

HETEROGENEOUS BELIEFS AND TESTS OF PRESENT VALUE MODELS*

Kenneth Kasa[†] Todd B. Walker[‡] Charles H. Whiteman[§]

November 2013

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 remain asymmetrically informed in equilibrium and must 'forecast the forecasts of others'. An econometrician, who incorrectly imposes a homogeneous beliefs equilibrium, will find that the asset price displays violations of variance bounds, predictability of excess returns, and rejections of cross-equation restrictions.

Keywords: Asset Pricing, Heterogeneous Beliefs, Excess Volatility
JEL Classification Numbers: G12, D82

*We would like to thank Imran Rasul and four anonymous referees for very useful comments and suggestions. We would also like to thank Tim Cogley, Martin Evans, Ron Michener, Bart Taub, Tom Sargent, 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*.

[†]Department of Economics, Simon Fraser University, kkasa@sfu.ca

[‡]Department of Economics, Indiana University, walkertb@indiana.edu

[§]Smeal College of Business, Penn State University, cwhiteman@smeal.psu.edu

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 a result, linear present value models have all but disappeared from serious academic research on asset pricing.¹ Attention has shifted to models with time-varying risk premia and other formulations that jettison the assumptions of a linear, constant discount rate model. Unfortunately, these models offer little improvement empirically.²

Our paper returns to the linear constant discount rate setting and argues that *informational* heterogeneity can account for many of the model's apparent empirical shortcomings. In particular, we make a simple change to the standard present value model: We assume speculative traders are heterogeneously informed about the observable fundamentals. Observed fundamentals consist of a sum of orthogonal components, and in addition to observing the sum, we assume each trader observes realizations of one of the underlying components. Coupled with additional noise that serves to break the no trade theorems of Milgrom and Stokey (1982) and Tirole (1982), we derive conditions under which this information structure preserves heterogeneous beliefs in a dynamic equilibrium. 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.

Following Futia (1981), we derive conditions under which the traders remain asymmetrically informed in equilibrium. Traders are only able to infer a weighted average of the other traders' innovations. 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 equilibrium ([Theorem 1](#)). This allows us to examine, in closed form, how heterogeneous beliefs alter equilibrium dynamics. A closed-form solution to a dynamic model with persistent heterogeneous beliefs is a non-trivial contribution. Much of the literature continues to follow the approximation of Townsend (1983), which assumes traders share information with a one- or two-period lag. This substantially mitigates the role of heterogeneous beliefs in dynamic models. [ii] We show how the additional dynamics of the heterogeneous beliefs equilibrium can play a crucial role in observed asset prices.³ More specifically, we show that excess volatility, predictability, and the rejection of cross-equation restrictions can all be reconciled with theory when asset price dynamics follow the persistent heterogeneous-beliefs equilibrium and the econometrician incorrectly imposes homogeneous-beliefs restrictions.

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.).

²Constantinides and Duffie (1996) achieve some success by introducing heterogeneity, in the form of nondiversifiable labor income risk.

³We focus on stock prices but see Piazzesi and Schneider (2009) and Xiong and Yan (2010) for applications to the housing and the bond markets, respectively.

Of course, this is not the first paper to study asymmetric information in asset markets.⁴ However, our paper is the first to combine several key ingredients. First, our model features *persistent* heterogeneous beliefs. We follow the “noisy-rational expectations” approach of Grossman and Stiglitz (1980) and Hellwig (1980) in which heterogeneous beliefs persist due to agents receiving different signals about underlying fundamentals. Noise traders serve to obscure the information coming from endogenous variables sufficiently so that the traders remain differentially informed in equilibrium.

An alternative way to preserve asymmetric information in equilibrium is to assume agents have heterogeneous prior beliefs about an unobservable economic variable or the informativeness of a signal. Difference-of-opinion models [Harris and Raviv (1993), Morris (1996), Detemple and Murthy (1994)] assume that agents receive the same information but “agree to disagree” about some fundamental aspect of the model (e.g., model specification or parameters). Heterogeneous priors can also arise from assuming agents have different prior knowledge about the informativeness of a signal [Scheinkman and Xiong (2003), Lam, Cecchetti, and Mark (2000), Dumas, Kurshev, and Uppal (2009), Xiong and Yan (2010)]. Acemoglu, Chernozhukov, and Yildiz (2008) show that when agents disagree about the informativeness of a signal of a random variable, then even an infinite sequence of signals will not lead to a convergence of beliefs. Heterogeneous prior models typically allow for closed-form solutions but this tractability comes at the cost of assuming agents do not learn from other agents’ actions through the price channel. Since a primary focus of this paper will be on understanding how these higher-order beliefs could lead to violations of variance bounds and predictability of asset prices, we do not want to abstract from this particular mechanism.⁵

Second, our model is dynamic and stationary. Following Grossman and Stiglitz (1980), most work using the noisy-rational expectations approach 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 (1994), He and Wang (1995), Foster and Viswanathan (1996), Albagli, Hellwig, and Tsyvinski (2011)]. Much 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 heterogeneous beliefs and link it directly to the empirical anomalies found in the asset pricing literature.

Third, our approach features *signal extraction from endogenous prices*. This distinguishes our paper from the work on global games with imperfect common knowledge and heterogeneous prior models [Harris and Raviv (1993), Morris and Shin (2003), Lam, Cecchetti, and Mark (2000), Dumas, Kurshev, and Uppal (2009), Xiong and Yan (2010)]. Although the global games literature has made important contributions to our understanding of higher-order beliefs, it is not directly applicable to asset pricing. As Atkeson (2000) notes, prices play an important role in aggregating information,

⁴See Brunnermeier (2001) for a review.

⁵Interestingly Banerjee, Kaniel, and Kremer (2009) argue that price drift in a heterogeneous prior model cannot be accomplished without higher-order difference of opinion.

and it remains to be seen how robust the work on global games is to the inclusion of asset markets.⁶ Similar in spirit to our paper, heterogeneous prior models have been successful at replicating excess volatility and other asset pricing anomalies, through the use of heterogeneous beliefs and sentiment shocks. But this approach does not allow signal extraction from prices. Rather than learning about other agents actions in equilibrium, agents either agree to disagree or do not condition on endogenous variables. Here, agents extract information from prices, which we view as an important extension of the literature.

Finally, 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 (2011)]. However, recent work by Pearlman and Sargent (2005) and Walker (2007) demonstrates that numerical approaches can be misleading.

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.⁷ 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 heterogenous beliefs are preserved in equilibrium.⁸

⁶Two notable exceptions are Angeletos and Werning (2006) and Hellwig, Mukherji, and Tsyvinski (2006), who incorporate signal extraction from prices into the Morris-Shin framework. These models differ from our setup in that agents must choose “actions” in a coordination game. Their models focus primarily on the issue of equilibrium uniqueness and are essentially static.

⁷Kasa (2000) uses frequency domain methods to solve the model of Townsend (1983). He shows, along with Pearlman and Sargent (2005), that the equilibrium of Townsend is fully revealing and therefore there is no need to forecast the forecasts of other agents. Here we derive conditions that guarantee non-revealing equilibria and isolate the higher-order belief components. The online appendix provides a review of the Riesz-Fischer Theorem.

⁸Makarov and Rytchkov (2012), Bernhardt, Seiler, and Taub (2010) and Rondina and Walker (2012) also use

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 (1991) 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 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

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

where time is discrete and indexed by $t = 0, 1, 2, \dots$; there is a continuum of investors on the unit interval indexed by i , 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), and u_t represents the influence of unobserved fundamentals (e.g., noise or liquidity traders). The parameter, $\beta < 1$, is a constant discount factor. The model is a stylized noisy rational expectations model that is standard in the asset pricing literature. It is a special case of Lucas (1978) (with risk-neutral investors and assuming shares are traded cum-dividend) and has the following micro-foundations.

frequency domain techniques to solve dynamic models with heterogeneously informed agents. Makarov and Rytchkov (2012) 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 (2012) extend Futia (1981) to the case of dispersed information.

2.1 MICRO-FOUNDATIONS 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 $1 + r$. The stock pays a dividend with value f_t . Shares of the stock are infinitely divisible and traded competitively. Following a 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 random 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). 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)].⁹ These noise traders only purpose is to break the no-trade theorem. We assume that the stochastic processes u_t and f_t are all stationary and Gaussian.

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

$$X_t^i = \mathbb{E}[Q_{t+1}|\Omega_t^i], \quad Q_{t+1} = (1 + r)f_t + p_{t+1} - (1 + r)p_t \quad (2.2)$$

where Ω_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. The linear trading rule can be derived by assuming trader i chooses the amount of stock to purchase in accordance with a constant, absolute risk aversion (CARA) preference structure over wealth, which dates back at least to Grossman and Stiglitz (1980). At time t , the budget constraint of investor i is given by

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

where $w_{i,t}$ denotes the wealth of agent i at t and $z_{i,t}$ is the number of units of the risky asset held by agent i at t . The investor will choose $z_{i,t}$ so as to maximize a constant absolute risk aversion utility function

$$-\mathbb{E}_t^i \exp(-\gamma w_{i,t+1}), \quad (2.3)$$

where γ is the risk aversion parameter, and \mathbb{E}_t^i denotes the time t conditional expectation of agent i . All random variables in the model are assumed to be distributed normally, so that (2.3) can be calculated from the (conditional) moment generating function for the normal random variable $-\gamma w_{i,t+1}$. That is,

$$-\mathbb{E}_t^i \exp(-\gamma w_{i,t+1}) = -\exp\{-\gamma \mathbb{E}_t^i(w_{i,t+1}) + (1/2)\gamma^2 v_t(w_{i,t+1})\}$$

where v_t denotes conditional variance. Note that $v_t(w_{i,t+1}) = z_{i,t}^2 v_t(p_{t+1})$. Stationarity implies the conditional variance term will be a constant; thus write $v_t(w_{i,t+1}) \equiv z_{i,t}^2 \delta$. The agent's demand

⁹Although noisy rational expectations models have a long history in finance and macroeconomics, prior applications assume *homogeneous* beliefs.

function for the risky asset follows from the first-order necessary conditions for maximization and is given by

$$z_{i,t} = \frac{1}{\gamma\delta} [\mathbb{E}_t^i p_{t+1} - \beta^{-1}(p_t - f_t)] \quad (2.4)$$

where $\beta^{-1} \equiv 1 + r > 1$. The difference between the trading rule (2.2) and (2.4) is that we have assumed the coefficient of risk aversion and the conditional variance term are normalized to unity. While this limits the role of risk aversion, it preserves linearity which allows us to focus on the linear discounted present value model, and permits closed-form solutions. Setting demand equal to the stochastic supply delivers (2.1).

2.2 INFORMATION STRUCTURE The major departure from the standard noisy rational expectations asset pricing model is that we show that agents have *persistent* heterogeneous beliefs about underlying market fundamentals. 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 an exogenous ongoing filtering process of heterogeneous information. The idea is that each period investors acquire information about some aspect of the risky asset's underlying observed fundamentals. We then derive conditions under which this belief heterogeneity is preserved in equilibrium.

Observed fundamentals are driven by the exogenous process:

$$f_t = a_1(L)\varepsilon_{1t} + a_2(L)\varepsilon_{2t} \quad (2.5)$$

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, unit variance Gaussian random variables, and are assumed to be uncorrelated both contemporaneously and across time.

We assume two trader types—Type 1 and Type 2. Each period both traders observe p_t and f_t . However, in addition, Type 1 traders observe the realizations of ε_{1t} , while Type 2 traders observe the realizations of ε_{2t} .¹⁰ Both the model and the information structure are common knowledge. The time- t information set of Type i is then

$$\Omega_t^i = \mathbb{V}_t(\varepsilon_i) \vee \mathbb{V}_t(f) \vee \mathbb{V}_t(p) \quad (2.6)$$

where the operator $\mathbb{V}_t(x)$ denotes the Hilbert space generated by the random sequence $\{x_{t-j}\}_{j=0}^\infty$ and \vee denotes the span (i.e., the smallest closed subspace which contains the subspaces) of the $\mathbb{V}_t(\varepsilon_i)$, $\mathbb{V}_t(f)$, and $\mathbb{V}_t(p)$ spaces. If the exogenous and endogenous information are disjoint, then the

¹⁰Rondina and Walker (2012) show that the aggregate equilibrium dynamics generated by this information structure is analogous to one in which there are a continuum of traders who each receive a private, noisy signal on an underlying fundamental.

linear span becomes a direct sum. We use similar notation as Futia (1981) in that $\mathbb{V}_t(x) = \mathbb{V}_t(y)$ means the space spanned by $\{x_{t-j}\}_{j=0}^{\infty}$ is equivalent, in mean square, to the space spanned by $\{y_{t-j}\}_{j=0}^{\infty}$.

A potential issue in models with heterogeneous beliefs is the long-run survival of both trader types. Survival is an important issue when the trader types differ in terms of information, over-confidence, risk-aversion, expectation formation, etc. [De Long, Shleifer, Summers, and Waldmann (1991).] However in our setup, agents are symmetrically uninformed. Type 1 agents observe ε_{1t} , while Type 2 agents observe ε_{2t} . Given that the shocks are both Gaussian with the same mean and unit variance, neither trader type has an informational advantage over the other one. Moreover, in the empirical analysis below we assume that the trader types are in equal proportion and $\tilde{a}_1(L) = \tilde{a}_2(L)$.

Specification of the information structure also requires assumptions about the noise process, u_t . Although neither trader observes u_t directly, both its existence and its law of motion are common knowledge.¹¹ We assume an i.i.d. structure for the noise with $u_t = a_3 v_t$ where v_t is a zero mean, Gaussian random variable that is uncorrelated with both ε_{1t} and ε_{2t} .

2.3 THE SIGNAL EXTRACTION PROBLEM Given the equilibrium pricing equation (2.1), agents form expectations about next period's price conditional on (2.6), $\mathbb{E}^i[p_{t+1}|\Omega_t^i]$. 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 do so, since nothing you do can influence the expectations of others. However, 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 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 the trader observes $(p_t, f_t, \varepsilon_{i,t})$ and the underlying structural shocks. 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 solution to Type 2's via symmetry.

For Type 1 traders, the mapping between observables and the underlying shocks takes the

¹¹Engel 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.

following form,

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

$$\mathbf{x}_{1t} = M_1(L)\boldsymbol{\epsilon}_{1t} \tag{2.7}$$

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), which yields 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.¹²

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. To do this, define the average expectation operator as $\bar{\mathbb{E}}_t^0 f_{t+1} = \int_0^1 \mathbb{E}[f_{t+1}|\Omega_t^i] di$ 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

$$p_t = f_t - u_t + \beta \sum_{k=0}^{\infty} \beta^k \bar{\mathbb{E}}_t^k (f_{t+k+1} - u_{t+k+1}) \tag{2.8}$$

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

¹²One of the key differences between the results derived here and those of Kasa (2000) is that we provide conditions that guarantee agents remain asymmetrically informed in equilibrium, implying that they will ‘forecast the forecasts of others’. As shown by Kasa (2000) and Pearlman and Sargent (2005), that is not true in the setup of Townsend (1983) that is studied by Kasa (2000). Thus we are able to isolate the affects of heterogeneous beliefs on asset prices.

problem. Existing approaches either approximate the solution of this infinite dimensional problem [e.g., Nimark (2011), 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].

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.¹³ 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.¹⁴

In order to preserve heterogeneous beliefs, we must ensure that the variables observed by each trader $(p_t, f_t, \varepsilon_t^i)$ do not fully reveal the underlying shocks driving the system in (2.7). The following lemma will be useful in establishing conditions that preserve heterogeneous beliefs in equilibrium.

Lemma 1: *If $\mathbf{M}_1(L)$ has a one-sided inverse in non-negative powers of L , then the Hilbert space generated by \mathbf{x}_{1t} is equivalent (in mean-square error) to the Hilbert space generated by $\boldsymbol{\varepsilon}_{1t}$. If $\mathbf{M}_1(L)$ is not invertible in non-negative powers of L , then the Hilbert space generated by \mathbf{x}_{1t} is a strictly smaller space than that generated by $\boldsymbol{\varepsilon}_{1t}$.*

Proof. See Corollary 2 on page 101 of Hoffman (1962). □

The lemma states that if $\mathbf{M}_1(L)$ is invertible in non-negative powers of L (which correspond to current and past values of p, f, ε), then asymmetric information cannot be preserved in equilibrium because the observed history of \mathbf{x}_t would reveal $\boldsymbol{\varepsilon}_{2t}$ to Type 1 traders. Therefore, they would be able to infer the private information of Type 2 traders. The invertibility condition can be easily checked by calculating the zeros of the determinant of the $\mathbf{M}_1(L)$ matrix. 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 . If $\det \mathbf{M}_1(L)$ contains at least one zero inside the unit circle, then \mathbf{x}_t spans a smaller space than $\boldsymbol{\varepsilon}_{1t}$ and Trader 1 would not be able to recover the private information of Trader 2.

Letting $\mathbf{M}_1(z)$ denote the z -transform of $\mathbf{M}_1(L)$, one can readily verify from (2.7) that

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

¹³The Appendix provides a brief discussion of this theorem and its implications. See Whiteman (1983) for a more detailed discussion.

¹⁴This point has been emphasized in particular in the work of Bart Taub. See, e.g., Taub (1990).

Analogously, for Type 2 traders we have $\det \mathbf{M}_2(z) = a_1(z)\pi_3(z)$. Hence, a sufficient condition for neither trader to be able to infer the other trader's private information is if the exogenous process $a_i(L)\varepsilon_{it}$ has a zero inside the unit circle.¹⁵ For simplicity, we assume that the zeros coincide.

Assumption 1: *The analytic functions $a_1(z)$ and $a_2(z)$ have a single, identical zero inside the unit circle.*

Assumption 1 allows us to write $a_i(z)$ as $(\lambda - z)\tilde{a}_i(z)$ where $|\lambda| < 1$ represents the common noninvertible root of the fundamentals process and $\tilde{a}_i(z)$ are analytic functions with all roots outside the unit circle.

Forming the conditional expectations for (2.1) requires the invertible representation of (2.7) (i.e., the Wold representation). The Wold representation delivers the unique mapping between the observables and the underlying shocks.¹⁶ Standard root flipping procedures [see Rozanov (1967) and Hansen and Sargent (1991)] can be applied to (2.7) to yield the Wold representation. Once this mapping is established, optimal prediction formulas of Wiener and Kolmogorov can be applied to obtain the optimal forecast of next period's price, $\mathbb{E}_{it}p_{t+1}$.

The Wold representation is derived by flipping the roots outside the unit circle via the Blaschke factor, $\mathcal{B}_\lambda(L) = (1 - \lambda L)/(\lambda - L)$ [see Rozanov (1967) and Hansen and Sargent (1991)]. For Type 1 traders this procedure amounts to

$$\begin{aligned} \mathbf{x}_{1t} &= \underbrace{M_1(L)WB(L)}_{M_1^*(L)} \underbrace{B(L)^{-1}W^{-1}\boldsymbol{\epsilon}_{1t}}_{\boldsymbol{\epsilon}_{1t}^*} \\ &= M_1^*(L) \boldsymbol{\epsilon}_{1t}^* \end{aligned} \tag{2.9}$$

where

$$W = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -y/\sqrt{1+y^2} & 1/\sqrt{1+y^2} \\ 0 & 1/\sqrt{1+y^2} & y/\sqrt{1+y^2} \end{bmatrix} \quad B(L) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \mathcal{B}_\lambda(L) & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

where $y = \pi_3(\lambda)/\pi_2(\lambda)$. Note that $\det M_1^*(L) = (1 - \lambda L)\tilde{a}_2(L)\pi_3(L)$. 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 $\boldsymbol{\epsilon}_{1t}^*$. Although it may appear as if $\boldsymbol{\epsilon}_{1t}^*$ is autocorrelated, it is in fact an i.i.d. innovation sequence due to the fact that Blaschke factors have unit modulus (i.e., the variance-covariance generating function for the Blaschke factor is $\mathcal{B}_\lambda(z)\mathcal{B}_\lambda(z)^{-1} = 1$).

With the Wold representation in hand, Type 1's optimal forecast of $\mathbf{x}_{1,t+1}$ is a straightforward

¹⁵Subsequent analysis assumes that in equilibrium the $\pi_3(z)$ function does not contain zeros inside the unit circle. We verify this numerically and provide a general proof when a_3 is small in Appendix A.

¹⁶The uniqueness is in terms of spanning conditions. While moving average representations are not unique, Wold representations deliver the unique space spanned by the observables.

application of the Wiener-Kolmogorov prediction formula

$$\mathbb{E}_t^1 \mathbf{x}_{1,t+1} = \left[\frac{\mathbf{M}_1^*(L)}{L} \right]_+ \boldsymbol{\varepsilon}_{1t}^* = L^{-1}[\mathbf{M}_1^*(L) - \mathbf{M}_1(0)^*] \boldsymbol{\varepsilon}_{1t}^* \quad (2.10)$$

A completely symmetrical expression characterizes Type 2's forecast, with the crucial difference that the unobservable shock is now ε_{1t} .

2.4 EQUILIBRIUM Assuming type 1 agents are in proportion $\kappa \in (0, 1)$, the model is given by

$$p_t = \kappa \beta \mathbb{E}_t^1 p_{t+1} + (1 - \kappa) \beta \mathbb{E}_t^2 p_{t+1} + f_t - \beta u_t \quad (2.11)$$

The equilibrium is found by substituting the expectations (2.10) into (2.11) and deriving the equilibrium pricing functions $\pi_i(z)$ for $i = 1, 2, 3$ in the space of z -transforms, following Futia (1981) and Whiteman (1983). The following theorem characterizes the equilibrium.

Theorem 1: *There exists a rational expectations equilibrium to (2.11) where expectations are given by (2.10) and $a(z)$ satisfies Assumption 1. The equilibrium pricing functions are given by*

$$\pi_1(L) \varepsilon_{1t} = \left[\frac{L(\lambda - L) \tilde{a}_1(L)}{L - \beta} - \frac{\beta(\lambda - \beta) \tilde{a}_1(\beta)}{L - \beta} + \frac{\chi_1 \beta (1 - \kappa) (1 - \lambda^2)}{\lambda(1 - \lambda\beta)(1 - \lambda L)} \right] \varepsilon_{1t} \quad (2.12)$$

$$\pi_2(L) \varepsilon_{2t} = \left[\frac{L(\lambda - L) \tilde{a}_2(L)}{L - \beta} - \frac{\beta(\lambda - \beta) \tilde{a}_2(\beta)}{L - \beta} + \frac{\chi_2 \beta \kappa (1 - \lambda^2)}{\lambda(1 - \lambda\beta)(1 - \lambda L)} \right] \varepsilon_{2t} \quad (2.13)$$

$$\pi_3(L) v_t = \left[a_3 - \frac{\beta(1 - \lambda^2)(\kappa \chi_2 + (1 - \kappa) \chi_1)}{\lambda(1 - \lambda\beta)(1 - \lambda L)} \right] v_t \quad (2.14)$$

where $\chi_i = (y_i(y_i \pi_i(0) - \pi_3(0)))/(1 + y_i^2)$ and $y_1 = \pi_3(\lambda)/\pi_2(\lambda)$, $y_2 = \pi_3(\lambda)/\pi_1(\lambda)$.

Proof. See Appendix A. □

Theorem 1 establishes the existence of a rational expectations equilibrium. If we further impose that the traders are in equal proportion ($\kappa = 0.5$), equivalence between $\tilde{a}_1(L)$ and $\tilde{a}_2(L)$, and if $\pi_3(z)$ contains no zeros inside the unit circle, we can establish uniqueness of the equilibrium.

Corollary 1: For $\kappa = 0.5$, $\tilde{a}_1(L) = \tilde{a}_2(L)$ and if $\pi_3(z)$ contains no zeros inside the unit circle, there exists a unique rational expectations equilibrium to (2.11) where expectations are given by (2.10) and $a(z)$ satisfies Assumption 1.

Proof. See Appendix A. □

In the empirical analysis that follows, we make these assumptions and check that $\pi_3(z)$ contains no zeros inside the unit circle in order to ensure that the equilibrium price is unique (see Appendix A). There are other benefits to setting $\kappa = 0.5$ and $\tilde{a}_1(L) = \tilde{a}_2(L)$. These assumptions simplify the algebra associated with the econometrician's Wold representation. These assumptions also

eliminate issues with survivorship of both agent types. Given that the agents are symmetrically uninformed, an equal proportion assumption implies both agent types will survive in equilibrium.

To interpret the heterogeneous beliefs equilibrium, it is useful to consider the benchmark cases associated with homogeneous beliefs. There are two such cases to consider: [i.] each trader observes both ε_{1t} and ε_{2t} ; and [ii.] neither trader observes ε_{1t} and ε_{2t} . Both cases can be characterized using the standard Hansen-Sargent optimal prediction formula.

Corollary 2: If both trader types observe ε_{1t} and ε_{2t} directly ($\Omega_t = \mathbb{V}_t(\varepsilon_1) \vee \mathbb{V}_t(\varepsilon_2) \vee \mathbb{V}_t(f) \vee \mathbb{V}_t(p)$), the equilibrium pricing functions are given by

$$\pi_i^\varepsilon(L)\varepsilon_{it} = \left[\frac{L(\lambda - L)\tilde{a}_i(L) - \beta(\lambda - \beta)\tilde{a}_i(\beta)}{L - \beta} \right] \varepsilon_{it} \quad (2.15)$$

If neither trader type observes ε_{1t} and ε_{2t} ($\Omega_t = \mathbb{V}_t(f) \vee \mathbb{V}_t(p)$), the equilibrium pricing functions are given by

$$\pi_i^e(L)e_{it} = \left[\frac{L(1 - \lambda L)\tilde{a}_i(L) - \beta(1 - \lambda\beta)\tilde{a}_i(\beta)}{L - \beta} \right] e_{it} \quad (2.16)$$

$$= \left[\frac{L(\lambda - L)\tilde{a}_i(L)}{L - \beta} - \frac{\beta(\lambda - \beta)\tilde{a}_i(\beta)}{L - \beta} - \frac{\tilde{a}_i(\beta)(1 - \lambda^2)}{1 - \lambda L} \right] \varepsilon_{it} \quad (2.17)$$

where $i = 1, 2$ and $e_{it} = (\lambda - L)/(1 - \lambda L)\varepsilon_{it}$.

Proof. See Appendix A. □

Equations (2.15) and (2.16) are the limiting value of (2.12) as $\kappa \rightarrow 1$ and $\kappa \rightarrow 0$, respectively. Thus, the first two terms on the RHS of (2.12) and (2.13) represent the complete information rational expectations equilibrium. From (2.17), the third term on the RHS of (2.12) and (2.13) encodes the uncertainty faced by the trader type that does not observe the underlying shock. For example, Trader Type 2, in proportion $1 - \kappa$, does not observe the ε_{1t} shock directly. As shown in the previous section, the best the agent can do is estimate the moving average associated with (2.9) defined by ϵ_{1t}^* . Note that the construction of the ϵ_{1t}^* relates the true shocks ε_{it} to the innovations in the traders' information set via a Blaschke factor $\mathcal{B}_\lambda(L) = (\lambda - L)/(1 - \lambda L)$. Comparing representation (2.16) with (2.17) shows the correspondence between the Blaschke factorized shocks e_{it} and the true innovations ε_{it} . The third term on the RHS of (2.17) therefore captures the optimal conditioning down when ε_{it} is unobserved. The uncertainty is persistent with autoregressive coefficient of λ ; and as λ approaches one from below, [Assumption 1](#) no longer holds (i.e., the economy becomes fully revealing), and this term vanishes.

3 EMPIRICAL IMPLICATIONS

As noted in the Introduction, there are 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 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 perspective. We focus on three empirical “anomalies” 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 1871-2012.

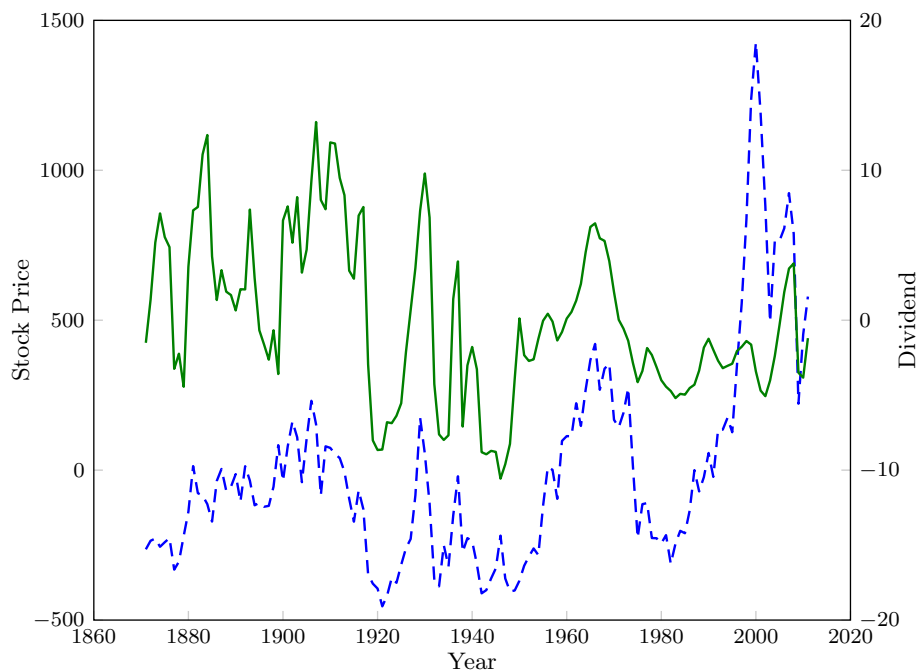


Figure 1: Annual S&P 500 Stock Price (dashed, left axis) and Dividends (solid, right axis). Both series are demeaned, detrended using an exponential trend and deflated using the CPI. Data Source: <http://www.econ.yale.edu/shiller/data.htm>.

Following the original work of Shiller (1981), a common exponential trend was removed from both series. Shiller’s work unleashed a deluge of responses, many of which pointed to biases in his statistical methodology. Removing a common deterministic trend is one of them. However, we are not concerned with this subsequent literature, for a couple of reasons. First, as documented in

¹⁷Recent work by Nimark (2012) 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.

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 more difficult.

3.1 WOLD REPRESENTATION As noted above, in this section we are taking the perspective of an outside econometrician. By ‘outside’ we mean that the econometrician is not an active participant in the market. In particular, he does not observe any of the underlying shocks driving observable fundamentals. Instead, he witnesses the realizations of prices and the fundamentals themselves. Using this information, he wants to test 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, his starting point is to fit VARs to the data. His mistake will arise when using and interpreting these VARs. Specifically, he will misinterpret the residuals as representing innovations to the information sets of *traders*, when in fact they are not.

The first step is to derive the econometrician’s Wold representation, which will be used throughout this section. This representation is the best the econometrician can do given his (false) assumptions about homogeneity of beliefs and his information set. Note that deriving the econometrician’s Wold representation requires a difficult 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 algebraic problem. However, we can greatly simplify it by setting $\tilde{a}_1(L) = \tilde{a}_2(L)$ and $\kappa = 0.5$. This implies $\pi_1(L) = \pi_2(L) = \pi(L)$, which allows us to form the composite shock $\bar{\varepsilon}_t = \varepsilon_{1t} + \varepsilon_{2t}$. Following [Corollary 1](#), the equilibrium exists and is unique.

Proposition 1: *For the equilibrium given in [Theorem 1](#), the econometrician’s Wold representation is given by*

$$\begin{bmatrix} f_t \\ p_t \end{bmatrix} = \begin{bmatrix} (\lambda - L)\tilde{a}(L) & 0 \\ \pi(L) & \pi_3(L) \end{bmatrix} \bar{\eta} \begin{bmatrix} -\eta & 1 \\ 1 & \eta \end{bmatrix} \begin{bmatrix} \frac{1-\lambda L}{\lambda-L} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \xi_{1,t} \\ \xi_{2,t} \end{bmatrix} \quad (3.1)$$

where

$$\begin{bmatrix} \xi_{1,t} \\ \xi_{2,t} \end{bmatrix} = \bar{\eta} \begin{bmatrix} \frac{\lambda-L}{1-\lambda L}(v_t - \eta\bar{\varepsilon}_t) \\ \bar{\varepsilon}_t + \eta v_t \end{bmatrix} \quad (3.2)$$

and $\pi(\cdot)$ and $\pi_3(\cdot)$ are given by [Theorem 1](#), $\eta = \pi_3(\lambda)/\pi(\lambda)$ and $\bar{\eta} = (\sqrt{1 + \eta^2})^{-1}$.

Proof. The proof follows the derivation of the Wold representation for each trader type discussed in [Section 2](#). Given that the fundamentals (f_t) share a common zero at $L = \lambda$, the Blaschke factor $(L - \lambda)/(1 - \lambda L)$ is used to flip the zero outside of the unit circle to $L = \lambda^{-1}$. The flipping of this zero results in the econometrician’s Wold representation [\(3.1\)](#). \square

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.

TABLE 1: ARMA(1,1) ESTIMATES WITH ANNUAL DIVIDEND DATA[†]

	AR(1)	MA(1)
d_t (1871–2011)	0.732 (0.068)	0.308 (0.098)
d_t (1946–2011)	0.872 (0.071)	0.526 (0.117)

[†] Notes: (1) Estimates pertain to the model, $d_t = (1 - \lambda L)/(1 - \rho L)e_t$. (2) Data are annual dividend data of the S&P 500. The data are demeaned, detrended using an exponential trend and deflated using the CPI. Data Source: <http://www.econ.yale.edu/shiller/data.htm>. (3) Asymptotic standard errors are in parentheses. (4) Based on the AIC criterion, the ARMA(1,1) specification was preferred to alternative lag structures (e.g., AR(1), ARMA(2,1)).

It is natural to ask 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. Table 1 displays estimates of univariate ARMA(1,1) model for dividends, which is what the model would predict if $\tilde{a}(L) = \tilde{a}(0)/(1 - \rho L)$. The data are demeaned and detrended annual dividend data of the S&P 500. The full data sample, from 1871–2011, suggest that $\lambda = 0.3$ and $\rho = 0.73$ would be reasonable values. The post-war sample yields a slightly higher value for $\rho = 0.872$ and $\lambda = 0.526$.

3.2 VARIANCE BOUNDS 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, they should be smoother than the realizations of discounted future fundamentals. Violations of variance bounds are a robust empirical finding, so much so that Shiller (1989) claims variance bound violations are the most damning piece of evidence against the efficient-markets hypothesis.

To develop intuition as to how our model will violate variance bounds, it is useful to work through the variance bound calculation for a homogeneous beliefs equilibrium. As discussed above, under the assumption of homogeneous beliefs, the equilibrium given by (2.11) will satisfy the standard Hansen-Sargent optimal prediction formula. Without loss of generality, we may write the equilibrium price as

$$p_t = \left(\frac{L\Gamma(L) - \beta\Gamma(\beta)}{L - \beta} \right) \nu_t \quad (3.3)$$

where fundamentals are given by $f_t = \Gamma(L)\nu_t$ with $\Gamma(L)$ satisfying square summability, and ν_t represents the innovation to the agents' information set. Note that with $\Gamma(L)$ and ν_t appropriately defined, (3.3) is a generalization of both the complete information, homogeneous beliefs equilibrium

($\kappa \rightarrow 1$) given by (2.15) and the incomplete information, homogeneous beliefs equilibrium ($\kappa \rightarrow 0$) given by (2.16) of Corollary 2.

We may decompose the equilibrium price into two components, the perfect-foresight price p_t^{pf} and a remainder term p_t^R :

$$p_t^{pf} = [L\Gamma(L)/(L - \beta)]\nu_t = \sum_{j=0}^{\infty} \beta^j f_{t+j} \quad (3.4)$$

$$p_t^R = [\beta\Gamma(\beta)/(L - \beta)]\nu_t \quad (3.5)$$

The equilibrium price will not be “excessively” volatile relative to the fundamentals if the variance of the perfect-foresight price exceeds that of the price. That is, if $var(p_t) = var(p_t^{pf}) + var(p_t^R) - 2cov(p_t^{pf}, p_t^R) < var(p_t^{pf})$, which holds when $var(p_t^R) - 2cov(p_t^{pf}, p_t^R)$ is negative. Using the residue calculus, any square-summable process $\Gamma(L)$ has the following correlation¹⁹

$$\begin{aligned} cov(p_t^{pf}, p_t^R) &= \frac{\beta\Gamma(\beta)}{2\pi i} \oint \frac{z\Gamma(z)}{(z - \beta)(z^{-1} - \beta)} \frac{dz}{z} = \frac{\beta\Gamma(\beta)}{2\pi i} \oint \frac{z\Gamma(z)}{(z - \beta)(1 - \beta z)} dz \\ &= \frac{\beta^2\Gamma(\beta)^2}{1 - \beta^2} \end{aligned} \quad (3.6)$$

This positive correlation is a direct result of the assumption of a homogeneous beliefs, rational expectations equilibrium. The Hansen-Sargent formula (3.3) represents the agents’ optimal prediction of the price given knowledge of the model structure. The perfect-foresight price (3.4) would be the equilibrium if the agent knew past, present and *future* realizations of ν_t . Future realizations are not part of the information set and therefore the remainder term (3.5) represents a particular linear combination of future ν_t ’s that must be subtracted off from the perfect foresight price.²⁰ This conditioning down is optimal but tolerates correlation in the forecast error

$$p_t - p_t^{pf} = -\frac{\beta\Gamma(\beta)}{L - \beta}\nu_t = -\beta\Gamma(\beta)\{\nu_{t+1} + \beta\nu_{t+2} + \beta^2\nu_{t+3}\} \quad (3.7)$$

The serial correlation in the forecast error cannot be exploited to improve the prediction of p_t^{pf} because it involves *future* values of ν_t [cf. Lewis and Whiteman (2008)]. Therefore it is optimal for the remainder term, p_t^R , to be positively correlated with the perfect foresight price.

Moreover, the variance of p_t^R is given by $\beta^2\Gamma(\beta)^2/(1 - \beta^2)$. This, coupled with (3.6), yields

$$var(p_t^R) - 2cov(p_t^{pf}, p_t^R) = -\frac{\beta^2\Gamma(\beta)^2}{1 - \beta^2} \quad (3.8)$$

which is clearly negative and delivers Shiller’s bound. Thus, we have established the following

¹⁹Using the residue calculus allows us to be very general with respect to the exogenous fundamentals process, f_t . Appendix B offers a primer on calculating contour integrals.

²⁰The linear combination that will be subtracted from p_t^{pf} is the principal part of the Laurent series expansion of p_t^{pf} . This is a result of optimal prediction formulas [Hansen and Sargent (1980)].

proposition.

Proposition 2: *If the asset price is generated by a homogeneous-beliefs equilibrium given by Corollary 2 and the econometrician’s information set is identical to that of the agents’, then the equilibrium price will not exhibit excess volatility.*

The derivations leading up to (3.8) assumed that the econometrician’s Wold representation from Proposition 1 could uncover the agents’ innovation ν_t . The following corollary to this proposition weakens the assumption that the econometrician shares the information set of the agents of the model. In fact, the econometrician can condition on a strict subset of the information available to traders and still accurately determine whether or not the asset price is prone to excess volatility. However, this is true if and only if the asset price is generated by homogeneous beliefs. This result is due to the conditioning down arguments of Hansen and Sargent (1991) and Campbell and Shiller (1987).

Corollary 3: *If the asset price is generated by the complete information, homogeneous-beliefs equilibrium given by (2.15) of Corollary 2 and the econometrician’s information set is given by the Wold representation of Proposition 1, then the econometrician’s inferences will be correct and the equilibrium price will not exhibit excess volatility.*

Proof. A complete proof can be found in Appendix A but the intuition for this result is that the conditional second moments of the equilibrium price are identical across the two information sets.

$$\text{var}(p_t^{pf}|e_t) = \text{var}(p_t^{pf}|\varepsilon_t), \quad \text{var}(p_t^R|e_t) = \text{var}(p_t^R|\varepsilon_t), \quad \text{cov}(p_t^R, p_t^{pf}|e_t) = \text{cov}(p_t^R, p_t^{pf}|\varepsilon_t) \quad (3.9)$$

where p_t^R and p_t^{pf} represent the perfect foresight price and the remainder term as identified in (3.4)–(3.5) and conditional on the information set of the agent ε_t and econometrician e_t . Thus the calculations leading to Proposition 2 continue to hold. \square

Proposition 2 and Corollary 3 imply that if the economy is characterized by a homogeneous beliefs equilibrium, then the outside econometrician’s inference will be correct and the price dynamics will not be excessively volatile. The intuition for why homogeneous belief economies will never exhibit excess volatility comes from the optimization inherent in a homogeneous beliefs rational expectations equilibrium. However, the logic behind variance bounds rests on two key assumptions—there are no missing fundamentals and forecast errors are orthogonal to forecasts. Hamilton and Whiteman (1985) emphasize the point that rational agents could forecast variables not included in the econometrician’s information set, which renders excess volatility tests statistically invalid. The “missing fundamentals” in our model are the higher-order beliefs formed by each agent type. Second, variance bounds are premised on the assumption that forecast errors are orthogonal to forecasts. This will always be true under the assumption of rational expectations and homogeneous beliefs. In models with heterogeneous beliefs however, this simply is not the case. Although each trader’s own individual forecast errors are orthogonal to his own information set, this argument

does not extend to the market forecast error. 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'.

As a heuristic example that is directly applicable to our setup, let $p_t = (\lambda - L)\pi(L)\varepsilon_t$ with $|\lambda| < 1$ and assume Type 1 agents observe ε_t directly, while Type 2 agents only observe current and past realizations of p .²¹ Using the Wiener-Kolmogorov optimal prediction formulas gives

$$\begin{aligned}\mathbb{E}[p_{t+1}|\varepsilon^t] &= L^{-1}[(\lambda - L)\pi(L) - \lambda\pi(0)]\varepsilon_t \\ \mathbb{E}[p_{t+1}|p^t] &= L^{-1}[(1 - \lambda L)\pi(L) - \pi(0)]e_t = L^{-1}[(\lambda - L)\pi(L) - \mathcal{B}_\lambda(L)\pi(0)]\varepsilon_t\end{aligned}$$

where $\mathcal{B}_\lambda(L) = (\lambda - L)/(1 - \lambda L)$ and $e_t = \mathcal{B}_\lambda(L)\varepsilon_t$. The equilibrium price is determined by the average market expectation, $\bar{\mathbb{E}}_t p_{t+1} = \kappa\mathbb{E}[p_{t+1}|\varepsilon^t] + (1 - \kappa)\mathbb{E}[p_{t+1}|p^t]$. Using the optimal prediction formulas above, it is straightforward to show that $\bar{\mathbb{E}}_t \bar{\mathbb{E}}_{t+1} p_{t+2} \neq \bar{\mathbb{E}}_t p_{t+2}$,

$$\bar{\mathbb{E}}_t \bar{\mathbb{E}}_{t+1} p_{t+2} - \bar{\mathbb{E}}_t p_{t+2} = -\kappa(1 - \kappa)\pi(0)\lambda \left(\frac{1 - \lambda^2}{1 - \lambda L} \right) \varepsilon_t \quad (3.10)$$

The intuition behind the law of iterated expectations is that it would be suboptimal for the agents to alter their forecast of the price over time. However, when agents form higher-order beliefs, it becomes optimal for them to adjust their forecast of the price based upon the actions of other agents. Hence, the average expectation of the price is changing over time. This result is now well known [cf., Allen, Morris, and Shin (2006) Bacchetta and van Wincoop (2006) Rondina and Walker (2012)] but, to the best of our knowledge, our paper is the first to examine the implications of this result in an econometric model of asset pricing.

Theorem 1 and **Corollary 2** implies that we may write the equilibrium price according to

$$p_t = p_t^{pf} - \underbrace{\kappa p_t^{R|\varepsilon} - (1 - \kappa)p_t^{R|e} - p_t^{HOBs}}_{p_t^{R|Het}} \quad (3.11)$$

$$= p_t^{pf} - p_t^{R|Het} \quad (3.12)$$

where $p_t^{R|\varepsilon}$ and $p_t^{R|e}$ denote the error term in forecasting future fundamentals for each agent type, and p_t^{HOBs} denotes the higher-order beliefs component of the equilibrium price. Equation (3.12) is how the outside econometrician, who erroneously assumes a homogeneous beliefs equilibrium, will perceive this forecast error. He will lump the higher-order beliefs component and the convex combination of forecast errors into one term, $p_t^{R|Het}$. Given that we have an analytical solution on hand, we know the precise functional form of this error term. From **Theorem 1**

$$p_t^{R|Het} = - \left[\frac{\beta(\lambda - \beta)\tilde{a}_1(\beta)}{L - \beta} - \frac{\chi_1\beta(1 - \kappa)(1 - \lambda^2)}{\lambda(1 - \lambda\beta)(1 - \lambda L)} \right] \varepsilon_{1t} \quad (3.13)$$

²¹This is a slightly modified version of the model of Section 2 because here we assume f_t is not observable, but the intuition is directly applicable.

with a similar representation for ε_{2t} . The heterogeneous beliefs forecast error is a combination of the complete information, homogeneous beliefs error term (first term), and a term that captures the higher-order belief elements and the incomplete information error term, weighted by $(1 - \kappa)$, (second term).

It is obvious that the variance of the error term will be higher in the heterogeneous beliefs equilibrium vis-à-vis the homogeneous beliefs equilibrium due to the presence of the higher-order beliefs term (second term on the RHS of (3.13)). What is not clear is if this additional volatility is enough to overcome the variance bound. The following proposition gives the conditions under which excess volatility is achieved.

Proposition 3: *If the asset price is generated by a heterogeneous-beliefs equilibrium given by Theorem 1, then the equilibrium price will exhibit excess volatility if*

$$\text{var}(p_t^{\text{R|Het}}) - 2\text{cov}(p_t^{\text{R|Het}}, p_t^{\text{pf}}) > 0 \tag{3.14}$$

which holds when

$$\frac{\chi_1 \beta^2 (1 - \lambda^2)(1 - \kappa)}{\lambda(1 - \lambda\beta)} \left[\frac{\chi_1(1 - \kappa)}{\lambda(1 - \lambda\beta)} - 2\tilde{a}_1(\beta) \right] - \frac{(\lambda - \beta)^2 \beta^2 a_1(\beta)^2}{1 - \beta^2} > 0 \tag{3.15}$$

Proof. See Appendix A. □

There is an analogous representation for ε_{2t} . The last term is the correlation attributable to homogeneous-beliefs equilibrium and is equivalent to (3.8). The additional term is due to heterogeneous beliefs. As λ (or κ) approaches one, the economy reverts to the full-information, homogeneous beliefs equilibrium and no violations would occur. Violations of the variance bound are not guaranteed. The variance of the heterogeneous-beliefs component of the price (second term on the RHS of (3.13)) must be sufficiently large. Of course, the issue here is whether the bound is breached according to the econometrician’s Wold representation given by (3.1). To investigate this we must resort to numerical simulations.

Figure 2 reports plots of $\text{var}(p)/\text{var}(p^{\text{pf}})$ for alternative parameter values of λ according to the Wold representation of the econometrician. Therefore, the same assumptions imposed to generate the econometrician’s Wold representation apply to Figure 2 (i.e., $\tilde{a}_1 = \tilde{a}_2$, and $\kappa = 0.5$). In accordance with Table 3.1, we assume f_t follows an ARMA(1,1) process with autoregressive coefficient of $\rho = 0.732$. Given annual data, we set $\beta = 0.95$. This is close to the original discount rates used by Shiller. Finally, in order to ensure that the additional volatility is not due to noise traders, Figure 2 subtracts off the additional variance attributable to the noise term a_3 in the equilibrium price p_t . The implied value for a_3 ranges from 1.58 to 2.73 in Figure 2.²²

²²The appendix establishes the uniqueness of the equilibrium given these parameter values and plots the implied value of a_3 and root of $\pi_3(z)$ in Figure 2.

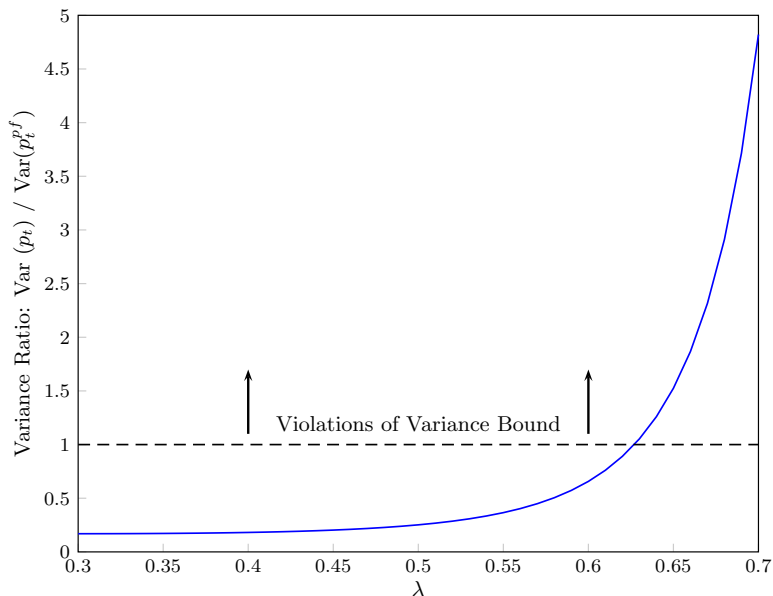


Figure 2: Violation of Variance Bound. The solid line plots the variance of the equilibrium price over the variance of the perfect foresight price less the volatility due to noise traders for $\beta = 0.95$, and $\rho = 0.732$. Violations of the variance bound occur for values greater than unity.

Notice that violations do not occur when $\lambda \approx 0.3$. Hence, the full-sample estimate of the ARMA process given by Table 3.1 will satisfy the variance bound. It is not 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 Theorem 1 and equations (3.10) and (3.13). Notice that the *persistence* of heterogeneous beliefs is dictated by λ . When λ is close to zero heterogeneous beliefs just add a small amount high frequency noise to prices.

Conversely, for values of $\lambda > 0.62$, the asset price violates the variance bound. This value is within one standard error of the post-war estimate for λ given by Table 3.1. As λ increases, we see significant violations of the bound of the same order of magnitude that Shiller found. For example, when $\lambda = 0.69$, the variance of observed prices is more than four times its hypothetical upper bound! Note this is not an artifact of biased statistical procedures, since we are comparing *population* moments.

Figure 3 compares the sample path of observed stock prices with their perfect foresight counterparts. The left panel updates Shiller's (1981) original plot. 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 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 right panel of Figure 3 follows the same procedure using data generated by our model. All the parameters are the same as Figure 2 with $\lambda = 0.65$. The plot gives the distinct impression

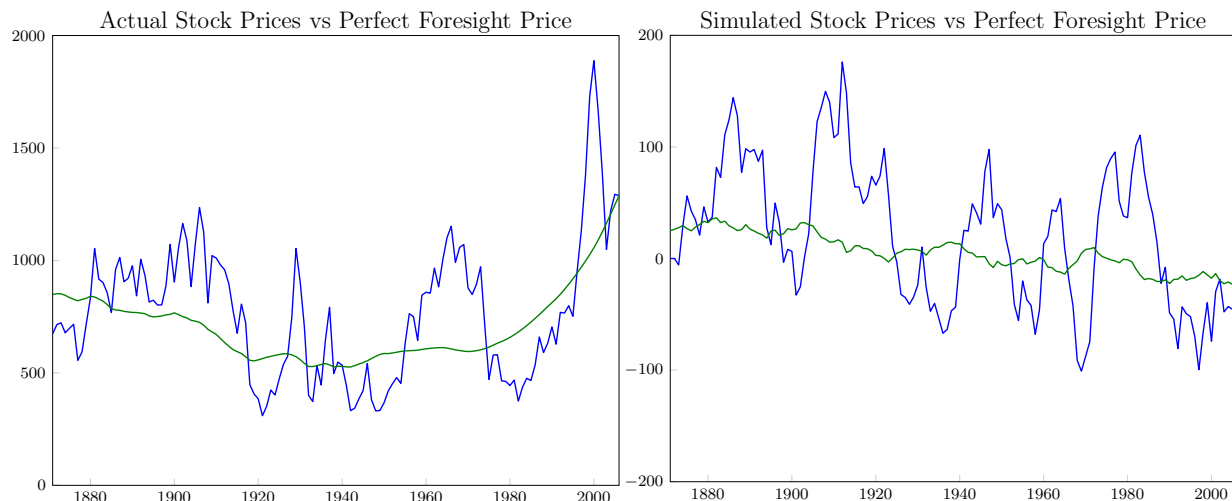


Figure 3: Stock Prices vs Present Value of Future Dividends (Left: Data, Right: Simulation)

that prices are too volatile relative to their fundamentals.

3.3 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.²³ 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 [Cochrane (2001) and Shiller (1989)]. 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 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 is satisfied. The equilibrium pricing functions were computed by imposing this condition. However, this is not the condition the 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 Wold representation in (3.1) to construct what he (falsely) believes is the market’s expectation of next period’s price. The Wold representation is a subset of the information available to the traders and therefore the econometrician’s projection of excess returns on his time t information set will not display orthogonality.

Define the excess return as $R_{t+1} = \beta \mathbb{E}_t p_{t+1} + f_t - p_t$, where the expectation is taken with respect to the Econometrician’s information set given by the Wold representation (3.1). Predictability suggests that if we regress R_t onto lagged information, we should find statistically significant coefficients. The following result summarizes what our econometrician would find

Proposition 4: *Given Assumption 1, and assuming a symmetric equilibrium (i.e., $\kappa = 0.5$,*

²³See Kojien and Van Nieuwerburgh (2010) for a recent review of the literature.

$\tilde{a}_1(L) = \tilde{a}_2(L) = \tilde{a}(L)$, the econometrician's Wold Representation in (3.1) generates the following projection of R_{t+1} onto the (econometrician's) time- t information set

$$\beta E_t p_{t+1} + f_t - p_t = -h_1 \xi_{1,t} - \beta(\lambda - \beta) \tilde{a}(\beta) \left(h_2 + \frac{\delta(1 - \beta\lambda)}{1 - \lambda L} \right) \xi_{2,t} \quad (3.16)$$

where $h_1 = \beta x(1 + .5y^2)/[(1 - \lambda\beta)(1 + y^2)] - \eta\beta\tilde{a}(\beta)$, $h_2 = ya_3/[\beta(\lambda - \beta)\tilde{a}(\beta)]$, $\delta = -.5xy(1 - \lambda^2)/[\lambda\tilde{a}(\beta)(1 + y^2)(1 - \lambda\beta)(\lambda - \beta)]$, and x and y are defined in the proof of Theorem 1.

Proof. See Appendix A. □

The upshot is that (3.16) is non-zero and the econometrician will find predictability in asset returns. One can readily verify that as $a_3 \rightarrow 0$, so that the equilibrium becomes revealing and expectations become homogeneous, we have $(h_1, h_2, \delta) \rightarrow 0$, and returns become unpredictable. Hence, the only reason the econometrician is finding predictability is that he is incorrectly assuming a homogeneous beliefs equilibrium. If he were to take into account the heterogeneity, the excess return (3.16) would be zero.

Note that by inverting the Wold representation we can express $(\xi_{1,t}, \xi_{2,t})$ in terms of the history of (p_t, f_t) , so that excess returns can be written as functions of current and past observables. Table 2 contains a small set of results from these kinds of regressions, using the heterogeneous beliefs equilibrium of Theorem 1 and 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}	P_{t-1}	P_{t-2}	\bar{R}^2
<u>Data</u>					
$R_t =$.317 (.209)				.094
$R_t =$.375 (.181)	-.179 (.161)			.116
$R_t =$			-.031 (.048)		.012
$R_t =$.274 (.194)	-.325 (.216)	.110
<u>Model</u>					
$R_t =$.544 (.076)				.274
$R_t =$.476 (.088)	.137 (.089)			.281
$R_t =$			-.050 (.011)		.134
$R_t =$.422 (.075)	-.479 (.076)	.331

Notes: (1) All regressions include a constant.

(2) Heteroskedascity-robust standard errors in parentheses.

The empirical results here are consistent with those of the literature.²⁴ There is evidence in favor of predictability and our model generated data matches this predictability relatively well. Evidence of predictability is somewhat stronger using model-simulated data. However, if the data are restricted to post-1945, predictability becomes more pronounced [Kojien and Van Nieuwerburgh (2010)]. Our results suggest that even if one had access to the *population* second moments, we would still observe predictability. Finding that returns are predictable is not a puzzle if investors have heterogeneous expectations of returns.

Recent papers have turned to present-value relationships like (2.8) and focused on the distinct role of return and dividend *growth* predictability to better understand this anomaly [Cochrane (2008), Lettau and Van Nieuwerburgh (2008), Kojien and Van Binsbergen (2010)]. Evidence

²⁴Kojien and Van Nieuwerburgh (2010), [Table 1] conduct several robustness checks and show that predictability is robust to log-linearized returns, excess returns (log returns in excess of log risk-free rate), adjusting for inflation, excluding NASDAQ stocks.

against dividend growth predictability implies evidence in favor of return predictability, and vice versa. These studies have found that dividend growth predictability is strong, but not very persistent. While return predictability is modest but very persistent. Over 90% of the variation in the price-dividend ratio is due to variation in expected returns [Kojien and Van Nieuwerburgh (2010)]. In a homogeneous-beliefs equilibrium, the law of iterated expectations applies and therefore a present-value relationship of the form, $p_t = \mathbb{E}_t \sum_{j=0}^{\infty} \beta^j f_{t+j}$, holds. However if an econometrician assumes homogeneous beliefs and proceeds to impose the law of iterated expectations on a heterogeneous-beliefs equilibrium, the price will include an additional term. [Theorem 1](#) shows that the heterogeneous-beliefs equilibrium price can be written as $p_t = p_t^* + p_t^R$, where p_t^* is the equilibrium that would emerge under a fully-revealing, homogeneous-beliefs equilibrium. The additional term, p_t^* , is due to heterogeneous beliefs and, as demonstrated by [Proposition 4](#), will lead to predictability and variation in expected returns.

3.4 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 (1991) 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.

The intuition behind our result comes from the discrepancy between the complete information, homogeneous beliefs pricing function and the heterogeneous beliefs counterpart outlined in [Theorem 1](#), which we repeat here for convenience,

$$\pi_1(L)\varepsilon_{1t} = \left[\frac{L(\lambda - L)\tilde{a}_1(L)}{L - \beta} \right] \varepsilon_{1t} - \left[\frac{\beta(\lambda - \beta)\tilde{a}_1(\beta)}{L - \beta} \right] \varepsilon_{1t} - \left[\frac{\chi_1\beta(1 - \kappa)(1 - \lambda^2)}{\lambda(1 - \lambda\beta)(1 - \lambda L)} \right] \varepsilon_{1t} \quad (3.17)$$

with an analogous representation for ε_{2t} . We can represent the equilibrium price as

$$p_t = K(\tilde{a}_1(L))\varepsilon_{1t} + K(\tilde{a}_2(L))\varepsilon_{2t} \quad (3.18)$$

where we now write the pricing function as $K(\tilde{a}(L))$ to emphasize the fact that it is the output of a linear operator, $K(\cdot)$. Note that the heterogeneous beliefs pricing operator consists of the sum of *two* linear operators, $K = K^s + K^h$, where K^s denotes the conventional complete information,

homogeneous beliefs operator (first and second term on the RHS of (3.17)) given by the Hansen-Sargent formula, and K^h denotes the heterogeneous beliefs operator (third term on the RHS of (3.17)). This delivers the following result,

Remark 1: 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 (i.e., the third term on the RHS of (3.17)). 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 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) specification is 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 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.

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	11.9	20.1	.474
$P_t =$	-1.28 (1.16)	.927 (.040)	.814	1.72			
<u>Simulation</u>							
$D_t =$.445 (.086)	.0017 (.0005)	.469	2.26	15.7	415.5	.992
$P_t =$	11.2 (2.56)	.953 (.014)	.985	1.66			

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.

The bottom panel reports the results from following the exact same procedures using data generated by the model, with the parameter values $\kappa = 0.5$, $\lambda = 0.65$, and $\rho = 0.73$. 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 left panel of Figure 4 is nearly identical to the left panel of Figure 3. On the other hand, the right panels of these two figures, pertaining to model-generated data, 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

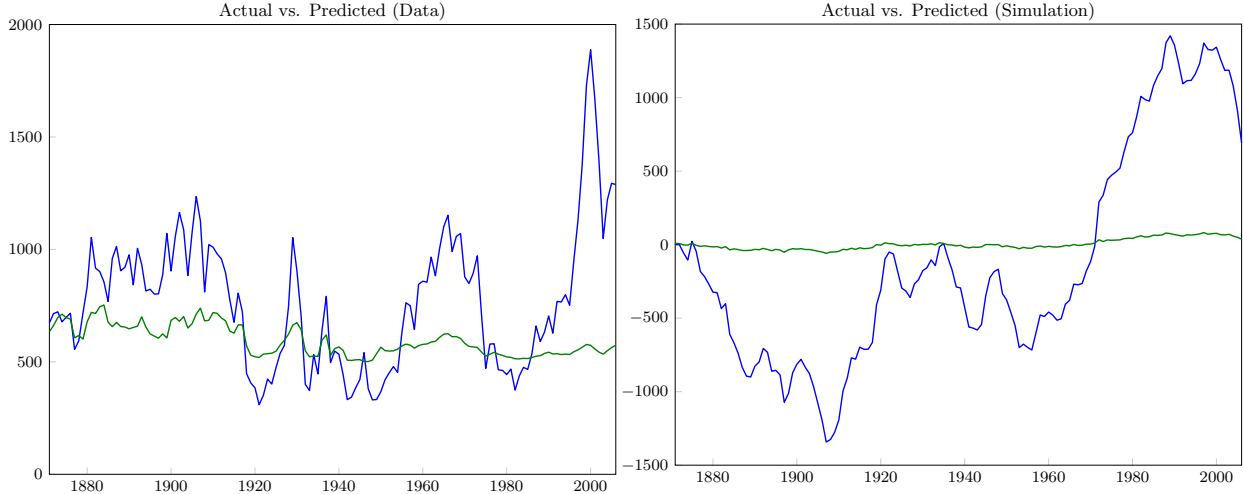


Figure 4: Actual vs. Predicted Stock Prices (Left: Data, Right: Simulation)

adhere to the law of iterated expectations. Unfortunately, the clever VAR procedures of Hansen and Sargent (1991) and Campbell and Shiller (1987) rely heavily on the law of iterated expectations.

4 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. 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. Moreover, Rondina and Walker (2012) show that our information structure is consistent with a setup in which there are a continuum of traders who each receive a private, noisy signal on an underlying fundamental. We interpret this as suggestive evidence that our results are robust to alternative information structures.

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. Three 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, issues of market microstructure might be critical to the propagation or mitigation of heterogeneous beliefs. Of particular interest would be relaxing the assumption of infinitesimal investors or examining issues of market clearing (e.g., relaxing the assumption that the riskless asset is inelastically supplied at a fixed price). Finally, 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.

REFERENCES

- ACEMOGLU, D., V. CHERNOZHUKOV, AND M. YILDIZ (2008): “Fragility of Asymptotic Agreement under Bayesian Learning,” MIT Department of Economics Working Paper.
- ALBAGLI, E., C. HELLWIG, AND A. TSYVINSKI (2011): “A Theory of Asset Prices Based on Heterogeneous Information,” NBER Working Paper No. 17548.
- ALLEN, F., S. MORRIS, AND H. SHIN (2006): “Beauty Contests and Iterated Expectations in Asset Markets,” *Review of Financial Studies*, 19(3), 719–752.
- ANGELETOS, G.-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.
- BANERJEE, S., R. KANIEL, AND I. KREMER (2009): “Price Drift as an Outcome of Differences in Higher-Order Beliefs,” *Review of Financial Studies*, 22(9), 3707–3734.
- 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, NJ.
- COCHRANE, J. H. (2008): “The Dog that Did Not Bark: A Defense of Return Predictability,” *Review of Financial Studies*, 21(4), 1533–1575.
- CONSTANTINIDES, G. M., AND D. DUFFIE (1996): “Asset Pricing with Heterogeneous Consumers,” *Journal of Political Economy*, 104, 219–240.
- DE LONG, J. B., A. SHLEIFER, L. H. SUMMERS, AND R. J. WALDMANN (1991): “The survival of noise traders in financial markets,” *Journal of Business*, 64(1), 1–19.
- DETEMPLE, J., AND S. MURTHY (1994): “Intertemporal asset pricing with heterogeneous beliefs,” *Journal of Economic Theory*, 62(2), 294–320.
- DUMAS, B., A. KURSHEV, AND R. UPPAL (2009): “Equilibrium portfolio strategies in the presence of sentiment risk and excess volatility,” *The Journal of Finance*, 64(2), 579–629.
- 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 (1980): “Formulating and Estimating Dynamic Linear Rational Expectations Models,” *Journal of Economic Dynamics and Control*, 2, 7–46.
- (1991): “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.

- HARRIS, M., AND A. RAVIV (1993): “Differences of opinion make a horse race,” *Review of Financial Studies*, 6(3), 473–506.
- 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.
- HELLWIG, M. F. (1980): “On the aggregation of information in competitive markets,” *Journal of economic theory*, 22(3), 477–498.
- HOFFMAN, K. (1962): *Banach Spaces of Analytic Functions*. Prentice-Hall, Englewood Cliffs, New Jersey.
- KASA, K. (2000): “Forecasting the Forecasts of Others in the Frequency Domain,” *Review of Economic Dynamics*, 3, 726–756.
- KOIJEN, R., AND J. VAN BINSBERGEN (2010): “Predictive Regressions: A Present-Value Approach,” *The Journal of Finance*, 65(4), 1439–1471.
- KOIJEN, R. S., AND S. VAN NIEUWERBURGH (2010): “Predictability of returns and cash flows,” NBER Working Paper No. 16648.
- LAM, P.-S., S. G. CECCHETTI, AND N. C. MARK (2000): “Asset Pricing with Distorted Beliefs: Are Equity Returns Too Good to Be True?,” *American Economic Review*, 90(4), 787–805.
- LETTAU, M., AND S. VAN NIEUWERBURGH (2008): “Reconciling the return predictability evidence,” *Review of Financial Studies*, 21(4), 1607–1652.
- LEWIS, K. F., AND C. H. WHITEMAN (2008): “Prediction Formulas,” in *The New Palgrave Dictionary of Economics*, ed. by S. N. Durlauf, and L. E. Blume. Palgrave Macmillan, Basingstoke.
- LUCAS, J. R. E. (1978): “Asset Prices in an Exchange Economy,” *Econometrica*, 46(6), 1429–1445.
- MAKAROV, I., AND O. RYTCHKOV (2012): “Forecasting the Forecasts of Others: Implications for Asset Pricing,” *Journal of Economic Theory*, 147, 941–966.
- 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. (2011): “Dynamic higher order expectations,” Working paper, Universitat Pompeu Fabra.
- NIMARK, K. P. (2012): “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 (2012): “Information Equilibria in Dynamic Economies with Dispersed Information,” Working Paper, Indiana University.
- ROZANOV, Y. A. (1967): *Stationary Random Processes*. Holden-Day, San Francisco.
- SCHEINKMAN, J. A., AND W. XIONG (2003): “Overconfidence and Speculative Bubbles,” *Journal of Political Economy*, 111, 1183–1219.
- SHILLER, R. (1989): *Market Volatility*. MIT Press, Cambridge, MA.
- 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.
- SINGLETON, K. 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(4), 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. (1994): “A Model of Competitive Stock Trading Volume,” *Journal of Political Economy*, 102, 127–168.
- 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.
- XIONG, W., AND H. YAN (2010): “Heterogeneous expectations and bond markets,” *Review of Financial Studies*, 23(4), 1433–1466.