, with the parameter values $\kappa_1 = .99$, $\lambda = .80$, and $\gamma = .98$. Interestingly, applying the Campbell-Shiller method to the model leads to even stronger rejections. The strange looking coefficient estimates in the price equation arise from the near ‘exactness’ of the model. With lagged dividends included, prices *should* have little additional explanatory power for future prices. As in the data, the correlation between the model’s predicted price and the actual price is fairly high. Instead, the failure in both cases stems from the excessive volatility of prices. This can be seen in Figure 4, which plots actual versus predicted prices for both observed data and model-generated data.

Note that the top half of Figure 4 is nearly identical to the top half of Figure 3. On the other hand, the bottom portions of these two Figures, pertaining to model-generated data,

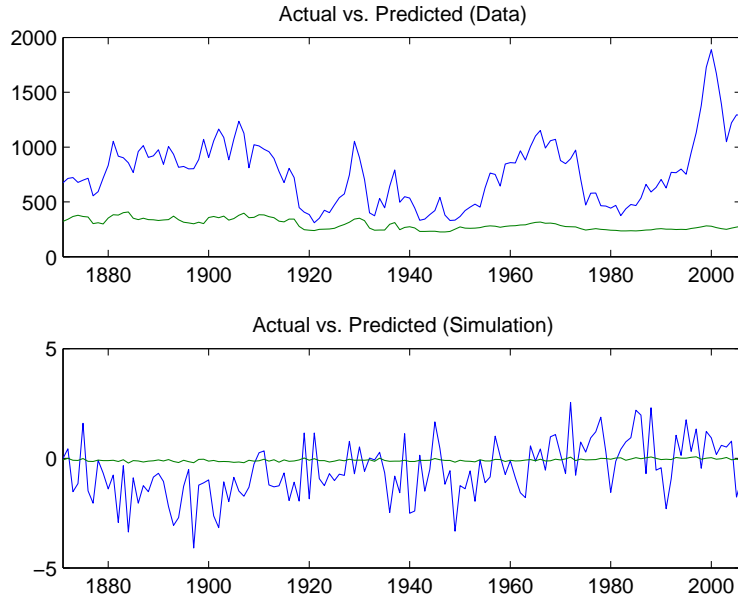


Figure 4: Actual vs. Predicted Stock Prices

are quite different. According to the VAR, dividends are not very persistent or forecastable. As a result, predicted prices, which are based on VAR forecasts of future dividends, are nearly flat.

It is noteworthy that we continue to reject the model despite using procedures that are robust to the possibility that agents have more information than the econometrician. This kind of asymmetric information is not the issue here. Rather it is the presence of asymmetric information among the agents themselves that is the source of the problem. When the agents themselves have asymmetric information, prices are determined by *average* expectations, and these averaged expectations do not adhere to the law of iterated expectations. Unfortunately, the clever VAR procedures of Hansen and Sargent (1991a) and Campbell and Shiller (1987) rely heavily on the law of iterated expectations.

5. CONCLUSION

For more than thirty years now, economists have been rejecting linear present value asset pricing models. These rejections have been interpreted as evidence in favor of time-varying risk premia. Unfortunately, linking risk premia to observable data has been quite challenging. Promising approaches for meeting the challenge involve introducing incomplete markets and agent heterogeneity into models.

This paper has suggested that a different sort of heterogeneity, an *informational heterogeneity*, offers an equally promising route toward reconciling asset prices with observed fundamentals. Unfortunately, heterogeneous information does not automatically translate into heterogeneous beliefs, and it is only the latter that generates the ‘excess volatility’ that

is so commonly seen in the data. The hard work in the analysis, therefore, is deriving the conditions that prevent market data from fully revealing the private information of agents in dynamic settings. We have argued that frequency-domain methods possess distinct advantages over time-domain methods in this regard. The key to keeping information from leaking out through observed asset prices is to ensure that the mappings between the two are ‘noninvertible’. These noninvertibility conditions are easy to derive and manipulate in the frequency domain.

Our results demonstrate how informational heterogeneity can in principle explain well-known empirical anomalies, such as excess volatility, excess return predictability, and rejections of cross-equation restrictions. Ever since Townsend (1983) and Singleton (1987), (or in fact, ever since Keynes!) economists have suspected that heterogeneous beliefs could be responsible for the apparent excess volatility in financial markets. Our results at last confirm these suspicions. Although we believe we have made substantial progress, there are still many avenues open for future research. Two seem particularly important. First, like the recent work of Engel, Mark, and West (2007), our paper offers some hope for linear present value models. Unlike their work, however, which is largely based on statistical and calibration issues, our paper points to a more radical reorientation of VAR methodology. In particular, it would be useful to develop and implement empirical procedures that are robust to heterogeneous beliefs, and perhaps even develop statistical tests that could reliably detect their presence. Second, the entire analysis here rests heavily on linearity. However, most macroeconomic models feature nonlinearities of one form or another. It is not at all clear whether standard linearization methods are applicable in models featuring heterogeneous beliefs. Resolving this issue will also be important for future applications.

APPENDIX A. PROOF OF PROPOSITION 4.1

The proof is by brute force computation. Let p_t^{pf} denote the perfect foresight price. It is given by

$$p_t^{pf} = \sum_{j=0}^{\infty} \beta^j f_{t+j} = \frac{L\tilde{a}_1(L)(1-\lambda L)}{(L-\beta)}\varepsilon_{1,t} + \frac{L\tilde{a}_2(L)(1-\lambda L)}{(L-\beta)}\varepsilon_{2,t}$$

As $\text{var}(\varepsilon_2) \rightarrow 0$, the econometrician's price process becomes

$$\begin{aligned} p_t &= (1-\lambda L) \left[\frac{L\tilde{a}_1(L) - \beta\tilde{a}_1(\beta)}{L-\beta} \right] \varepsilon_{1,t} + \frac{\kappa_2}{\kappa_1} \frac{b_1\lambda}{1-\lambda\beta} \varepsilon_{1,t} \\ &= p_t^{pf} - \beta\tilde{a}_1(\beta) \frac{1-\lambda L}{L-\beta} \varepsilon_{1,t} + \frac{\kappa_2}{\kappa_1} \frac{b_1\lambda}{1-\lambda\beta} \varepsilon_{1,t} \end{aligned}$$

Using the existence condition in Assumption 3.3 to substitute out $\beta\tilde{a}_1(\beta)$ we can write this as,

$$p_t = p_t^{pf} + \frac{b_1(\kappa_1^{-1} - \lambda L)}{L-\beta} \varepsilon_{1,t}$$

Using the residue calculus, we have

$$\begin{aligned} \text{var}(p) &= \text{var}(p^{pf}) + b_1^2 \oint \frac{(\kappa_1^{-1} - \lambda z)}{z-\beta} \frac{(\kappa_1^{-1} - \lambda z^{-1})}{z^{-1}-\beta} \frac{dz}{z} \\ &+ 2b_1 \oint \left(\frac{z\tilde{a}_1(z)(1-\lambda z)}{z-\beta} \right) \left(\frac{\kappa_1^{-1} - \lambda z^{-1}}{z^{-1}-\beta} \right) \frac{dz}{z} \\ &= \text{var}(p^{pf}) + \frac{b_1^2}{\beta(1-\beta^2)} [2\lambda\kappa_1^{-1} - \kappa_1^{-2}\beta - \lambda^2\beta] \end{aligned}$$

From this it is clear that $\text{var}(p) > \text{var}(p^{pf})$ iff $\beta < 2\lambda\kappa_1/(1 + \lambda^2\kappa_1^2)$.

APPENDIX B. FREQUENCY DOMAIN TECHNIQUES (NOT FOR PUBLICATION)

This appendix offers a brief introduction to the frequency domain techniques used to solve the model. Rather than match an infinite sequence of unknown coefficients, we employ the following theorem and solve for a fixed point in a function space.

Theorem (Riesz-Fischer): *Let $\{c_n\}$ be a square summable sequence of complex numbers (i.e., $\sum_{n=-\infty}^{\infty} |c_n|^2 < \infty$). Then there exists a complex-valued function, $g(\omega)$, defined for $\omega \in [-\pi, \pi]$, such that*

$$g(\omega) = \sum_{j=-\infty}^{\infty} c_j e^{-i\omega j} \tag{B.1}$$

where convergence is in the mean-square sense

$$\lim_{n \rightarrow \infty} \int_{-\pi}^{\pi} \left| \sum_{j=-n}^n c_j e^{-i\omega j} - g(\omega) \right|^2 d\omega = 0$$

and $g(\omega)$ is square (Lebesgue) integrable

$$\int_{-\pi}^{\pi} |g(\omega)|^2 d\omega < \infty$$

Conversely, given a square integrable $g(\omega)$ there exists a square summable sequence such that

$$c_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} g(\omega) e^{i\omega k} d\omega \quad (\text{B.2})$$

The Fourier transform pair in (B.1) and (B.2) defines an isometric isomorphism (i.e., a one-to-one onto transformation that preserves distance and linear structure) between the space of square summable sequences, $\ell^2(-\infty, \infty)$, and the space of square integrable functions, $L^2[-\pi, \pi]$. The sequence space, ℓ^2 , is referred to as the ‘time domain’ and the function space, L^2 , is referred to as the ‘frequency domain’. The equivalence between these two spaces allows us to work in whichever is most convenient. A basic premise of this paper is that in models featuring higher-order beliefs, the frequency domain is analytically more convenient.

In the context of linear prediction and signal extraction, it is useful to work with a version of Riesz-Fischer theorem that is generalized in one sense and specialized in another. In particular, it is possible to show, via Poisson’s integral formula, that the statement of the theorem applies not only to functions defined on an interval (the boundary of the unit circle), but to analytic functions defined within the entire unit circle of the complex plane. However, when extending the theorem in this way we exclude functions with Fourier coefficients that are nonzero for negative k . That is, we limit ourselves to functions where $c_{-k} = 0$ in equations (B.1) and (B.2). This turns out to be useful, since it is precisely these functions that represent the ‘past’ in the time domain. A space of analytic functions in the unit disk defined in this way is called a Hardy space, with an inner product defined by the contour integral,

$$(g_1, g_2) = \frac{1}{2\pi i} \oint g_1(z) \overline{g_2(z)} \frac{dz}{z}.$$

Rather than postulate a functional form and match coefficients, we solve for a single analytic function which represents, in the sense of the Riesz-Fischer theorem, this unknown pricing function. The approach is still ‘guess and verify’, but it takes place in a function space, and it works because the Riesz-Fischer theorem tells us that two stochastic processes are ‘equal’ (in the sense of mean-squared convergence) if and only if their z -transforms are identical as analytic functions inside the (open) unit disk. The real advantage of this approach stems from the ease with which it handles noninvertibility (i.e., nonrevealing information) issues. Invertibility hinges on the absence of zeros inside the unit circle of the

z -transform of the observed market data. By characterizing these zeros, we characterize the information revealing properties of the equilibrium.

APPENDIX C. MICROFOUNDATIONS OF THE MODEL (NOT FOR PUBLICATION)

This appendix develops a microfounded justification for the present value model of section 2. The model is a stylized noisy rational expectations model that is standard in the asset pricing literature. We examine an infinite horizon trading model where time is discrete and indexed by $t = 0, 1, 2, \dots$

C.1. Assets. There is a risky asset (stock) and riskless asset (bond) that is traded at each date. The riskless asset is in perfectly elastic supply with rate of return $1 + r_t \forall t$. The stock pays a dividend with value f_t . Shares of the stock are infinitely divisible and traded competitively. Following the standard assumption in the literature, we assume that the number of shares available to the market is random, u_t . This assumption follows the usual noise trading story in which a fraction of the traders are liquidity traders with inelastic demand of $1 - u_t$ shares of stock at t , leaving u_t shares to be traded (normalizing the total shares to one). We assume that the stochastic processes u_t and f_t are all stationary and Gaussian.

C.2. Investors. We assume all traders are price takers in that they are not large enough to influence the price. We assume investors submit demand schedules according to the linear trading rule

$$X_t^i = E[Q_{t+1}|F_t^i], \quad Q_{t+1} = (1 + r)f_t + p_{t+1} - (1 + r)p_t \quad (\text{C.1})$$

where F_t^i is the information set of trader i at t , p_t is the price of the stock, and Q_{t+1} is the excess return of the stock. Note that this demand can be derived from the usual trading rule assuming trader i chooses the amount of stock to purchase in accordance with a CARA preference structure over wealth. The difference is that we have assumed the coefficient of risk aversion and the conditional variance term are normalized to unity. Setting demand equal to the stochastic supply delivers (2.1).

REFERENCES

- ALBAGLI, E., C. HELLWIG, AND A. TSYVINSKI (2011): “A Theory of Asset Prices Based on Heterogeneous Information,” Cowles Foundation Discussion Paper 1827.
- ALLEN, F., S. MORRIS, AND H. S. SHIN (2006): “Beauty Contests and Iterated Expectations in Asset Markets,” *Review of Financial Studies*, 19, 719–752.
- ANDERSON, E., E. GHYSELS, AND J. JUERGENS (2005): “Do Heterogeneous Beliefs Matter for Asset Pricing?,” *Review of Financial Studies*, 18(3), 875.
- ANGELETOS, G.-M., AND I. WERNING (2006): “Crises and Prices: Information Aggregation, Multiplicity and Volatility,” *American Economic Review*, 96, 1720–36.
- ATKESON, A. (2000): “Comment on ‘Rethinking Multiple Equilibria in Macroeconomic Modeling’,” in *NBER Macroeconomics Annual 2000*, ed. by B. S. Bernanke, and K. Rogoff. MIT Press.
- BACCHETTA, P., AND E. VAN WINCOOP (2006): “Can Information Heterogeneity Explain the Exchange Rate Determination Puzzle?,” *American Economic Review*, 96(3), 552–576.
- (2008): “Higher Order Expectations in Asset Pricing,” *Journal of Money, Credit, and Banking*, 40(5), 837–866.
- BERNHARDT, D., P. SEILER, AND B. TAUB (2010): “Speculative Dynamics,” *Economic Theory*, 44, 1–52.
- BRUNNERMEIER, M. K. (2001): *Asset Pricing under Asymmetric Information: Bubbles, Crashes, Technical Analysis, and Herding*. Oxford University Press, Oxford.
- CAMPBELL, J. Y., AND R. J. SHILLER (1987): “Cointegration and Tests of Present Value Models,” *Journal of Political Economy*, 95, 1062–1088.
- COCHRANE, J. H. (2001): *Asset Pricing*. Princeton University Press, Princeton.
- CONSTANTINIDES, G. M., AND D. DUFFIE (1996): “Asset Pricing with Heterogeneous Consumers,” *Journal of Political Economy*, 104, 219–240.
- ENGEL, C., N. MARK, AND K. D. WEST (2007): “Exchange Rate Models Are Not As Bad As You Think,” *NBER Macroeconomics Annual*, 22, 381–441.
- ENGEL, C., AND K. D. WEST (2005): “Exchange Rates and Fundamentals,” *Journal of Political Economy*, 113(3), 485–517.
- FOSTER, F. D., AND S. VISWANATHAN (1996): “Strategic Trading when Agents Forecast the Forecasts of Others,” *Journal of Finance*, 51(4), 1437–1478.
- FUTIA, C. A. (1981): “Rational Expectations in Stationary Linear Models,” *Econometrica*, 49(1), 171–192.
- GROSSMAN, S. J., AND J. E. STIGLITZ (1980): “On the Impossibility of Informationally Efficient Markets,” *American Economic Review*, 70, 393–408.
- HAMILTON, J. D., AND C. H. WHITEMAN (1985): “The Observable Implications of Self-Fulfilling Expectations,” *Journal of Monetary Economics*, 16, 353–373.
- HANSEN, L. P., AND T. J. SARGENT (1991a): “Exact Linear Rational Expectations Models: Specification and Estimation,” in *Rational Expectations Econometrics*, ed. by L. P. Hansen, and T. J. Sargent, pp. 45–76. Westview Press.
- (1991b): “Two Difficulties in Interpreting Vector Autoregressions,” in *Rational Expectations Econometrics*, ed. by L. P. Hansen, and T. J. Sargent, pp. 77–119. Westview Press.
- HARRISON, J. M., AND D. M. KREPS (1978): “Speculative Investor Behavior in a Stock Market with Heterogeneous Expectations,” *Quarterly Journal of Economics*, 89, 323–36.
- HE, H., AND J. WANG (1995): “Differential Information and Dynamic Behavior of Stock Trading Volume,” *Review of Financial Studies*, 8, 919–972.

- HELLWIG, C., A. MUKHERJI, AND A. TSYVINSKI (2006): "Self-Fulfilling Currency Crises: The Role of Interest Rates," *American Economic Review*, 96, 1769–1787.
- KASA, K. (2000): "Forecasting the Forecasts of Others in the Frequency Domain," *Review of Economic Dynamics*, 3, 726–756.
- MAKAROV, I., AND O. RYTCHKOV (2008): "Forecasting the Forecasts of Others: Implications for Asset Pricing," Working Paper, MIT.
- MANKIW, N. G., D. ROMER, AND M. D. SHAPIRO (1985): "An Unbiased Reexamination of Stock Market Volatility," *Journal of Finance*, 40, 677–87.
- MILGROM, P., AND N. STOKEY (1982): "Information, Trade and Common Knowledge," *Journal of Economic Theory*, 26, 17–27.
- MORRIS, S. (1996): "Speculative Investor Behavior and Learning," *Quarterly Journal of Economics*, 111, 1111–33.
- MORRIS, S., AND H. S. SHIN (2003): "Global Games: Theory and Applications," in *Advances in Economics and Econometrics: Theory and Applications*, ed. by S. T. M. Dewatripont, L.P. Hansen. Cambridge University Press, Cambridge.
- NIMARK, K. P. (2007): "Dynamic Higher Order Expectations," Working Paper, Research Department, Reserve Bank of Australia.
- (2010): "Speculative Dynamics in the Term Structure of Interest Rates," Working Paper, Universitat Pompeu Fabra.
- PEARLMAN, J. G., AND T. J. SARGENT (2005): "Knowing the Forecasts of Others," *Review of Economic Dynamics*, 8(2), 480–497.
- PIAZZESI, M., AND M. SCHNEIDER (2009): "Momentum Traders in the Housing Market: Survey Evidence and a Search Model," *American Economic Review*, 99(2), 406–411.
- RONDINA, G., AND T. B. WALKER (2011): "Information Equilibria in Dynamic Economies with Dispersed Information," Working Paper, U.C. San Diego.
- ROZANOV, Y. A. (1967): *Stationary Random Processes*. Holden-Day, San Francisco.
- SHILLER, R. J. (1981): "Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?," *American Economic Review*, 71(3), 421–436.
- (1989): *Market Volatility*. MIT Press.
- SINGLETON, KENNETH, J. (1987): "Asset Prices in a Time Series Model with Disparately Informed, Competitive Traders," in *New Approaches to Monetary Economics*, ed. by W. Barnett, and K. Singleton. Cambridge University Press, Cambridge.
- TAUB, B. (1990): "The Equivalence of Lending Equilibria and Signalling-Based Insurance under Asymmetric Information," *Rand Journal of Economics*, 21, 388–408.
- TIROLE, J. (1982): "On the Possibility of Speculation under Rational Expectations," *Econometrica*, 50, 1163–82.
- TOWNSEND, R. M. (1983): "Forecasting the Forecasts of Others," *Journal of Political Economy*, 91, 546–588.
- WALKER, T. B. (2007): "How Equilibrium Prices Reveal Information in Time Series Models with Disparately Informed, Competitive Traders," *Journal of Economic Theory*, 137, 512–537.
- WANG, J. (1993): "A Model of Asset Prices under Asymmetric Information," *Review of Economic Studies*, 60(2), 249–282.
- WEST, K. D. (1988): "Dividend Innovations and Stock Price Volatility," *Econometrica*, 56, 37–61.

- WHITEMAN, C. (1983): *Linear Rational Expectations Models: A User's Guide*. University of Minnesota Press, Minneapolis.
- WOODFORD, M. (2003): "Imperfect Common Knowledge and the Effects of Monetary Policy," in *Knowledge, Information, and Expectations in Modern Macroeconomics*, ed. by P. Aghion, R. Frydman, J. Stiglitz, and M. Woodford. Princeton University Press, Princeton, N.J.