Heterogeneous Expectations and Bond Markets

Wei Xiong
Princeton University and NBER

Hongjun Yan
Yale University

This paper presents a dynamic equilibrium model of bond markets in which two groups of agents hold heterogeneous expectations about future economic conditions. The heterogeneous expectations cause agents to take on speculative positions against each other and therefore generate endogenous relative wealth fluctuation. The relative wealth fluctuation amplifies asset price volatility and contributes to the time variation in bond premia. Our model shows that a modest amount of heterogeneous expectations can help explain several puzzling phenomena, including the "excessive volatility" of bond yields, the failure of the expectations hypothesis, and the ability of a tent-shaped linear combination of forward rates to predict bond returns. (*JEL* D50, D90, G12)

1. Introduction

Standard asset pricing models use a representative-agent-based framework. Although this framework leads to tractable asset pricing formulas, it ignores important interactions among heterogeneous agents. There is ample evidence supporting the existence of heterogeneous expectations among agents, as shown in various surveys of professional forecasters and economists.[1] Casual observations also suggest that agents take on different investment positions, and trading between them can have important effects on asset prices. For example,

[1] For example, Mankiw, Reis, and Wolfers (2004) find that the interquartile range among professional economists' inflation expectations, as shown in the Livingston Survey and the Survey of Professional Forecasters, varies from above 2% in the early 1980s to around 0.5% in the early 2000s. Swanson (2006) finds that in the Blue Chip Economic Indicators survey of major U.S. corporations and financial institutions between 1991 and 2004, the difference between the 90th and 10th percentile forecasts of the next-quarter real U.S. GDP growth rate fluctuates between 1.5% and 5%, and the 90th and 10th percentile forecasts of the four-quarter-behind three-month Treasury bill rate fluctuates between 0.8% and 2.2%.

Advance Access publication November 15, 2009

in the ongoing financial crisis of 2007–2008, while many financial institutions invested heavily in securities related to subprime mortgages, some hedge funds instead took on large short positions. When the prices of these securities started to fall, these funds were able to make a large profit at the expense of those optimistic financial institutions.[2] The increased capital allowed these hedge funds to take even larger short positions and to push the prices further down.

In this paper, we analyze a dynamic equilibrium model in which heterogeneous expectations cause investors to trade with each other. We show that the endogenous wealth fluctuations caused by investors' trading can help to resolve several challenges encountered by standard representative-agent models, including the "excessive volatility" of bond yields, the failure of the expectations hypothesis, and the ability of a tent-shaped linear combination of forward rates to predict bond returns.

We adopt the standard endowment economy of Lucas (1978). To bridge our model with empirical studies of nominal interest rates, we introduce an exogenous price inflation process. We allow investors to hold heterogeneous expectations of future economic conditions. Specifically, we assume that there are two groups of investors using different learning models to estimate the unobservable inflation target, which determines future inflation and nominal short rates. Consequently, the two groups of investors hold heterogeneous expectations about future interest rates. This disagreement motivates investors to take on speculative positions against each other, and market clearing conditions determine the equilibrium prices. Following Detemple and Murthy (1994), we solve the equilibrium in a closed form and show that the equilibrium bond price is a wealth-weighted average of bond prices in homogeneous economies, in each of which only one type of investor is present.

Our model implies that the relative wealth fluctuation caused by investors' speculative positions amplifies bond yield volatility and contributes to the time variation in bond premia. The intuition is as follows. Suppose a group of investors is optimistic about future short interest rates (i.e., their expectation of future short rates is higher than that of the other group's). Then, investors in the optimistic group would bet on rates rising against those pessimistic investors. In equilibrium, bond prices aggregate investors' heterogeneous beliefs and, in particular, reflect their *wealth-weighted* average belief. When investors' wealth-weighted average belief about future short rates is higher than the econometrician's belief, they will discount bonds more heavily and the equilibrium bond prices would appear "cheap" to the econometrician—i.e., the bond premium is high. Similarly, the bond premium is low when investors' wealth-weighted average belief is lower than the econometrician's belief. Thus, the bond premium varies with the two groups' beliefs and wealth distribution. Note that the two

---

2 Two hedge funds managed by Paulson & Co were reported to make $15 billion from shorting subprime mortgages with a return over 600%; see, e.g., "The Wall Street Investor Who Shorted Subprime and Made $15bn," *Money Week*, January 28, 2008.

groups' wealth distribution is endogenously determined by their trading. When a positive shock hits the market, it favors optimistic investors and causes wealth to flow from pessimistic investors to optimistic investors, giving the optimistic belief a larger weight in determining bond prices. The relative-wealth fluctuation thus amplifies the effect of the initial news on bond prices and makes bond premia more variable.

We provide a calibration exercise to show that even with a modest amount of belief dispersion, the volatility amplification effect of investors' relative wealth fluctuation is significant enough to explain the "excess volatility puzzle" documented by Shiller (1979); Gurkaynak, Sack, and Swanson (2005); and Piazzesi and Schneider (2006). These studies find that long-term yields appear to be too volatile relative to the levels implied by standard representative-agent models.

We also show that heterogeneous expectations can help explain the failure of the classic expectations hypothesis in the data. The expectations hypothesis suggests that when the yield spread (long-term bond yield minus the short rate) is positive, the long-term bond yield is expected to rise (or the long-term bond price is expected to fall) because, otherwise, an investor cannot be indifferent about investing in the long bond or the short rate. However, this hypothesis has been rejected by many empirical studies. To mention one here, Campbell and Shiller (1991) find that when the yield spread is positive, the long-term bond yield tends to fall rather than rise. This pattern is a natural implication of our model: Suppose the wealth-weighted average belief about the future short rates is higher than the econometrician's belief. On the one hand, it implies that investors discount long-term bonds more heavily, which leads to higher long-term bond yields and so larger yield spreads; on the other hand, it also implies that the long-term bond prices appear "cheap" from the econometrician's point of view—i.e., the long-term bond prices are expected to rise and bond yields are expected to fall. Taken together, a high wealth-weighted average belief implies both large yield spreads and falling long-term bond yields in the future. Indeed, our simulations show that a reasonable amount of belief dispersion is able to generate regression results similar to those of Campbell and Shiller (1991).

Our model can also shed light on the recent finding of Cochrane and Piazzesi (2005) that a single tent-shaped linear combination of forward rates predicts excess returns on two- to five-year bonds. On the one hand, as we elaborate later, this tent-shaped factor tracks investors' wealth-weighted belief: the higher the weighted average belief about future short rates, the bigger the value of the tent-shaped factor. On the other hand, a higher wealth-weighted average belief about future short rates also makes bond prices cheap from the econometrician's point of view, and thus predicts higher future bond returns. As a result, the tent-shaped factor predicts bond premia. Our simulations confirm that a reasonable amount of belief dispersion is able to generate bond return predictability results comparable with those of Cochrane and Piazzesi (2005).

Our paper complements the growing literature on equilibrium effects of heterogeneous beliefs. Detemple and Murthy (1994) are the first to demonstrate

that equilibrium prices have a wealth-weighted average structure. Zapatero (1998); Basak (2000); Dumas, Kurshev, and Uppal (2009); Jouini and Napp (2007); Buraschi and Jiltsov (2006); David (2008); Li (2007); and Gallmeyer and Hollifield (2008) provide equilibrium models to study the effects of heterogeneous beliefs on a variety of issues, including asset price volatility, interest rates, equity premium, and the option-implied volatility. More recently, Shefrin (2008) provides a textbook treatment of belief heterogeneity, emphasizing its implications on a number of aspects in asset pricing. Our model differs from these models in two aspects. First, our model specification allows us to isolate belief-dispersion effects from other learning-related effects that also arise in these earlier models, such as effects caused by underestimation of risk and by erroneous average beliefs. Second, and more important, our model provides new implications of heterogeneous beliefs on bond yield movement and bond return predictability.

Our model also differs from the literature that studies the effect of investor preference heterogeneity on asset prices (e.g., Dumas 1989; Wang 1996; Chan and Kogan 2002; Bhamra and Uppal 2009). In particular, Wang analyzes the effect of preference heterogeneity on the yield curve. In another related study, Vayanos and Vila (2007) analyze the effect of the difference in investors' preferred habitats on bond markets. In contrast to these studies, our model generates new implications based on investors' belief dispersion.

The rest of the paper is organized as follows. Section 2 presents the model. Section 3 discusses the effect of heterogeneous expectations on bond market dynamics. Section 4 concludes. We provide all the technical proofs in the Appendix.

2. The Model

We adopt the standard endowment economy of Lucas (1978). To bridge our model with empirical studies of nominal interest rates, we also introduce a price-inflation process and allow two groups of investors holding heterogeneous expectations regarding an unobservable variable that determines the long-run inflation rate. Because of the disagreement on long-run inflation rates, investors speculate in the capital markets. We study a competitive equilibrium in which each investor optimizes consumption and investment decisions based on his or her own expectation. Market clearing conditions determine the equilibrium short rate and asset prices.

2.1 The economy
There is a single consumption good. The aggregate endowment of the consumption good follows

$$\frac{dD_t}{D_t} = \mu_D dt + \sigma_D dZ_D(t), \tag{1}$$

where μ_D and σ_D are constants, and $Z_D(t)$ is a standard Brownian motion process. We assume that the price level p_t (e.g., the CPI index) follows

$$\frac{dp_t}{p_t} = \pi_t dt, \tag{2}$$

where π_t is the inflation rate. Note that π_t follows a linear diffusion process

$$d\pi_t = -\lambda_\pi (\pi_t - \theta_t) dt + \sigma_\pi dZ_\pi(t), \tag{3}$$

where λ_π is the mean-reverting parameter, θ_t is the long-run mean of the inflation rate, σ_π is a volatility parameter, and $Z_\pi(t)$ is a standard Brownian motion independent of $Z_D(t)$. The long-run mean θ_t is unobservable and follows an *Ornstein–Uhlenbeck* process

$$d\theta_t = -\lambda_\theta (\theta_t - \bar{\theta}) dt + \sigma_\theta dZ_\theta(t), \tag{4}$$

where λ_θ is the mean-reverting parameter, $\bar{\theta}$ is the long-run mean of θ_t, σ_θ is a volatility parameter, and $Z_\theta(t)$ is a standard Brownian motion independent of $Z_D(t)$ and $Z_\pi(t)$. Intuitively, we interpret π_t as the current inflation rate and θ_t as the monetary authority's inflation target, which is not directly observable by the public. For convenience, we will refer to θ_t as the inflation target in the rest of the paper.

The aforementioned model structure—inflation rate π_t chasing a time-varying inflation target θ_t—is motivated by the findings of Gurkaynak, Sack, and Swanson (2005). They show that a model in which agents' expectations of long-run inflation stay constant is inconsistent with the significant responses of long-run forward rates to unexpected macroeconomic data releases. Making the inflation target time-varying and unobservable provides a convenient way of modeling investors' heterogeneous expectations. The central part of our analysis is to show that the heterogeneity in investors' expectations can generate significant effects on bond markets.[3]

2.2 Heterogeneous expectations

The economics and finance literature has widely adopted the Bayesian inference framework to model investors' learning processes about unobservable economic variables, such as productivity of the economy and profitability of a specific firm. One line of the literature (e.g., Harris and Raviv 1993; Detemple and Murthy 1994; Morris 1996; Basak 2000) assumes that investors hold heterogeneous prior beliefs about unobservable economic variables. In these models, investors continue to disagree with each other even after they update their beliefs using identical information, but the difference in their beliefs deterministically converges to zero.

---

[3] For simplicity, this paper focuses on agents' disagreement about future inflation rates. In additional analysis, available upon request, we study the effects of agents' disagreement about the real side of the economy, and the main insights are similar.

In another strand of the literature (e.g., Scheinkman and Xiong 2003; Dumas, Kurshev, and Uppal 2009; Buraschi and Jiltsov 2006; David 2008) heterogeneous beliefs arise from investors' different prior knowledge about the informativeness of signals and the dynamics of unobservable economic variables. In support of this approach, Kurz (1994) argues that nonstationarity of economic systems and limited data make it difficult for rational investors to identify the correct model of the economy from alternative ones. More recently, Acemoglu, Chernozhukov, and Yildiz (2007) show that when investors are uncertain about a random variable and the informativeness of a source of signal regarding the random variable, even an infinite sequence of signals from this same source does not lead investors' heterogeneous prior beliefs about the random variable to converge. This is because investors have to update beliefs about two sources of uncertainty using one sequence of signals. Finally, behavioral biases such as overconfidence could also prevent investors from efficiently learning about the informativeness of their signals.

Following this approach, we analyze two groups of investors who hold different prior knowledge about the informativeness of a flow of signals on the inflation target θ_t. In particular, we assume that the signals are not informative, but the heterogeneous prior knowledge leads the two groups to react differently to the signal flow and therefore to possess heterogeneous expectations about future inflation rates. This approach is tractable and generates a stationary process for the difference in investors' beliefs.

More specifically, we assume that all investors observe the following signal flow:

$$dS_t = dZ_S(t),$$

where $Z_S(t)$ is a standard Brownian motion independent of all the Brownian motions introduced earlier. That is, S_t is pure noise. However, investors in the economy believe that S_t is partially correlated with the fundamental shock to θ_t and thus contains useful information.

There are two groups of investors: group A and group B. They have different interpretations of S_t. In particular, group-i ($i \in \{A, B\}$) investors believe that the signal-generating process is

$$dS_t = \phi_i dZ_\theta(t) + \sqrt{1 - \phi_i^2} dZ_S(t), \tag{5}$$

where the parameter $\phi_i \in [0, 1)$ measures the perceived correlation between the signal dS_t and the fundamental shock $dZ_\theta(t)$. For generality, we assume that group-i investors believe that the unobservable process θ_t follows

$$d\theta_t = -\lambda_\theta(\theta_t - \bar{\theta})dt + k_i\sigma_\theta dZ_\theta(t), \tag{6}$$

where $k_i > 0$. That is, group-i investors misperceive the volatility of θ_t by a factor of k_i, and $k_i = 1$ corresponds to the case in which group-i investors

correctly perceive the volatility. This assumption is not crucial for the main implications of this paper, and the reason for introducing it is that this general specification of processes (5) and (6), as will become clear later, allows us to isolate the impact of heterogeneous beliefs.

2.2.1 Benchmark belief. We will evaluate the effects of investors' heterogeneous beliefs on asset price dynamics from the viewpoint of an outside observer, an econometrician, who understands that the signals are not informative. Hence, we first derive the belief of the econometrician. His or her information set at time t is $\{\pi_\tau\}_{\tau=0}^t$. We assume that the econometrician's prior belief about θ_0 has a Gaussian distribution. Since his or her information flow also follows Gaussian processes, his or her posterior beliefs about θ_t must likewise be Gaussian. According to the standard results in linear filtering—e.g., Theorem 12.7 of Liptser and Shiryaev (1977)—the econometrician's belief variance converges to a stationary level at an exponential rate. For our analysis, we will focus on the steady state, in which the belief variance has already reached its stationary level, denoted by $\bar{v}$, which is the positive root of the following quadratic equation of v:

$$\frac{\lambda_\pi^2}{\sigma_\pi^2}v^2 + 2\lambda_\theta v - \sigma_\theta^2 = 0. \tag{7}$$

We denote the econometrician's posterior distribution about θ_t at time t by

$$\theta_t|\{\pi_\tau\}_{\tau=0}^t \sim N\left(\hat{\theta}_t^R, \bar{v}\right),$$

where $\hat{\theta}_t^R$ is the mean of the posterior distribution. Applying Theorem 12.7 of Liptser and Shiryaev (1977), we obtain that

$$d\hat{\theta}_t^R = -\lambda_\theta\left(\hat{\theta}_t^R - \bar{\theta}\right)dt + \frac{\lambda_\pi}{\sigma_\pi}\bar{v}d\widehat{Z}_\pi^R(t), \tag{8}$$

where

$$d\widehat{Z}_\pi^R = \frac{1}{\sigma_\pi}\left[d\pi_t + \lambda_\pi\left(\pi_t - \hat{\theta}_t^R\right)dt\right] \tag{9}$$

is the information shock in $d\pi_t$. Note that $\widehat{Z}_\pi^R$ is a standard Brownian motion from the econometrician's point of view.

2.2.2 Heterogeneous beliefs. We now derive group-A and group-B investors' beliefs about θ_t. These investors' information set at time t includes $\{\pi_\tau, S_\tau\}_{\tau=0}^t$. We denote group-$i$ investors' posterior distribution about θ_t at time t by

$$\theta_t|\{\pi_\tau, S_\tau\}_{\tau=0}^t \sim N\left(\hat{\theta}_t^i, \bar{v}_i\right), \quad i \in \{A, B\},$$

where $\hat{\theta}_t^i$ is the mean of group-i investors' posterior distribution and $\bar{v}_i$ is the steady level of their belief variance. Hereafter we will refer to $\hat{\theta}_t^i$ as their belief. Again, according to Theorem 12.7 of Liptser and Shiryaev (1977), $\bar{v}_i$ is the positive root of the following quadratic equation of v:

$$\frac{\lambda_\pi^2}{\sigma_\pi^2} v^2 + 2\lambda_\theta v - (1 - \phi_i^2) k_i^2 \sigma_\theta^2 = 0, \tag{10}$$

and $\hat{\theta}_t^i$ follows

$$d\hat{\theta}_t^i = -\lambda_\theta(\hat{\theta}_t^i - \bar{\theta}) dt + \frac{\lambda_\pi}{\sigma_\pi} \bar{v}_i d\widehat{Z}_\pi^i(t) + \phi_i k_i \sigma_\theta dS_t, \tag{11}$$

where

$$d\widehat{Z}_\pi^i = \frac{1}{\sigma_\pi} \left[d\pi_t + \lambda_\pi (\pi_t - \hat{\theta}_t^i) dt \right] \tag{12}$$

is the information shock in $d\pi_t$ to group-i investors. Here $\widehat{Z}_\pi^i$ is a standard Brownian motion from group-i investors' point of view. Note that the difference in the two groups' perception, summarized by ϕ_i and k_i, causes them to react differently to the signal flow.

This heterogeneous economy differs from a homogeneous economy through several channels. First, disagreement induces speculative trading between the two groups and trading leads to endogenous relative wealth fluctuation, which in turn affects the equilibrium asset prices. Second, the average belief of the two groups could differ from the econometrician's belief, and the erroneous average belief would affect equilibrium asset price dynamics. Third, the investors' posterior belief variance $\bar{v}_i$ can differ from that of the econometrician's $\bar{v}$. On the one hand, investors feel they have extracted useful information from S_t, which reduces their posterior variance; on the other hand, their subjective perception about the volatility of θ_t (as in Equation (6)) may increase or decrease their posterior variance depending on whether the parameter k_i is larger or smaller than 1.

Although the effects generated by the second and third channels are interesting in their own right, one can capture these effects using a representative-agent framework. We are primarily interested in the disagreement effects through the first channel. The advantage of our specification in Equations (5) and (6) is that it allows us to shut down the second and third channels to isolate the effects from the first one by setting k_i and ϕ_i as follows:

$$\phi_A = -\phi_B = \phi, \tag{13}$$

$$k_A = k_B = \frac{1}{\sqrt{1 - \phi^2}} > 1. \tag{14}$$

These conditions imply that the responses of group-A and -B investors to S_t are in opposite directions but with the same magnitude. As a result, their beliefs diverge in response to S_t. It is clear from Equations (7) and (10) that condition (14) implies that each investor has the same posterior variance as the econometrician because the effect of his or her overestimation of the signal quality on the posterior variance is exactly offset by that of his or her overestimation of the fundamental volatility of θ_t. Taken together, conditions (13) and (14) imply that if the average of the two groups' prior beliefs at time zero is equal to the econometrician's, then in the future their average belief will always keep track of the econometrician's belief. We formally state these properties of investors' beliefs in the following proposition and provide a proof in the Appendix.

Proposition 1. *When conditions* (13) *and* (14) *are satisfied, the difference in the two groups' beliefs has the following process:*

$$d\left(\hat{\theta}_t^A - \hat{\theta}_t^B\right) = -\left(\lambda_\theta + \frac{\lambda_\pi^2}{\sigma_\pi^2}\bar{v}\right)\left(\hat{\theta}_t^A - \hat{\theta}_t^B\right)dt + \frac{2\phi\sigma_\theta}{\sqrt{1-\phi^2}}dS_t.$$

Furthermore, if $\frac{1}{2}(\hat{\theta}_0^A + \hat{\theta}_0^B) = \hat{\theta}_0^R$, then the average of the two groups' beliefs about θ_t always tracks the econometrician's belief:

$$\frac{1}{2}\left(\hat{\theta}_t^A + \hat{\theta}_t^B\right) = \hat{\theta}_t^R.$$

Proposition 1 shows that we can isolate belief-dispersion effects from other learning-related effects, such as those caused by erroneous average belief and underestimation of risk, by choosing a particular set of parameter values in investors' learning processes. This result, to the best of our knowledge, is new to the literature. In the following, we will derive the equilibrium based on the general learning specification of processes (5) and (6). To highlight the belief-dispersion effects, we impose conditions (13) and (14) in the baseline case of our calibration exercise.

2.3 Equilibrium

The difference in investors' beliefs causes speculative trading among themselves. Investors who are more optimistic about θ_t will bet on the rise of future inflation against those more pessimistic investors. Note that from each group's perspective, there are three sources of risks. For group-i investors, the shocks are dZ_D, $d\hat{Z}_\pi^i$, and dS_t. Thus, the markets are complete if investors can trade a risk-free asset and three risky assets that span these three sources of risks. In reality, financial markets offer many securities, such as bonds with different maturities, for investors to construct their bets and to complete the markets. We, hence, analyze an equilibrium with dynamically complete financial markets. As is well known, the prices in a complete-markets equilibrium are not affected

by the structure of the financial markets. Therefore, for brevity, we omit the exact financial securities that investors can trade in this section and leave more details about those securities to the Appendix.

At time zero, group-i investors ($i \in \{A, B\}$) are endowed with α^i fraction of the total wealth of the economy, with $\alpha^i \in (0, 1)$ and $\alpha^A + \alpha^B = 1$. The objective of group-i investors is to maximize their lifetime utility from consumption c_t^i, according to their belief and subject to their budget constraint:

$$\max_{c_t^i} E^i \int_0^\infty e^{-\rho t} u(c_t^i) dt,$$

where E^i is the expectation operator under their probability measure, and ρ is their time-preference parameter. All investors have the logarithmic utility function

$$u(c_t^i) = \log(c_t^i).$$

We use W_t^i to denote group-i investors' wealth at time t, and $\eta(t)$ to denote the wealth ratio between the two groups of investors:

$$\eta(t) \equiv \frac{W_t^B}{W_t^A}.$$

In equilibrium, both groups of investors make their optimal consumption and portfolio choices based on their beliefs, and all financial and good markets clear. We apply the martingale technique (e.g., Cox and Huang 1989; Karatzas, Lehoczky, and Shreve 1987) to construct the equilibrium. The detailed derivation of the equilibrium is reported in the Appendix, and we summarize the most relevant properties of the equilibrium in the following theorem.

Theorem 1. *For the economy defined above, the equilibrium nominal short rate is given by*

$$r_t = \pi_t + \rho + \mu_D - \sigma_D^2. \tag{15}$$

The logarithm of $\eta(t)$ follows a diffusion process in the econometrician's measure:

$$d \ln(\eta_t) = \frac{\lambda_\pi}{\sigma_\pi} (\hat{\theta}_t^B - \hat{\theta}_t^A) \left(\hat{\theta}_t^R - \frac{\hat{\theta}_t^A + \hat{\theta}_t^B}{2} \right) dt + \frac{\lambda_\pi}{\sigma_\pi} (\hat{\theta}_t^B - \hat{\theta}_t^A) d\hat{Z}_\pi^R(t). \tag{16}$$

Furthermore, the time-t nominal price of an asset, which provides a single nominal payoff X_T at time T, is given by

$$P_X(t) = \omega_t^A P_X^A(t) + \omega_t^B P_X^B(t), \tag{17}$$

where ω_t^i is the time-t wealth share of group-i investors, and $P_X^i(t)$ is the nominal price of the asset in a hypothetical economy, in which only group-i investors are present.

Theorem 1 highlights the fluctuation of the wealth distribution in the economy and its effect on the equilibrium prices. Equation (15) shows that the nominal short rate is determined by short-term inflation π_t and the parameters for investors' time preference and the endowment. Equation (16) highlights the impact of belief heterogeneity. In the absence of disagreement, the complete financial markets allow investors to perfectly share risk. Thus, the wealth ratio $\eta(t)$ is constant. Perfect risk sharing becomes impossible when the two groups disagree. As shown in Equation (16), $\eta(t)$ fluctuates over time and is driven by the disagreement between the two groups, $\hat{\theta}_t^A - \hat{\theta}_t^B$. For instance, if group-A investors hold a higher belief about θ_t ($\hat{\theta}_t^A > \hat{\theta}_t^B$), they would bet against group-B investors on the rise of future inflation. A positive shock to future inflation $d\hat{Z}_\pi^R$ would then favor group-A investors and cause wealth to flow from group-B to group-A—i.e., $\eta(t)$ decreases. Note that under conditions (13) and (14), from the viewpoint of the econometrician, the two groups' wealth ratio does not converge to either zero or infinity—i.e., no group is able to eventually dominate the economy in the long run. This is because in our symmetric setting, neither group has a superior learning model.[4]

Finally, Equation (17) shows that the price of an asset can be decomposed into the wealth-weighted average of each group's valuation of the asset in a hypothetical homogeneous economy. This result allows us to represent asset prices in a heterogeneous economy using prices in much simpler homogeneous economies. Thus, the equilibrium is remarkably simple to characterize even in this complex environment with heterogeneous investors. Although this price representation depends on investors' logarithmic preference, it is independent of the specific information structure in our model. Several earlier models (e.g., Detemple and Murthy 1994; Basak 2000) provide similar decompositions of real prices under different information structures. Our model further shows that such a decomposition works for both nominal and real prices.

2.4 The representative investor's belief
As is well known, one can construct a representative investor to replicate the price dynamics in a complete-markets equilibrium with heterogeneous investors. Does this mean that we can simply focus on the representative agent's belief process and ignore the heterogeneity between investors? The answer is no. To understand why, we construct a representative investor for our model.[5] If

---

[4] See Kogan et al. (2006); Dumas, Kurshev, and Uppal (2009); and Yan (2008) for the detailed analysis of irrational traders' survival in asymmetric settings.

[5] See Jouini and Napp (2007) for a recent analysis of the existence of a "consensus" belief for the representative agent in an exchange economy with agents holding heterogeneous beliefs.

we restrict the representative investor to having the same logarithmic preference as the group-A and group-B investors, we can back out the implied belief of the representative agent so that we obtain the same equilibrium dynamics as before. We summarize the result in the following proposition with a proof in the Appendix.

Proposition 2. *To replicate the competitive equilibrium derived in Theorem 1, we can construct a representative agent who has the same logarithmic preference as those investors in the heterogeneous economy. At any point of time, the representative agent's belief about θ_t, denoted by $\hat{\theta}_t^N$, has to be the wealth-weighted average belief of group-A and group-B investors:*

$$\hat{\theta}_t^N = \omega_t^A \hat{\theta}_t^A + \omega_t^B \hat{\theta}_t^B. \tag{18}$$

It is important to stress that the representative agent's belief must equal the wealth-weighted average belief, not only at one point of time but also at all future points. Thus, over time, the representative agent's belief of the inflation target would change in response not only to each group's belief fluctuation but also to the two groups' relative wealth fluctuation, which could be driven by factors that are unrelated to inflation. This implies that in an economy with heterogeneous beliefs, the representative agent's belief does not follow a standard Bayesian learning process with a reasonable prior.

3. Effects of Heterogeneous Expectations on Bond Markets

In this section, we analyze the effects of the two groups' heterogeneous beliefs on bond markets. Theorem 1 allows us to express the price of a bond as the wealth-weighted average of the two groups' bond valuations in homogeneous economies. Hence, we first derive bond prices in homogeneous economies.

3.1 Bond prices in homogeneous economies
Bond prices in homogeneous economies are reported in the following proposition; the proof is in the Appendix.

Proposition 3. *If the economy defined in Section 2 is populated by group-i investors only, the nominal price of a zero-coupon bond with a face value of* $1 *and a maturity of τ years is determined by*

$$B^H\left(\tau, \pi_t, \hat{\theta}_t^i\right) = e^{-a_\pi(\tau)\pi_t - a_\theta(\tau)\hat{\theta}_t^i - b(\tau)}, \tag{19}$$

where

$$a_\pi(\tau) = \frac{1}{\lambda_\pi}(1 - e^{-\lambda_\pi \tau}), \tag{20}$$

$$a_\theta(\tau) = \frac{1}{\lambda_\theta}(1 - e^{-\lambda_\theta \tau}) + \frac{1}{\lambda_\pi - \lambda_\theta}(e^{-\lambda_\pi \tau} - e^{-\lambda_\theta \tau}), \tag{21}$$

$$b(\tau) = \int_0^\tau \left[\lambda_\theta \bar{\theta} a_\theta(s) - \frac{1}{2}\sigma_\pi^2 a_\pi^2(s) - \frac{1}{2}\left(k_i^2 \sigma_\theta^2 - 2\lambda_\theta \bar{v}_i\right) a_\theta^2(s) \right.$$

$$\left. - \lambda_\pi \bar{v}_i a_\pi(s) a_\theta(s) + \rho + \mu_D - \sigma_D^2 \right] ds.$$

Proposition 3 implies that the yield of a τ-year bond in a homogeneous economy,

$$Y^H\left(\tau, \pi_t, \hat{\theta}_t^i\right) = -\frac{1}{\tau}\log(B^H) = \frac{a_\pi(\tau)}{\tau}\pi_t + \frac{a_\theta(\tau)}{\tau}\hat{\theta}_t^i + \frac{b(\tau)}{\tau},$$

is a linear function of two fundamental factors: π_t and $\hat{\theta}_t^i$, which represent the current inflation rate and investors' belief about the inflation target. This linear form belongs to the general affine structure proposed by Duffie and Kan (1996).

The loading on factor π_t, $a_\pi(\tau)/\tau$, has a value of one when the bond maturity τ is zero and monotonically decreases to zero as the maturity increases, suggesting that short-term yields are more exposed to the fluctuations in π_t. This is because of the mean reversion of the current inflation rate to its long-run mean θ_t.

Investors' belief about θ_t determines their expectation of the future short rates. The loading of the bond yield on $\hat{\theta}_t^i$, $a_\theta(\tau)/\tau$, has a hump shape if θ_t mean reverts ($\lambda_\theta > 0$).[6] Since θ_t describes the long-run mean of π_t, as the bond maturity increases from zero, the bond yield becomes more sensitive to the belief about θ_t—that is, as the bond maturity increases from zero to an intermediate value, the loading $a_\theta(\tau)/\tau$ increases. As the bond maturity increases further, the loading $a_\theta(\tau)/\tau$ falls. This is because of the mean reversion of θ_t, which causes any shock to θ_t to eventually die out. This force causes the yields of very long-term bonds to have low exposure to investors' belief about θ_t. This hump shape in the bond yield's loading on $\hat{\theta}_t^i$ is important for understanding later results such as the shape of the bond yield volatility curve and the factor for predicting bond returns.

By combining Theorem 1 and Proposition 3, we can express the nominal price of a τ-year zero-coupon bond at time t as

$$B_t = \omega_t^A B^H\left(\tau, \pi_t, \hat{\theta}_t^A\right) + \omega_t^B B^H\left(\tau, \pi_t, \hat{\theta}_t^B\right), \tag{22}$$

---

[6] In the case where mean reversion is not present ($\lambda_\theta = 0$), the factor loading $a_\theta(\tau)/\tau$ is a monotonically increasing function of bond maturity.

where ω_t^A and ω_t^B are the two groups' wealth shares in the economy, and $B^H(\tau, \pi_t, \hat{\theta}_t^i)$ is the bond price in a homogeneous economy in which only group-i investors are present. The implied bond yield in this heterogeneous economy is

$$Y_t(\tau) = \frac{a_\pi(\tau)}{\tau}\pi_t + \frac{b(\tau)}{\tau} - \frac{1}{\tau}\log\left[\omega_t^A e^{-a_\theta(\tau)\hat{\theta}_t^A} + \omega_t^B e^{-a_\theta(\tau)\hat{\theta}_t^B}\right]. \quad (23)$$

Note that Y_t is not a linear function of investors' beliefs $\hat{\theta}_t^A$ and $\hat{\theta}_t^B$. That is, bond yields in this heterogeneous economy have a nonaffine structure. This structure derives from the market aggregation of investors' heterogeneous valuations of the bond.

3.2 Volatility amplification

Heterogeneous expectations cause investors to take on speculative positions against each other in the financial markets. These speculative positions generate fluctuations in investors' wealth shares, which amplify bond yield volatility. The intuition is as follows. Bond yields are roughly determined by investors' wealth-weighted average belief about future interest rates, as in Equation (23). Since investors who are more optimistic about future rates bet on these rates rising against those more pessimistic investors, any positive news about future rates would cause wealth to flow from pessimistic investors to optimistic investors, giving the optimistic belief a greater weight in bond yields. The relative-wealth fluctuation thus amplifies the impact of the initial news on bond yields. As a result, a higher belief dispersion increases the relative-wealth fluctuation, which in turn increases the bond yield volatility. We summarize this intuition in Proposition 4 and provide a proof in the Appendix.

Proposition 4. *Bond yield volatility increases with the two groups' belief dispersion.*

This volatility amplification mechanism can help to explain the "excess volatility puzzle" of bond yields. Shiller (1979) shows that the observed bond yield volatility exceeds the upper limits implied by the expectations hypothesis and the observed persistence in short rates. Gurkaynak, Sack, and Swanson (2005) also document that bond yields exhibit excess sensitivity to particular shocks, such as macroeconomic announcements. Furthermore, Piazzesi and Schneider (2006) find that by estimating a representative-agent based asset pricing model with recursive utility preferences and exogenous consumption growth and inflation, the model explains a smaller fraction of the observed volatility in long yields than in short yields. Relating to this literature, Proposition 4 shows that extending standard representative-agent models with heterogeneous expectations can help to account for the observed high bond yield

volatility. In Section 3.4, we provide a calibration exercise to illustrate the magnitude of this mechanism.

3.3 Time-varying risk premium

Fluctuations in investors' belief dispersion and relative wealth also cause risk premia to vary over time. To examine the time variation in risk premia, we derive the dynamics of the stochastic discount factor in the following proposition, with a proof in the Appendix.

Proposition 5. *From the viewpoint of the econometrician, the state price density for real cash flow has the following process:*

$$\frac{dM_t}{M_t} = -(\rho + \mu_D - \sigma_D^2)dt - \sigma_D dZ_D - \frac{\lambda_\pi}{\sigma_\pi}\left(\hat{\theta}_t^R - \sum_{i=A}^{B} \omega_t^i \hat{\theta}_t^i\right)d\hat{Z}_\pi^R, \quad (24)$$

where $\hat{\theta}_t^R$ is the econometrician's belief about θ_t, and $d\hat{Z}_\pi^R$ is the information shock defined in Equation (9).

Proposition 5 shows that from the viewpoint of the econometrician, the market price of risk (risk premium per unit of risk) for the aggregate endowment shock dZ_D is σ_D, while the market price of risk for the information shock $d\hat{Z}_\pi^R$ is proportional to $\hat{\theta}_t^R - \sum_{i=A}^{B} \omega_t^i \hat{\theta}_t^i$, the difference between the econometrician's belief about θ_t and the two groups' wealth-weighted average belief.

In the benchmark case where investors are homogeneous and have the same belief as the econometrician ($\hat{\theta}_t^A = \hat{\theta}_t^B = \hat{\theta}_t^R$), the risk premium for the information shock $d\hat{Z}_\pi^R$ is zero and the market offers a constant price of risk only for the aggregate endowment shock dZ_D. When the two groups' beliefs are different, however, there is a nonzero risk premium for the information shock $d\hat{Z}_\pi^R$. Moreover, this premium varies over time with the relative wealth fluctuation across the two groups of investors.

The intuition is as follows. Suppose that the two groups' wealth-weighted average belief about θ_t is above the econometrician's belief. Then, relative to the econometrician, investors are more optimistic about the rise of π_t in the future, and so they are more optimistic about assets that are positively exposed to $d\hat{Z}_\pi^R$ (i.e., those prices are positively correlated with π_t). From the econometrician's point of view, those assets appear "expensive" and have low risk premia. Similarly, those assets would have high risk premia if the wealth-weighted average belief is below the econometrician's belief. The relative wealth fluctuation across the two groups affects the difference between their wealth-weighted average belief and the econometrician's belief, and thus contributes to the time variation in the risk premium.

In the next subsection, we provide a calibration exercise to show that a modest amount of belief dispersion can generate sufficient time variation in the

risk premium to explain the failure of the expectations hypothesis, and that the time variation of the risk premium is related to a tent-shaped linear combination of forward rates.

3.4 Calibration

This subsection illustrates the impact of investors' heterogeneous expectations on bond markets by simulating the heterogeneous economy based on a set of calibrated model parameters. Theorem 1 implies that the nominal short rate follows

$$dr_t = -\lambda_\pi \big[r_t - \big(\theta_t + \rho + \mu_D - \sigma_D^2\big)\big]dt + \sigma_\pi dZ_\pi.$$

The short rate mean reverts to a time-varying long-run mean $\theta_t + \rho + \mu_D - \sigma_D^2$. Balduzzi, Das, and Foresi (1998) estimate a two-factor model of the short rate and its long-run mean, with the same structure as ours. Based on the data between 1952 and 1993, they find that the long-run mean of the short rate moves slowly, with a mean-reversion parameter of 0.07. Since this mean-reversion parameter corresponds to λ_θ in our model, we choose λ_θ to be 0.07, which implies that it takes $\ln(2)/\lambda_\theta = 9.9$ years for the effect of a shock to the long-run mean of the short rate to die out by half. Balduzzi, Das, and Foresi (1998) also show that the mean-reversion parameter of the short rate (λ_π in our model) ranges from 0.2 to 3 in different sample periods. We choose λ_π to be 1, which implies that it takes $\ln(2)/\lambda_\pi = 0.69$ year for the difference between the short rate and its long-run mean to die out by half.[7]

We choose $\sigma_\pi = 1.25\%$ to match the short-rate volatility in the data, and set $\sigma_\theta = 1.2\%$ so that the volatility of θ_t is 0.35% per month, the middle point of the range from 0.1% to 0.6% estimated by Balduzzi, Das, and Foresi (1998). Furthermore, we choose $\mu_D = 2\%$ and $\sigma_D = 2\%$ to match the aggregate consumption growth rate and volatility in the data. Investors' time preference parameter ρ is set at 2%. We choose the following initial conditions for our simulation. The two groups have an equal wealth share at $t = 0$, i.e., $\omega_0^A = \omega_0^B = 0.5$; both π_0 and θ_0 start from their steady-state value $\bar{\theta}$, which we set at 2%; the two groups also share an identical prior belief about θ_0 equal to the steady-state value $\bar{\theta}$: $\hat{\theta}_0^A = \hat{\theta}_0^B = \bar{\theta}$. Finally, we impose conditions (13) and (14) on parameters ϕ_A, ϕ_B, k_A, and k_B so that the posterior variance of all investors is the same as that of the econometrician and the average belief of the two groups of investors coincides with the econometrician's belief.

Parameter ϕ directly affects the amount of belief dispersion between the two groups. We choose $\phi = 0.75$ to generate a modest amount of belief dispersion: In our simulated data, the average dispersion between the two groups, $|\hat{\theta}_t^A - \hat{\theta}_t^B|$,

---

[7] These two mean-reversion parameters affect the magnitude of agents' belief-dispersion effect. Intuitively, a larger λ_θ parameter causes θ_t to revert faster to its long-run mean $\bar{\theta}$, therefore making agents' belief dispersion about θ_t less important for bond prices; a larger λ_π parameter causes π_t to revert faster to θ_t, therefore making agents' belief dispersion about θ_t more important for bond prices.

is only 1.70%. This amount is rather modest compared with the typical dispersion observed in surveys of professional economists' inflation expectations (see footnote 1).[8] All the model parameters are summarized below:

$$\lambda_\theta = 0.07, \ \lambda_\pi = 1, \ \sigma_\pi = 1.25\%, \ \sigma_\theta = 1.2\%, \ \mu_D = 2\%, \ \sigma_D = 2\%,$$
$$\phi = 0.75, \ \omega_0^A = \omega_0^B = 0.5, \ \pi_0 = \theta_0 = \hat{\theta}_0^A = \hat{\theta}_0^B = \bar{\theta} = 2\%, \ \rho = 2\%. \ (25)$$

Based on these model parameters, we simulate the economy for 50 years at daily frequency and extract bond yields and forward rates for various maturities at the end of each month. The length of 50 years roughly matches the sample duration used in most empirical studies of the yield curve. The simulation is repeated 10,000 times.

To ensure that our calibration results are not driven by the specific parameter values chosen in the list of parameters (25), we also perform a series of robustness checks by varying the values of several key parameters: λ_θ, λ_π, σ_θ, ϕ, and k. These checks show that our main results are robust to a wide range of parameter values. Only in the extreme cases in which the parameter combination implies little disagreement, some results are weakened or disappear. For brevity, we do not report these robustness checks in the paper and, instead, make the results available by request.

3.4.1 Yield volatility curve. Figure 1 plots the monthly bond yield volatility, defined as the standard deviation of yield changes, for different maturities from zero to ten years. The upper solid line corresponds to the yield volatility in the heterogeneous economy. The two dashed lines around the volatility curve provide the 95th and 5th percentiles of the volatility estimates across the 10,000 simulated paths. As the maturity increases from zero to three years, the yield volatility increases from 36 to above 41 basis points per month. As the maturity further increases, the yield volatility then starts to fall slightly. The magnitude and shape of this volatility curve are similar to those estimated in Dai and Singleton (2003).

To illustrate the volatility amplification effect discussed in Section 3.2, we also compute the volatility curve in a hypothetical homogeneous economy in which all investors hold the equal-weighted average belief of the two groups in the above simulated heterogeneous economy (this average belief also coincides with the econometrician's belief, as shown in Proposition 1). Note that the average belief reflects the changes in the two groups' beliefs, but not their relative wealth fluctuation. As a result, the volatility curve in the homogeneous economy does not capture the volatility amplification effect caused by the two groups' relative wealth fluctuation. The lower solid line in Figure 1 plots the

---

[8] This amount of belief dispersion also leads to an average wealth share volatility of 15% per year for each group in the simulated data. Although we are not aware of any formal estimate of the relative wealth fluctuation between different investor groups in the economy, the 15% volatility appears feasible. Moreover, in the mutual fund and hedge fund industry, it is common to see large swings in asset under management among strategies (e.g., value versus growth strategies).
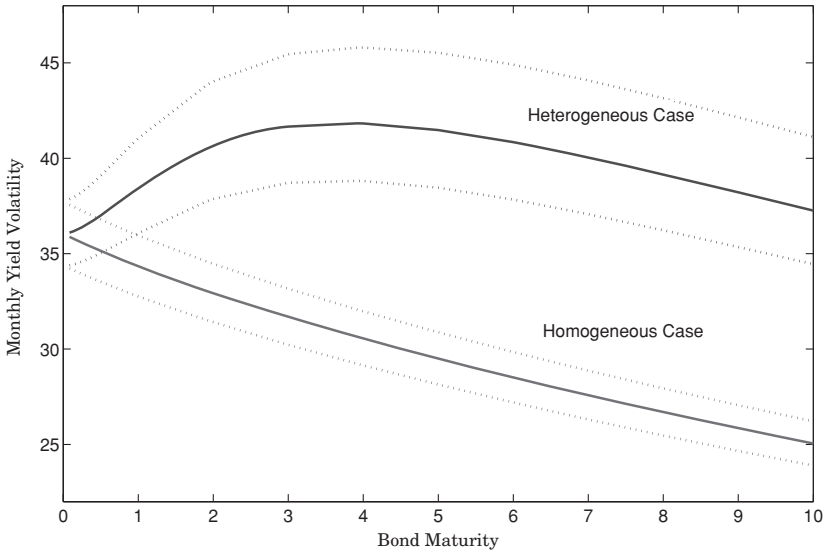
Figure 1
The Term Structure of Bond Yield Volatility
Using parameters specified in Equation (25), the economy is simulated for 50 years to calculate bond yield volatility, defined as the standard deviation of yield changes, for zero-coupon bonds with maturities ranging from 0 to 10 years. The simulation is iterated 10,000 times, and the figure plots the average (solid line), 95th percentile, and 5th percentile (dashed lines) of the estimated volatility across the 10,000 paths on bond maturity. Similar simulations are also performed on a homogeneous economy with a representative agent holding the *equal*-weighted average belief of the two groups in the heterogeneous economy. The plots at the bottom of the figure correspond to the average (solid line), 95th percentile, and 5th percentile (dashed lines) of the estimated volatility across the 10,000 paths in the homogeneous economy.

volatility curve in this homogeneous economy. The volatility drops monotonically from 36 to 25 basis points per month as the bond maturity increases from zero to ten years. The difference between the two solid lines measures the volatility amplification effect induced by wealth fluctuation. This effect is small at short maturities but becomes substantial when bond maturity increases. For the 10-year bond, this amplification effect is 12 basis points per month, or roughly one-third of the total bond yield volatility.

Why does the volatility curve have a hump shape in the heterogeneous case, but a monotonically decreasing shape in the homogeneous case? In the homogeneous case, the bond yield is a linear combination of two factors:

$$Y_t(\tau) = \frac{a_\pi(\tau)}{\tau}\pi_t + \frac{a_\theta(\tau)}{\tau}\hat{\theta}_t^R + \frac{b(\tau)}{\tau}.$$

From our earlier discussion, the loading on the first factor π_t, $\frac{a_\pi(\tau)}{\tau}$, decreases monotonically with τ, while the loading on the second factor $\hat{\theta}_t^R$, $\frac{a_\theta(\tau)}{\tau}$, has a hump shape. The monotonically decreasing shape of the volatility curve reflects that the contribution of the first factor of the bond yield volatility dominates that of the second factor.

1450

To simplify our discussion of the heterogeneous case, we approximate Equation (23) by a linear form:

$$Y_t(\tau) \approx \frac{a_\pi(\tau)}{\tau}\pi_t + \frac{a_\theta(\tau)}{\tau}\left(\omega_t^A\hat{\theta}_t^A + \omega_t^B\hat{\theta}_t^B\right) + \frac{b(\tau)}{\tau}.$$

Note that the second factor now becomes the wealth-weighted average belief $\omega_t^A\hat{\theta}_t^A + \omega_t^B\hat{\theta}_t^B$, which is more volatile than the second factor in the homogeneous economy, $\hat{\theta}_t^R$. In other words, the wealth fluctuation effect makes the second factor more volatile. The volatility curve displays a hump shape when the wealth fluctuation effect is strong enough.

3.4.2 Campbell–Shiller bond yield regression. This section demonstrates that the time variation in the risk premium in our model can help explain the failure of the expectations hypothesis. The expectations hypothesis posits that an investor in the bond market should be indifferent about the investment in the short rate or in a long-term bond over the same short period. Despite its intuitive appeal, this prediction is rejected by many empirical studies (e.g., Fama and Bliss 1987; Campbell and Shiller 1991).

In particular, Campbell and Shiller (1991) run the following regression:

$$Y_{t+1}(n-1) - Y_t(n) = \alpha_n + \beta_n\frac{Y_t(n) - Y_t(1)}{n-1}, \qquad (26)$$

where $Y_t(n)$ is the n-month yield at month t, α_n is the regression constant, and β_n is the regression coefficient. They show that the expectations hypothesis is equivalent to the following null hypothesis for regression (26):

$$\beta_n = 1.$$

Intuitively, when the yield spread, $Y_t(n) - Y_t(1)$, is positive, the long-term bond yield is expected to rise (or the long-term bond price is expected to fall), because otherwise an investor cannot be indifferent about investing in the long-term bond or the short rate.

The regression results in Panel A of Table 1 are collected from Table 10.3 of Campbell, Lo, and MacKinlay (1997), which uses 40 years of U.S. Treasury bond yield data from 1952 to 1991. It shows that β_n starts with a value of 0.003 for a two-month yield, and then monotonically decreases as the bond maturity increases. Note that β_n eventually takes a value of -4.226 for a 10-year yield. All these coefficients are significantly different from 1 (the null), and the coefficient of the 10-year yield is significantly negative. Taken together, these regression results reject the expectations hypothesis: When the yield spread is positive, the long-term bond yield tends to fall rather than rise.

This pattern, however, is a natural implication of our model: Suppose that the wealth-weighted average belief about the future short rates is higher than the econometrician's belief. On the one hand, this implies that investors discount

Table 1
Coefficients of yield change regressions

n	2	3	6	12	24	48	120
			Panel A: Results from Campbell–Lo–MacKinlay				
β_n	0.003	−0.145	−0.835	−1.435	−1.448	−2.262	−4.226
s.e.	(0.191)	(0.282)	(0.442)	(0.599)	(1.004)	(1.458)	(2.076)
		Panel B: Results from our simulation of a heterogeneous economy					
β_n	−1.037	−1.104	−1.304	−1.699	−2.435	−3.231	−3.499
s.e.	(0.516)	(0.541)	(0.620)	(0.785)	(1.110)	(1.632)	(1.986)

This table reports the estimates of β_n in Equation (26) and their standard errors for bond maturities of n months. Panel A is taken from Table 10.3 of Campbell, Lo, and MacKinlay (1997), which uses U.S. Treasury bond yield data from 1952–1991. Panel B reports the mean and standard deviation of the estimates of β_n across the 10,000 simulated paths of the heterogeneous economy with parameters from (25).

long-term bonds more heavily, which leads to higher long-term bond yields and so larger yield spreads; on the other hand, it also implies that the long-term bond prices appear "cheap" from the econometrician's point of view—i.e., the long-term bond prices are expected to rise and bond yields are expected to fall. Taken together, a high wealth-weighted average belief implies both large yield spreads and falling long-term bond yields in the future.

To examine whether this mechanism can explain the failure of the expectations hypothesis, we simulate our economy 10,000 times using the parameters summarized in (25). For each simulated path, we run regression (26) using our simulated bond yield data. Panel B of Table 1 reports the average regression coefficients and their standard errors. The average of the regression coefficients decreases monotonically from −1.037 to −3.499 as bond maturity increases from two months to ten years, with a similar trend and magnitude to that in Panel A. These coefficients are also significantly lower than zero based on the standard errors across the 10,000 sample paths.[9] Note that the null hypothesis holds in a homogeneous economy with each investor holding the same belief as the econometrician. Therefore, extending a standard asset pricing model with modest heterogeneous expectations offers a potential explanation for the failure of the expectations hypothesis in the data.

The literature often attributes the failure of the expectations hypothesis to time-varying risk premia. Dai and Singleton (2002) find that certain classes of affine term structure models with time-varying risk premia are able to match the aforementioned bond yield regression results. However, the economic determinants of the time-varying risk premia still remain elusive. A few recent studies argue for time-varying risk preference of the representative investor (e.g., Wachter 2006; Buraschi and Jiltsov 2007) and time-varying rare disaster risk (e.g., Gabaix 2008), while our model proposes a new mechanism based on investors' heterogeneous expectations.

---

[9] By simulating 10,000 paths, we are able to control for simulation errors and show with sufficient confidence that our model implications match with empirical findings. However, it may not be appropriate to directly compare the standard errors computed from the cross-section of simulated paths with those in empirical studies based on one path of data.

3.4.3 Cochrane–Piazzesi bond return regression.
Cochrane and Piazzesi (2005) find that a single factor based on a tent-shaped linear combination of forward rates predicts excess returns on bonds with maturity ranging from two to five years.[10] Moreover, this single factor substantially improves the predictive power of the forward spread (an *n*-year forward rate minus a one-year spot rate) in Fama and Bliss (1987), who regress excess returns of *n*-year bonds on *n*-year forward spreads. Can our model explain this interesting phenomenon? To examine this question, we run the regressions by Cochrane and Piazzesi (2005) and Fama and Bliss (1987) using our simulated data.

Following Cochrane and Piazzesi (2005), for each of the 10,000 simulated paths of our heterogeneous economy, we regress excess bond returns on the one-year bond yield and three- and five-year forward rates:[11]

$$rx_{t+1}(n) = \beta_0(n) + \beta_1(n)Y_t(1) + \beta_3(n)F_t(3) + \beta_5(n)F_t(5) + \varepsilon_{t+1}(n),$$
$$n = 2, 3, 4, 5, \tag{27}$$

where $rx_{t+1}(n)$ is the one-year excess bond return defined by

$$rx_{t+1}(n) \equiv \log B_{t+1}(n-1) - \log B_t(n) - Y_t(1),$$

where $B_t(n)$ is the time-*t* price of an *n*-year zero-coupon bond, $Y_t(1)$ is the one-year bond yield, and

$$F_t(n) \equiv \log B_t(n-1) - \log B_t(n)$$

is the log forward rate at time *t* for loans to be made between time $t + n - 1$ and $t + n$.

The top panel of Figure 2 plots the average (across simulated paths) slope coefficients

$$[\beta_1(n), \beta_3(n), \beta_5(n)]$$

for different bond maturities ($n = 2, 3, 4, 5$). The plot shows a pattern that is strikingly similar to the finding of Cochrane and Piazzesi (2005): A tent-shaped combination of forward rates forecasts holding period returns of bonds at all maturities, with longer maturity bonds having greater loadings on this factor. Panel A of Table 2 reports the average coefficients, together with the standard errors and average regression R^2. All coefficients are statistically different from zero and the regression R^2 is around 20% for all maturities.

We also follow the two-stage regression by Cochrane and Piazzesi (2005) to describe bond premia of all maturities by a single factor. First, we regress the

---

[10] See Dai, Singleton, and Yang (2004) and Cochrane and Piazzesi (2004) for more discussions on this result.

[11] To avoid collinearity problems, we do not include the two- and four-year forward rates in the regression.
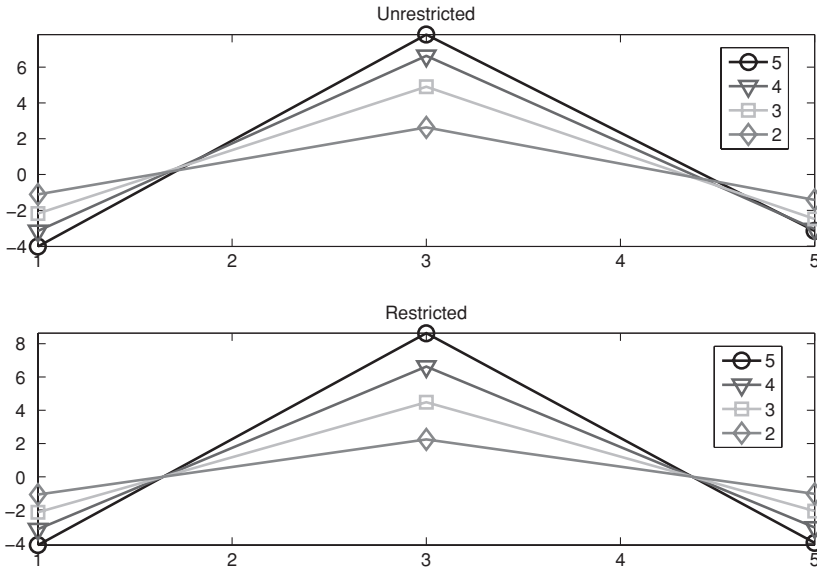
Figure 2
Coefficients in Cochrane–Piazzesi Bond Return Regression
Using parameters specified in Equation (25), we simulate the heterogeneous economy for 50 years. For each simulated path, we run Cochrane–Piazzesi regressions (27) and (29). We iterate the simulation and regressions 10,000 times, and this figure plots the average regression coefficients across the simulated paths. The top panel is based on the unrestricted coefficients $[\beta_1(n), \beta_3(n), \beta_5(n)]$ in regression (27) for two- to five-year bonds, while the bottom panel is based on the restricted coefficients $[b(n)\gamma_1, b(n)\gamma_3, b(n)\gamma_5]$ in regression (29).

average (across maturity) excess return on the forward rates:

$$\frac{1}{4}\sum_{n=2}^{5} rx_{t+1}(n) = \gamma_0 + \gamma_1 Y_t(1) + \gamma_3 F_t(3) + \gamma_5 F_t(5) + \bar{\varepsilon}_{t+1}$$

to identify the tent-shaped factor TF_t,

$$TF_t = \gamma_0 + \gamma_1 Y_t(1) + \gamma_3 F_t(3) + \gamma_5 F_t(5). \tag{28}$$

Then, we regress the individual excess returns on the common factor identified in the first step:

$$rx_{t+1}(n) = b(n)TF_t + \varepsilon_{t+1}(n), \quad n = 2, 3, 4, 5. \tag{29}$$

As in Cochrane and Piazzesi (2005), we also impose $1/4\sum_{n=2}^{5} b(n) = 1$ to separately identify the values of γ_i ($i = 0, 1, 3, 5$) and $b(n)$ ($n = 2, 3, 4, 5$). This two-stage regression puts the following restrictions on the slope coefficients of regression (27):

$$\beta_1(n) = b(n)\gamma_1, \quad \beta_3(n) = b(n)\gamma_3, \quad \beta_5(n) = b(n)\gamma_5.$$

Table 2
Coefficients in Cochrane–Piazzesi bond return regression

	Panel A: Unrestricted				Panel B: Restricted			
	$\beta_1(n)$	$\beta_3(n)$	$\beta_5(n)$	R^2	$\beta_1(n)$	$\beta_3(n)$	$\beta_5(n)$	R^2
$rx(2)$	−1.11	2.63	−1.42	20.8%	−1.07	2.24	−1.03	20.7%
s.e.	(0.05)	(0.38)	(0.38)		(0.05)	(0.40)	(0.40)	
$rx(3)$	−2.18	4.91	−2.48	20.1%	−2.13	4.48	−2.06	20.0%
s.e.	(0.09)	(0.79)	(0.78)		(0.09)	(0.80)	(0.79)	
$rx(4)$	−3.15	6.63	−3.05	19.5%	−3.15	6.62	−3.05	19.5%
s.e.	(0.14)	(1.18)	(1.18)		(0.14)	(1.18)	(1.17)	
$rx(5)$	−4.02	7.82	−3.15	19.1%	−4.11	8.64	−3.97	19.1%
s.e.	(0.18)	(1.56)	(1.56)		(0.18)	(1.54)	(1.53)	

Using parameters specified in Equation (25), we simulate the heterogeneous economy for 50 years. For each simulated path, we run Cochrane–Piazzesi regressions (27) and (29). We iterate the simulation and regressions 10,000 times, and this table reports the average and standard errors of regression coefficients across the simulated paths. Panel A reports the unrestricted coefficients $[\beta_1(n), \beta_3(n), \beta_5(n)]$ and R^2 for regression (27) for two- to five-year bonds, while Panel B reports the restricted coefficients $[b(n)\gamma_1, b(n)\gamma_3, b(n)\gamma_5]$ and R^2 for regression (29).

The bottom panel of Figure 2 plots the average (across the simulated paths) restricted slope coefficients

$$[b(n)\gamma_1, b(n)\gamma_3, b(n)\gamma_5]$$

for different bond maturities. The plot shows a clear tent-shaped pattern similar to that in the unrestricted regressions, confirming that the same single factor predicts returns of all bonds. Panel B of Table 2 reports the standard errors of these restricted coefficients (across the simulated paths), together with the average coefficients, and regression R^2. All coefficients are statistically different from zero. The R^2 of each restricted regression is almost identical to the corresponding unrestricted regression R^2, suggesting that the single factor summarizes most of the predictive power in all the forward rates.

We also run the Fama–Bliss regressions: regressing n-year bond excess returns on n-year forward spreads. While the forward yields forecast bond premia, the predictive power is substantially weaker: The regression R^2 is less than 10% for all maturities. This result is also consistent with the finding of Cochrane and Piazzesi (2005) that the linear combination of forward rates has a stronger return predictive power than the maturity-specific forward spreads.

The intuition behind these results can be summarized as follows. As discussed earlier, a higher wealth-weighted average belief about the future short rates leads to higher future bond returns. Moreover, as will be elaborated next, a higher wealth-weighted average belief means a larger value of the tent-shaped factor TF_t. As a result, a larger value of the tent-shaped factor TF_t predicts higher future bond returns.

To understand why a higher wealth-weighted average belief means a larger value of TF_t, we first use Equation (22) to derive the τ-year ahead forward rate

at time t, $F_t(\tau)$:

$$F_t(\tau) = a'_\pi(\tau)\pi_t + b'(\tau) + a'_\theta(\tau)\tilde{\theta}_t, \tag{30}$$

where $\tilde{\theta}_t$ is a weighted average of the two groups' beliefs:

$$\tilde{\theta}_t = \frac{B^H\left(\tau, \pi_t, \hat{\theta}_t^A\right)}{B_t}\omega_t^A\hat{\theta}_t^A + \frac{B^H\left(\tau, \pi_t, \hat{\theta}_t^B\right)}{B_t}\omega_t^B\hat{\theta}_t^B. \tag{31}$$

Note that $a'_\theta(\tau)$, the forward rate's loading on $\tilde{\theta}_t$, has a hump shape with respect to τ. That is, the forward rates for the intermediate future are more sensitive to the belief about θ_t than the forward rates for the near and very distant future. This is a direct implication from the result, noted earlier in Section 3.1, that the intermediate term bond yields are most sensitive to the belief about θ_t.[12] Under the parameters specified in (25), $a'_\theta(\tau)$ attains its maximum at around $\tau = 3$. That is, the three-year forward rate is more sensitive to $\tilde{\theta}_t$ than the one-year and five-year forward rates. Note that $\tilde{\theta}_t$ fluctuates with the wealth distribution of the two groups. As the optimistic group's wealth share goes up, $\tilde{\theta}_t$ increases, and so the hump shape of $a'_\theta(\tau)$ implies that the three-year forward rate increases more than the one-year and five-year forward rates. This leads to a higher value of the tent-shaped factor since it has a high loading on the three-year forward rate but low loadings on the one- and five-year forward rates: According to the estimates of regression (28) from our simulated data, $TF_t = -2.1 - 2.6Y_t(1) + 5.3F_t(3) - 2.3F_t(5)$.[13]

Interestingly, one of the key elements in our explanation—the hump shape of $a'_\theta(\tau)$—is consistent with the empirical finding of Gurkaynak, Sack, and Swanson (2005) about the response of forward rates to various macroeconomic shocks. These authors regress daily changes in forward rates on the surprise component of various macroeconomic data releases, including capacity utilization, consumer confidence, CPI (core), employment cost index, GDP (advance), initial claims, leading indicators, NAPM, new home sales, non-farm payroll, PPI (core), retail sales, and unemployment rate. The surprise component of each release is computed as the released value less the market expectation, which is measured by the median market forecast as compiled and published by Money Market Services. They find that 11 out of the 13 surprise variables have significant impact on forward rates up to five years ahead. Of more relevance to our study, Figure 2 of their paper shows that the impact of eight of these variables is most pronounced on the three-year-ahead forward rate.

---

[12] Note that there is a no-arbitrage relationship between the spot rates and forward rates: the time-t forward rate for a loan from $t + \tau_1$ to $t + \tau_2$ is $-(Y_t(\tau_2)\tau_2 - Y_t(\tau_1)\tau_1)/(\tau_2 - \tau_1)$. Hence, the fact that the intermediate term bond yields are more sensitive to the beliefs about θ_t than the short-term and very long-term yields implies that the intermediate forward rates have higher exposure to the beliefs about θ_t than the forward rates in the near and distant future.

[13] The estimates of γ_0 through γ_5 are the average of the estimates across the 10,000 simulated paths, and are all significantly different from zero.

In other words, the response of the forward curve has a hump shape with a peak around the three-year-ahead forward rate. To the extent that these surprise variables represent shocks to $\tilde{\theta}_t$ in Equation (31), this finding suggests that, consistent with our calibrations, $a'_\theta(\tau)$ has a hump shape with a peak at around three years.

3.5 Trading volume

Heterogeneous expectations cause investors to take on speculative positions against each other in the financial markets. These speculative positions can cause fluctuations in investors' wealth shares upon the arrival of new information. Investors then trade with each other to rebalance their positions. Intuitively, when belief dispersion increases, the size of their speculative positions becomes larger. This, in turn, leads to a higher volatility of investors' wealth and therefore a larger trading volume in the markets. Clearly, trading volume is affected by the structure of the financial securities available to investors. In order to analyze trading volume, we consider a set of securities whose return processes are given by Equations (A1)–(A3) in the Appendix. These securities' returns have constant volatility and are independent of each other. Moreover, these securities make the markets dynamically complete. We use the volatility of one group's position changes as a measure of trading volume.[14] This measure corresponds to the conventional volume measure in a discrete-time setup. We summarize the effect of investors' belief dispersion on trading volume in Proposition 5, and provide a formal derivation and further discussion on our volume measure in the Appendix.

Proposition 5. *The trading volume of these securities* (*the fluctuation in investors' speculative positions*) *increases with the belief dispersion between the two groups of investors.*

This proposition links trading volume to belief dispersion and hence provides implications on the joint dynamics of trading volume with other variables. For example, as noted in Proposition 4, higher belief dispersion also leads to higher bond yield volatility. Moreover, as shown in Equation (19), $B^H(\tau, \pi_t, \hat{\theta}_t^i)$ is convex in $\hat{\theta}_t^i$. Hence, applying the Jensen's inequality to Equation (22) implies that higher belief dispersion leads to a higher bond price, and therefore a lower bond yield. This effect disappears when bond maturity goes to zero since the convexity in $B^H(\tau, \pi_t, \hat{\theta}_t^i)$ disappears when τ goes to zero. That is, higher belief dispersion reduces long-term bond yields but has little impact on the yields of short-term bonds. Therefore, higher belief dispersion reduces the slope of the

---

[14] Note that trading volume is specific to the menu of securities available in the markets. In complete markets, introducing a different menu of securities does not affect the equilibrium prices, but it could change trading volume. Nevertheless, analyzing trading volume can still lead to useful insights about the market dynamics because it is easily measurable and the menu of securities available in the markets does not change dramatically over time.

yield curve. Taken together, our model provides a testable prediction on the joint dynamics of trading volume and the yield curve: higher belief dispersion simultaneously implies higher trading volume, higher bond yield volatility, and lower yield curve slope.

4. Conclusion

In this paper, we present a dynamic equilibrium model with two groups of investors holding heterogeneous expectations about future economic conditions. The heterogeneous expectations cause investors to take on speculative positions against each other and therefore generate endogenous relative wealth fluctuation. The relative wealth fluctuation amplifies asset price volatility and contributes to the time variation in equilibrium risk premia. We show that a modest amount of heterogeneous expectations can help resolve several challenges encountered by standard representative-agent models, including the "excessive volatility" of bond yields, the failure of the expectations hypothesis, and the ability of a tent-shaped linear combination of forward rates to predict bond returns.

Appendix: Proofs

Proof of Proposition 1
From Equations (11), (13), and (14), we take the difference of $d\hat{\theta}_t^A$ and $d\hat{\theta}_t^B$:

$$d\hat{\theta}_t^A - d\hat{\theta}_t^B = -\lambda_\theta(\hat{\theta}_t^A - \hat{\theta}_t^B)dt + \frac{\lambda_\pi}{\sigma_\pi}\bar{v}\left[d\widehat{Z}_\pi^A(t) - d\widehat{Z}_\pi^B(t)\right] + \frac{2\phi\sigma_\theta}{\sqrt{1-\phi^2}}dS_t$$

$$= -\left(\lambda_\theta + \frac{\lambda_\pi^2}{\sigma_\pi^2}\bar{v}\right)(\hat{\theta}_t^A - \hat{\theta}_t^B)dt + \frac{2\phi\sigma_\theta}{\sqrt{1-\phi^2}}dS_t.$$

Based on Equations (11), (13), and (14) and the initial condition $(\hat{\theta}_0^A + \hat{\theta}_0^B)/2 = \hat{\theta}_0^R$, it is direct to verify that $(\hat{\theta}_t^A + \hat{\theta}_t^B)/2 = \hat{\theta}_t^R$.

Derivation of the equilibrium and proof of Theorem 1
We introduce three zero-net supply risky securities to complete the markets:

$$\frac{dp_\pi}{p_\pi} = \mu_\pi(t)dt + d\pi_t, \tag{A1}$$

$$\frac{dp_S}{p_S} = \mu_S(t)dt + dS_t, \tag{A2}$$

$$\frac{dp_c}{p_c} = \mu_c(t)dt + dZ_D(t), \tag{A3}$$

which we refer to as security π, security S, and security C, respectively. One can also introduce any other three securities that complete the markets, and the resulting equilibrium is not affected. To group-i investors, the dynamics of π_t is

$$d\pi_t = -\lambda_\pi(\pi_t - \hat{\theta}_t^i)dt + \sigma_\pi d\widehat{Z}_\pi^i(t). \tag{A4}$$

1458

Thus, Equation (A1) implies that the price dynamics of security π to group-i investors is

$$\frac{dp_\pi}{p_\pi} = \left[\mu_\pi(t) - \lambda_\pi \pi_t + \lambda_\pi \hat{\theta}^i_t\right]dt + \sigma_\pi d\widehat{Z}^i_\pi(t).$$

Therefore, group-i investors' perceived expected return is $\mu^i_\pi(t) = \mu_\pi(t) - \lambda_\pi \pi_t + \lambda_\pi \hat{\theta}^i_t$. Also note that all investors observe dS_t and $dZ_D(t)$. Hence, they agree on the expected returns of securities S and C (i.e., $\mu_S(t)$ and $\mu_c(t)$).

Market completeness implies the existence of a unique state price density process to each group. We denote $\xi_i(t)$ as state price density of group-i investors, which, according to no arbitrage, is given by

$$d\xi_i(t) = -\xi_i(t)\left[\iota_t dt + \kappa_i(t)dZ^i_t\right], \tag{A5}$$

where ι_t is the real short rate, $\kappa_i(t)$ is a vector of market prices, perceived by group-i investors, of the three sources of risk $Z^i_t = [\widehat{Z}^i_\pi(t), Z_D(t), S_t]^\mathsf{T}$, and $\xi_i(0) = 1$.

Using the martingale technique of Cox and Huang (1989) and Karatzas, Lehoczky, and Shreve (1987), each group's dynamic optimization problem can be rewritten as a static one at time zero:

$$\max_{c_i} E^i_0 \left[\int_0^\infty e^{-\rho t} u\left(c^i_t\right)dt\right], \tag{A6}$$

subject to

$$E^i_0 \left[\int_0^\infty \xi_i(t)c^i_t dt\right] \le \alpha_i E^i_0 \left[\int_0^\infty \xi_i(t)D(t)dt\right], \tag{A7}$$

where the budget constraint is determined by the initial fraction of group-i's endowment α_i. The necessary and sufficient conditions for the optimal consumption stream are

$$e^{-\rho t}\frac{1}{c^i_t} = \varsigma_i \xi_i(t), \quad i = 1, 2, \ldots, N, \tag{A8}$$

where $\varsigma_i > 0$ is the Lagrange multiplier associated with the initial budget constraint. By combining Equation (A8) and the binding budget constraint (Equation (A7) with equality), we obtain the following condition for ς_i:

$$\rho \varsigma_i \alpha_i E^i_0 \left[\int_0^\infty \xi_i(t)D(t)dt\right] = 1. \tag{A9}$$

As is well known from Merton (1971), an investor with a logarithmic preference always consumes his or her wealth at a constant rate equal to his or her time preference parameter: $c^i_t = \rho W^i_t$. This consumption policy and Equation (A8) imply that

$$\eta(t) = \frac{c^B_t}{c^A_t} = \frac{\varsigma_A \xi_A(t)}{\varsigma_B \xi_B(t)}. \tag{A10}$$

By applying Ito's lemma to the above expression, we obtain

$$\frac{d\eta(t)}{\eta(t)} = \left(\hat{\theta}^B_t - \hat{\theta}^A_t\right)\frac{\lambda_\pi}{\sigma_\pi}d\widehat{Z}^A_\pi, \tag{A11}$$

with the initial condition $\eta(0) = \varsigma_A/\varsigma_B$. Rewriting Equation (A11) in the econometrician's measure, we obtain Equation (16).

Equation (A10) and the market clearing condition $c_t^A + c_t^B = D_t$ imply that

$$c_t^A = \frac{1}{1 + \eta_t} D_t, \tag{A12}$$

$$c_t^B = \frac{\eta_t}{1 + \eta_t} D_t. \tag{A13}$$

By substituting the consumption rules above into Equation (A8), we obtain an explicit expression for the state price density $\xi_i(t)$. Applying Ito's lemma to $\xi_i(t)$, we obtain the real short rate ι_t. Since $r_t = \iota_t + \pi_t$, we obtain Equation (15). To prove Equation (17), we need the following lemma.

Lemma A1. *If X_T is a random variable to be realized at time $T > t$ and $E^B[X_T] < \infty$, then $E_t^B[X_T] = E_t^A[\frac{\eta_T}{\eta_t} X_T]$.*

Proof. For any random variable X_T with $E^B[X_T] < \infty$, we can define $Y_T = \frac{W_T^B}{W_t^B} X_T$. Suppose there is a financial security that is a claim to the cash flow Y_T. Then group-B investors' valuation of this security is

$$E_t^B \left[e^{-\rho(T-t)} \frac{u'(c_T^B)}{u'(c_t^B)} Y_T \right] = e^{-\rho(T-t)} E_t^B \left[\frac{c_t^B}{c_T^B} Y_T \right] = e^{-\rho(T-t)} E_t^B \left[\frac{W_t^B}{W_T^B} Y_T \right] = e^{-\rho(T-t)} E_t^B[X_T],$$

where the second equality follows from these investors' consumption rule $c_t^B = \rho W_t^B$. Similarly, group-A investors' valuation of this security is

$$E_t^A \left[e^{-\rho(T-t)} \frac{u'(c_T^A)}{u'(c_t^A)} Y_T \right] = e^{-\rho(T-t)} E_t^A \left[\frac{c_t^A}{c_T^A} Y_T \right] = e^{-\rho(T-t)} E_t^A \left[\frac{W_t^A}{W_T^A} Y_T \right] = e^{-\rho(T-t)} E_t^A \left[\frac{\eta_T}{\eta_t} X_T \right].$$

Since group-A and group-B investors should have the same security valuation in equilibrium, we must have

$$E_t^B[X_T] = E_t^A \left[\frac{\eta_T}{\eta_t} X_T \right]. \qquad \blacksquare$$

This lemma shows that the wealth ratio between the two groups can also act as the Radon–Nikodym derivative between the two groups' probability measures. With this lemma, we can now prove Equation (17). The time-t nominal price of an asset, which provides a single nominal payoff X_T at time T, is given by

$$P_X(t) = p_t E_t^A \left[e^{-\rho(T-t)} \frac{c_t^A}{c_T^A} \frac{X_T}{p_T} \right].$$

Substituting the consumption share (A12) into the above equation, after some algebra, we obtain

$$P_X(t) = \frac{1}{1 + \eta_t} p_t E_t^A \left[e^{-\rho(T-t)} \frac{Dt}{D_T} \frac{X_T}{p_T} \right] + \frac{\eta_t}{1 + \eta_t} p_t E_t^A \left[\frac{\eta_T}{\eta_t} e^{-\rho(T-t)} \frac{Dt}{D_T} \frac{X_T}{p_T} \right]$$

$$= \frac{1}{1 + \eta_t} p_t E_t^A \left[e^{-\rho(T-t)} \frac{Dt}{D_T} \frac{X_T}{p_T} \right] + \frac{\eta_t}{1 + \eta_t} p_t E_t^B \left[e^{-\rho(T-t)} \frac{Dt}{D_T} \frac{X_T}{p_T} \right].$$

Note that $\frac{1}{1+\eta_t}$ is group-A investors' wealth share and $\frac{\eta_t}{1+\eta_t}$ is group-B investors' wealth share. Moreover, in a homogeneous economy with group-i ($i \in \{A, B\}$) investors only, the nominal price of this security is $p_t E_t^i [e^{-\rho(T-t)} \frac{Dt}{D_T} \frac{X_T}{p_T}]$. Hence, we obtain Equation (17).

Proof of Proposition 2

To replicate the price dynamics in the heterogeneous-investor economy, the representative investor's stochastic discount factor should be the same as that of investors in each group (say, group A) after adjusting for the difference in the probability measures. That is, the representative investor's marginal utility should have the following property in any future state:

$$u'(c_t^A) = \eta_t^N u'(c_t^N), \tag{A14}$$

where $u'(c_t^A)$ is the group-A investors' marginal utility, $u'(c_t^N)$ is the representative agent's marginal utility, and η_t^N is the change of measure from the representative agent's measure to group-A investors' measure. Similar to Equation (A11), we now obtain

$$\frac{d\eta_t^N}{\eta_t^N} = (\hat{\theta}_t^N - \hat{\theta}_t^A)\frac{\lambda_\pi}{\sigma_\pi}d\widehat{Z}_\pi^A. \tag{A15}$$

Note that investors with a logarithmic preference always consume their wealth at a constant rate: $c^A = \rho W_t^A$ and $c^N = \rho(W_t^A + W_t^B)$. By substituting these into Equation (A14) and after some algebra, we obtain

$$\eta_t^N = \eta_t + 1.$$

This further implies

$$\frac{d\eta_t^N}{\eta_t^N} = \frac{\eta_t}{\eta_t^N}\frac{d\eta_t}{\eta_t} = \frac{\eta_t}{1 + \eta_t}\frac{d\eta_t}{\eta_t}.$$

By substituting Equations (A11) and (A15) into the above expression and after some algebra, we obtain Equation (18).

Proof of Proposition 3

The price of the bond in a homogeneous economy has the following function form:

$$B_t^i = B^H\left(\tau, \pi_t, \hat{\theta}_t^i\right). \tag{A16}$$

In the homogeneous economy with group-i investors only, the real bond return $\frac{dB^H/p_t}{B^H/p_t}$ has to satisfy the following relationship with the real state price density M_t^H:

$$E_t^i\left(\frac{dB^H/p_t}{B^H/p_t}\right) + E_t^i\left(\frac{dM_t^H}{M_t^H}\right) + E_t^i\left(\frac{dB^H/p_t}{B^H/p_t}\frac{dM_t^H}{M_t^H}\right) = 0. \tag{A17}$$

From (2) and (A16), it is easy to see that

$$E_t^i\left(\frac{dB^H/p_t}{B^H/p_t}\right) = E_t^i\left(\frac{dB^H}{B^H}\right) - \pi_t dt. \tag{A18}$$

The real state price density in the homogeneous economy is

$$M_t^H = e^{-\rho t}\frac{1}{D_t}.$$

Applying Ito's lemma to the above expression, we obtain

$$\frac{dM_t^H}{M_t^H} = \left(-\rho - \mu_D + \sigma_D^2\right)dt - \sigma_D dZ_D.$$

We then have

$$E_t^i \left(\frac{dM_t^H}{M_t^H} \right) = \left(-\rho - \mu_D + \sigma_D^2 \right) dt, \tag{A19}$$

$$E_t^i \left(\frac{dB^H / p_t}{B^H / p_t} \frac{dM_t^H}{M_t^H} \right) = 0. \tag{A20}$$

Applying Ito's lemma to Equation (A16), we obtain

$$E_t^i \left(\frac{dB^H}{B^H} \right) = \left\{ -\frac{B_\tau^H}{B^H} - \lambda_\pi \left(\pi_t - \hat{\theta}_t^i \right) \frac{B_\pi^H}{B^H} - \lambda_\theta \left(\hat{\theta}_t^i - \bar{\theta} \right) \frac{B_\theta^H}{B^H} + \frac{1}{2} \sigma_\pi^2 \frac{B_{\pi\pi}^H}{B^H} \right.$$

$$\left. + \frac{1}{2} \left(k_i^2 \sigma_\theta^2 - 2\lambda_\theta \bar{v}_i \right) \frac{B_{\theta\theta}^H}{B^H} + \lambda_\pi \bar{v}_i \frac{B_{\pi\theta}^H}{B^H} \right\} dt. \tag{A21}$$

Substituting Equations (A18)–(A21) into Equation (A17), we obtain

$$0 = -\frac{B_\tau^H}{B^H} - \lambda_\pi \left(\pi_t - \hat{\theta}_t^i \right) \frac{B_\pi^H}{B^H} - \lambda_\theta \left(\hat{\theta}_t^i - \bar{\theta} \right) \frac{B_\theta^H}{B^H} + \frac{1}{2} \sigma_\pi^2 \frac{B_{\pi\pi}^H}{B^H}$$

$$\frac{1}{2} \left(k_i^2 \sigma_\theta^2 - 2\lambda_\theta \bar{v}_i \right) \frac{B_{\theta\theta}^H}{B^H} + \lambda_\pi \bar{v}_i \frac{B_{\pi\theta}^H}{B^H} - \pi_t - \rho - \mu_D + \sigma_D^2. \tag{A22}$$

We conjecture the following solution:

$$B^H \left(\tau, \pi_t, \hat{\theta}_t^i \right) = e^{-a_\pi(\tau)\pi_t - a_\theta(\tau)\hat{\theta}_t^i - b(\tau)}.$$

By substituting the conjectured solution into the differential equation in (A22) and collecting common terms, we obtain the following algebra equation:

$$0 = [a_\pi'(\tau) + \lambda_\pi a_\pi(\tau) - 1]\pi_t + [a_\theta'(\tau) - \lambda_\pi a_\pi(\tau) + \lambda_\theta a_\theta(\tau)]\hat{\theta}_t^i$$

$$+ \left[b'(\tau) - \lambda_\theta \bar{\theta} a_\theta(\tau) + \frac{1}{2} \sigma_\pi^2 a_\pi(\tau)^2 + \frac{1}{2} \left(k_i^2 \sigma_\theta^2 - 2\lambda_\theta \bar{v}_i \right) a_\theta(\tau)^2 \right.$$

$$+ \left. \lambda_\pi \bar{v}_i a_\pi(\tau) a_\theta(\tau) - \rho - \mu_D + \sigma_D^2 \right].$$

Since this equation has to hold for any values of π_t and $\hat{\theta}_t^i$, their coefficients must be 0. Thus, $a_\pi(\tau)$, $a_\theta(\tau)$, and $b(\tau)$ satisfy the following differential equations:

$$a_\pi'(\tau) + \lambda_\pi a_\pi(\tau) - 1 = 0,$$
$$a_\theta'(\tau) - \lambda_\pi a_\pi(\tau) + \lambda_\theta a_\theta(\tau) = 0,$$
$$b'(\tau) - \lambda_\theta \bar{\theta} a_\theta(\tau) + \frac{1}{2} \sigma_\pi^2 a_\pi^2(\tau) + \frac{1}{2} \left(k_i^2 \sigma_\theta^2 - 2\lambda_\theta \bar{v}_i \right) a_\theta^2(\tau)$$
$$+ \lambda_\pi \bar{v}_i a_\pi(\tau) a_\theta(\tau) - \rho - \mu_D + \sigma_D^2 = 0,$$

subject to the boundary conditions

$$a_\pi(0) = a_\theta(0) = b(0) = 0.$$

Solving these equations provides the bond price formula given in Proposition 3.

Proof of Proposition 4

Define $g_t \equiv \hat{\theta}_t^A - \hat{\theta}_t^B$. Since bond yield $Y_t(\tau) = -\frac{1}{\tau} \log(B_t)$, its volatility is

$$Vol[dY(\tau)_t] = \frac{1}{\tau} Vol(dB_t/B_t).$$

Applying Ito's lemma to Equation (22) in the econometrician's probability measure provides the following diffusion terms of $\frac{dB_t}{B_t}$:

$$-\left[a_\pi(\tau)\sigma_\pi + a_\theta(\tau)\lambda_\pi\sigma_\pi^{-1}\bar{v} - \frac{\lambda_\pi}{\sigma_\pi}\frac{\eta_t}{(1+\eta_t)^2}g_t \frac{e^{-a_\theta(\tau)g_t/2} - e^{a_\theta(\tau)g_t/2}}{\eta_t e^{-a_\theta(\tau)g_t/2} + e^{a_\theta(\tau)g_t/2}}\right]d\hat{Z}_\pi^R$$

$$-a_\theta(\tau)\frac{\phi\sigma_\theta}{\sqrt{1-\phi^2}}\frac{\eta_t e^{-a_\theta(\tau)g_t/2} - e^{a_\theta(\tau)g_t/2}}{\eta_t e^{-a_\theta(\tau)g_t/2} + e^{a_\theta(\tau)g_t/2}}dZ_S(t).$$

Since the diffusion term in each row is independent of each other, we obtain

$$\left(\frac{dB_t}{B_t}\right)^2 = \left[a_\pi(\tau)\sigma_\pi + a_\theta(\tau)\lambda_\theta\sigma_\pi^{-1}\bar{v} + \frac{\lambda_\pi}{\sigma_\pi}\frac{\eta_t}{(\eta_t+1)^2}K_1(g_t)\right]^2 dt + a_\theta^2(\tau)\frac{\phi^2\sigma_\theta^2}{(1-\phi^2)}K_2(g_t)dt,$$

where

$$K_1(g_t) = -g_t \frac{e^{-a_\theta(\tau)g_t/2} - e^{a_\theta(\tau)g_t/2}}{\eta_t e^{-a_\theta(\tau)g_t/2} + e^{a_\theta(\tau)g_t/2}},$$

and

$$K_2(g_t) = \left[\frac{\eta_t e^{-a_\theta(\tau)g_t/2} - e^{a_\theta(\tau)g_t/2}}{\eta_t e^{-a_\theta(\tau)g_t/2} + e^{a_\theta(\tau)g_t/2}}\right]^2.$$

Direct derivations of K_1 and K_2 provide that both of them increase as $|g_t|$ increases. Thus, the conditional variance of the bond return increases in the belief dispersion.

Proof of Proposition 5

As noted in the proof of Theorem 1, group-*B* investor's consumption is $c_t^B = \frac{\eta_t}{1+\eta_t}D_t$, and hence his or her marginal utility is $e^{-\rho t}\frac{1+\eta_t}{\eta_t D_t}$. By applying Ito's lemma to it and substituting in Equations (1) and (16), we obtain Equation (24).

Proof of Proposition 6

From Merton (1971), we know that group-*i* investors' position in security π is

$$n_\pi^i(t) = W_t^i \frac{\mu_\pi^i(t) - r_t}{\sigma_\pi^2} = W_t\omega_t^i \frac{\mu_\pi(t) - \lambda_\pi\pi_t + \lambda_\pi\hat{\theta}_t^i - r_t}{\sigma_\pi^2}.$$

Imposing the market-clearing condition $n_\pi^A(t) + n_\pi^B(t) = 0$, we obtain $\mu_\pi(t)$. Denote $W_t \equiv (W_t^A + W_t^B)$. Substituting $\mu_\pi(t)$ into the above expression, we obtain

$$n_\pi^A(t) = \frac{\lambda_\pi}{\sigma_\pi^2} W_t \frac{\eta_t}{(\eta_t+1)^2}g_t.$$

Note that group-*A* investors' position in security π is proportional to g_t. This implies that as the belief dispersion $|g_t|$ widens, group-*A* investors take a larger position in security π. Note that there is no disagreement on the expected returns of securities S and C. Hence, both groups have zero positions in those securities. Thus, we need to consider only trading volume in security π.

The absolute value of the change in group-*A* investors' position determines trading volume. In our model, the change in investors' position follows a diffusion process. It is well known that diffusion processes have infinite variation over a given time interval. However, since actual trading occurs in discrete time, it is reasonable to analyze trading volume through the change in investors' position across a finite time interval. Since the absolute value of a realized position change across a finite but small interval is finite and on average increases with the volatility of the position change, this motivates us to use the volatility as a measure of trading volume.

We now examine the change in group-*A* investors' position in security π, $dn_\pi(t)$, whose diffusion terms are

$$\frac{\lambda_\pi}{\sigma_\pi^2} \left[\frac{\eta_t}{(\eta_t + 1)^2} g_t dW_t - W_t g_t \frac{\eta_t - 1}{(\eta_t + 1)^3} d\eta_t + W_t \frac{\eta_t}{(\eta_t + 1)^2} dg_t \right].$$

By deriving the diffusion processes of dW_t, $d\eta_t$, and dg_t, and substituting them into the equation above, we can derive an expression of the variance of the position change, which increases with g_t^2. Thus, trading volume increases with investors' belief dispersion.

References

Acemoglu, D., V. Chernozhukov, and M. Yildiz. 2007. Learning and Disagreement in an Uncertain World. Working Paper, MIT.

Balduzzi, P., D. Das, and S. Foresi. 1998. The Central Tendency: A Second Factor in Bond Yields. *Review of Economics and Statistics* 80:62–72.

Basak, S. 2000. A Model of Dynamic Equilibrium Asset Pricing with Heterogeneous Beliefs and Extraneous Risk. *Journal of Economic Dynamics and Control* 24:63–95.

Bhamra, H., and R. Uppal. 2009. The Effect of Introducing a Non-Redundant Derivative on the Volatility of Stock-Market Returns When Agents Differ in Risk Aversion. *Review of Financial Studies* 22:2303–30.

Buraschi, A., and A. Jiltsov. 2006. Model Uncertainty and Option Markets with Heterogeneous Beliefs. *Journal of Finance* 61:2841–97.

Buraschi, A., and A. Jiltsov. 2007. Term Structure of Interest Rates: Implications of Habit Persistence. *Journal of Finance* 62:3009–63.

Campbell, J., A. Lo, and C. MacKinlay. 1997. *The Econometrics of Financial Markets*. Princeton, NJ: Princeton University Press.

Campbell, J., and R. Shiller. 1991. Yield Spreads and Interest Rate Movements: A Bird's Eye View. *Review of Economic Studies* 58:495–514.

Chan, L., and L. Kogan. 2002. Catching Up with the Joneses: Heterogeneous Preferences and the Dynamics of Asset Prices. *Journal of Political Economy* 110:1255–85.

Cochrane, J., and M. Piazzesi. 2004. Reply to Dai, Singleton, and Yang. Working Paper, University of Chicago.

Cochrane, J., and M. Piazzesi. 2005. Bond Risk Premia. *American Economic Review* 95:138–60.

Cox, J., and C.-F. Huang. 1989. Optimal Consumption and Investment Policies When Asset Prices Follow a Diffusion Process. *Journal of Economic Theory* 49:33–83.

Dai, Q., and K. Singleton. 2002. Expectation Puzzles, Time-Varying Risk Premia, and Affine Models of the Term Structure. *Journal of Financial Economics* 63:415–41.

Dai, Q., and K. Singleton. 2003. Term Structure Dynamics in Theory and Reality. *Review of Financial Studies* 16:631–78.

Dai, Q., K. Singleton, and W. Yang. 2004. Predictability of Bond Risk Premia and Affine Term Structure Models. Working Paper, Stanford University.

David, A. 2008. Heterogeneous Beliefs, Speculation, and the Equity Premium. *Journal of Finance* 63:41–83.

Detemple, J., and S. Murthy. 1994. Intertemporal Asset Pricing with Heterogeneous Beliefs. *Journal of Economic Theory* 62:294–320.

Duffie, D., and R. Kan. 1996. A Yield-Factor Model of Interest Rates. *Mathematical Finance* 4:379–406.

Dumas, B. 1989. Two-Person Dynamic Equilibrium in the Capital Market. *Review of Financial Studies* 2:157–88.

Dumas, B., A. Kurshev, and R. Uppal. 2009. Equilibrium Portfolio Strategies in the Presence of Sentiment Risk and Excess Volatility. *Journal of Finance* 64:579–629.

Fama, E., and R. Bliss. 1987. The Information in Long-Maturity Forward Rates. *American Economic Review* 77:680–92.

Gabaix, X. 2008. Variable Rare Disasters: An Exactly Solved Framework for Ten Puzzles in Macro-Finance. Working Paper, New York University.

Gallmeyer, M., and B. Hollifield. 2008. An Examination of Heterogeneous Beliefs with a Short Sale Constraint. *Review of Finance* 12:323–64.

Gurkaynak, R., B. Sack, and E. Swanson. 2005. The Excess Sensitivity of Long-Term Interest Rates: Evidence and Implications for Macroeconomic Models. *American Economic Review* 95:425–36.

Harris, M., and A. Raviv. 1993. Differences of Opinion Make a Horse Race. *Review of Financial Studies* 6:473–506.

Jouini, E., and C. Napp. 2007. Consensus Consumer and Intertemporal Asset Pricing with Heterogeneous Beliefs. *Review of Economic Studies* 74:1149–74.

Karatzas, I., J. Lehoczky, and S. Shreve. 1987. Optimal Portfolio and Consumption Decisions for a "Small Investor" on a Finite Horizon. *SIAM Journal on Control and Optimization* 25:1557–86.

Kogan, L., S. Ross, J. Wang, and M. Westerfield. 2006. The Price Impact and Survival of Irrational Traders. *Journal of Finance* 61:195–229.

Kurz, M. 1994. On the Structure and Diversity of Rational Beliefs. *Economic Theory* 4:877–900.

Li, T. 2007. Heterogeneous Beliefs, Asset Prices, and Volatility in a Pure Exchange Economy. *Journal of Economic Dynamics and Control* 31:1697–727.

Liptser, R., and A. Shiryaev. 1977. *Statistics of Random Processes*, 2 vols. New York: Springer.

Lucas, R. 1978. Asset Prices in an Exchange Economy. *Econometrica* 46:1429–45.

Mankiw, G., R. Reis, and J. Wolfers. 2004. Disagreement about Inflation Expectations. *NBER Macroeconomics Annual 2003*. Cambridge, MA: MIT Press.

Merton, R. 1971. Optimum Consumption and Portfolio Rules in a Continuous-Time Model. *Journal of Economic Theory* 3:373–413.

Morris, S. 1996. Speculative Investor Behavior and Learning. *Quarterly Journal of Economics* 111:1111–33.

Piazzesi, M., and M. Schneider. 2006. Equilibrium Yield Curve. *NBER Macroeconomics Annual* 389–442.

Scheinkman, J., and W. Xiong. 2003. Overconfidence and Speculative Bubbles. *Journal of Political Economy* 111:1183–219.

Shefrin, H. 2008. *A Behavioral Approach to Asset Pricing*. Burlington, MA: Academic Press.

Shiller, R. 1979. The Volatility of Long-Term Interest Rates and Expectations Models of the Term Structure. *Journal of Political Economy* 87:1190–219.

Swanson, E. 2006. Have Increases in Federal Reserve Transparency Improved Private Sector Interest Rate Forecasts? *Journal of Money, Credit and Banking* 38:791–819.

Vayanos, D., and J.-L. Vila. 2007. A Preferred-Habitat Model of the Term Structure of Interest Rates. Working Paper, London School of Economics.

Wachter, J. 2006. A Consumption-Based Model of the Term Structure of Interest Rates. *Journal of Financial Economics* 79:365–99.

Wang, J. 1996. The Term Structure of Interest Rates in a Pure Exchange Economy with Heterogeneous Investors. *Journal of Financial Economics* 41:75–110.

Yan, H. 2008. Natural Selection in Financial Markets: Does It Work? *Management Science* 54:1935–50.

Zapatero, F. 1998. Effects of Financial Innovation on Market Volatility When Beliefs Are Heterogeneous. *Journal of Economic Dynamics and Control* 22:597–626.