Abstract

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 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 expectation 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.

We are grateful to Nick Barberis, Markus Brunnermeier, Bernard Dumas, Nicolae Garleanu, Jon Ingersoll, Arvind Krishnamurthy, Owen Lamont, Debbie Lucas, Lin Peng, Monika Piazzesi, Chris Sims, Hyun Shin, Jeremy Stein, Stijn Van Nieuwerburgh, Neng Wang, Moto Yogo, and seminar participants at Bank of Italy, Federal Reserve Bank of New York, Harvard University, NBER Summer Institute, New York University, Northwestern University, Princeton University, University of British Columbia, University of Chicago, University of Illinois-Chicago, Wharton School, and Yale University for their helpful discussions and comments.

1Princeton University and NBER. Email: wxiong@princeton.edu.

2Yale University. Email: hongjun.yan@yale.edu.
1 Introduction

Following the seminal work of Cox, Ingersoll and Ross (1985), most academic studies in the economics and finance literature use representative-agent models to analyze yield curve dynamics. While this approach leads to tractable parametric models of yield curve dynamics, representative-agent models ignore important interactions among heterogeneous agents. In this paper, we analyze an equilibrium model of bond markets in which heterogeneous expectations cause agents to trade with each other. We show that the endogenous wealth fluctuations caused by agents’ trading can help resolve several challenges encountered by standard representative-agent models, including the “excessive volatility” of bond yields and the failure of the expectations hypothesis. Moreover, our model also sheds light on the recent finding of Cochrane and Piazzesi (2005) that a single tent-shaped linear combination of forward rates predicts excess bond returns.

We adopt the standard equilibrium framework of Cox, Ingersoll and Ross (1985) with log-utility agents and a constant-return-to-scale risky investment technology. Unlike their model, we allow agents to hold heterogeneous expectations of future economic conditions.\footnote{There is ample evidence supporting the existence of heterogeneous expectations among agents. 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 (2005) 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 next-quarter real US GDP growth rate fluctuates between 1.5% and 5%, and the 90th and 10th percentile forecasts of four-quarter-behind 3-month Treasury bill rate fluctuates between 0.8% and 2.2%.} Specifically, we assume that there are two groups of agents using different learning models to infer the value of an unobservable variable that determines future equilibrium short rates. Consequently, the two groups of agents hold heterogeneous expectations about future interest rates. A group is said to be more optimistic (pessimistic) about future short rates if its expectation of future short rates is higher (lower) than the other group’s.

Note that when agents adopt different learning models, not only will their beliefs differ from each other, but their average belief can also deviate from the belief of an outside objective observer of the economy, whom we call an econometrician. In order to isolate the effects caused by the two groups’ belief dispersion from those caused by their erroneous average
belief, we adopt a novel specification in which the two groups’ beliefs are divergent but their average is always identical to the econometrician’s benchmark belief. Heterogeneous beliefs motivate agents to take speculative positions against each other in bond markets, and market clearing conditions determine bond prices. We 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 agent is present.

Standard representative-agent models have difficulty generating the large bond yield volatility and highly variable risk premia observed in actual data because the aggregate consumption is rather smooth. Our model shows that the relative wealth fluctuation caused by agents’ speculative positions amplifies bond yield volatility and contributes to the time variation in bond premia.

The intuition is as follows. Agents who are optimistic about future interest rates would bet on rates rising against those pessimistic agents. In equilibrium, bond prices aggregate agents’ heterogeneous beliefs and, in particular, reflect their wealth-weighted average belief. When agents’ wealth-weighted average belief about the future short rates is higher than the econometrician’s belief, they would 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 agents’ 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 groups’ wealth distribution is endogenously determined by their trading. When a positive shock hits the market, it favors optimistic agents and causes wealth to flow from pessimistic agents to optimistic agents, 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 agents’ 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-
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 agent 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 in 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 agents discount long term bonds more heavily, which leads to higher long term bond yields and so larger yield spreads; on the other, 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, we show in our simulations that a reasonable amount of belief dispersion is able to generate regression results similar to those of Campbell and Shiller.

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. As elaborated later in the paper, this tent-shaped factor tracks agents’ wealth-weighted belief: the higher the weighted average belief about the future short rates, the bigger value of the tent-shaped factor. Moreover, a higher belief about the 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 simulation confirms that a reasonable amount of belief dispersion is able to generate bond return predictability results comparable to those of Cochrane and Piazzesi.

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. More recently, Basak (2000), Dumas, Kurshev and
Uppal (2005), Jouini and Napp (2005), Buraschi and Jiltsov (2006), David (2007), and Li (2007) 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. 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 under-estimation of risk and by erroneous average beliefs. Second, and more importantly, 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), and Chan and Kogan (2002). 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. Finally, Section 4 concludes. We provide all the technical proofs in the Appendix.

2 The Model

We adopt the equilibrium framework of Cox, Ingersoll and Ross (1985) with log-utility agents and a constant-return-to-scale risky investment technology. Unlike their model, ours assumes that agents cannot directly observe a random variable that determines future returns of the risky technology, and that agents can only estimate its value. There are two groups of agents holding heterogeneous expectations regarding this variable. Because of the belief dispersion, agents speculate in capital markets. We study a competitive equilibrium in which each agent optimizes consumption and investment decisions based on his own expectation. Market clearing conditions determine the equilibrium short rate and asset prices.
2.1 The economy

We consider an economy with only one constant-return-to-scale technology. The return of the technology follows a diffusion process:

\[
\frac{dI_t}{I_t} = f_t dt + \sigma_I dZ_I(t), \tag{1}
\]

where \( f_t \) is the expected instantaneous return, \( \sigma_I \) a volatility parameter, and \( Z_I(t) \) a standard Brownian motion process.

The expected instantaneous return from the risky technology, \( f_t \), follows another linear diffusion process:

\[
df_t = -\lambda_f (f_t - l_t) dt + \sigma_f dZ_f(t), \tag{2}
\]

where \( \lambda_f \) is a constant governing the mean-reverting speed of \( f_t \), \( l_t \) a moving long-run mean of the risky technology’s expected return, \( \sigma_f \) a volatility parameter, and \( Z_f(t) \) a standard Brownian motion process independent of \( Z_I(t) \). The long-run mean \( l_t \) is unobservable and follows an Ornstein-Uhlenbeck process:

\[
dl_t = -\lambda_l (l_t - \bar{l}) dt + \sigma_l dZ_l(t), \tag{3}
\]

where \( \lambda_l \) is a parameter governing the mean-reverting speed of \( l_t \), \( \bar{l} \) the long-run mean of \( l_t \), \( \sigma_l \) a volatility parameter, and \( Z_l(t) \) a standard Brownian motion process independent of \( Z_I(t) \) and \( Z_f(t) \).

As we show later, both variables \( f_t \) and \( l_t \) affect the dynamics of short-term and long-term interest rates. Since the risky technology represents an alternative investment to investing in the short term bond, the technology’s expected instantaneous return \( f_t \), after adjusted for risk, determines the short rate. \( l_t \) is the level to which \( f_t \) mean-reverts and so it affects the future short rates. In reality, investors often face great uncertainty regarding the future short rates and need to form expectations in order to trade long-term bonds. Introducing the variable \( l_t \) and making it unobservable provides a convenient way of modelling agents’ expectations. The central part of our analysis is to show that agents’ heterogeneous expectations can generate significant effects on bond markets.\(^2\)

\(^2\)For simplicity, this paper focuses on agents’ disagreement about the real side of the economy. In an
2.2 Heterogeneous expectations

The economics and finance literature has widely adopted the Bayesian inference framework to model agents’ 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) and Basak (2000), assumes that agents hold heterogeneous prior beliefs about unobservable economic variables. In these models, agents 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.

In another strand of the literature, e.g., Scheinkman and Xiong (2003), Dumas, Kurshev and Uppal (2005), Buraschi and Jiltsov (2006) and David (2007), heterogeneous beliefs arise from agents’ 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 agents to identify the correct model of the economy from alternative ones. More recently, Acemoglu, Chernozhukov, and Yildiz (2007) show that when agents are uncertain about a random variable and about the informativeness of a source of signal regarding the random variable, even an infinite sequence of signals from this same source does not lead agents’ heterogeneous prior beliefs about the random variable to converge. This is because agents have to update beliefs about two sources of uncertainty using one sequence of signals. Finally, behavioral biases such as overconfidence could also prevent agents 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 long-run mean of risky technology returns \( l_t \). In particular, we assume that the signals are not informative about \( l_t \), but the heterogeneous prior knowledge leads the two groups to react differently to the earlier version, we extended our model to incorporate the monetary side of the economy by specifying two additional random processes, with a similar structure to that of \((f_t, l_t)\). More specifically, one observable process determines the short-term inflation rate and the other unobservable process determines the long-run mean to which the short-term inflation rate mean-reverts. By allowing agents to hold heterogeneous beliefs about the value of this unobservable long-run mean of inflation rate, we can show that agents’ disagreement about the monetary side of the economy has an effect on bond markets similar to their disagreement about the real side. Since this analysis does not present additional conceptual insight, we do not report it in this paper.
signal flow and therefore to possess heterogeneous expectations. This approach is tractable and generates a stationary process for the difference in agents’ beliefs.

2.2.1 Benchmark belief

We will evaluate the effects of agents’ heterogeneous beliefs on asset price dynamics from the viewpoint of an outside observer, an econometrician. Hence, we first derive the belief of the econometrician. Since the econometrician understands that the signals are not informative, his information set at time $t$ is $\{f_\tau\}_{\tau=0}^t$. We assume that the econometrician’s prior belief about $l_0$ has a Gaussian distribution. Since their information flow also follows Gaussian processes, their posterior beliefs about $l_t$ 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 $\bar{\gamma}$, which is the positive root to the following quadratic equation of $\gamma$:

$$
\frac{\lambda_f^2}{\sigma_f^2} \gamma^2 + 2\lambda_f \gamma - \sigma_l^2 = 0.
$$

We denote the econometrician’s posterior distribution about $l_t$ at time $t$ by

$$
l_t|\{f_\tau, S(\tau)\}_{\tau=0}^t \sim N(\hat{l}_R^t, \bar{\gamma}),
$$

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

$$
d\hat{l}_R^t = -\lambda_l(\hat{l}_R^t - \bar{l})dt + \frac{\lambda_f}{\sigma_f} \gamma d\hat{Z}_R^f(t),
$$

where

$$
d\hat{Z}_R^f = \frac{1}{\sigma_f} \left[ df_t + \lambda_f(f_t - \hat{l}_R^t)dt \right]
$$

is the information shock in $df_t$. $\hat{Z}_R^f$ is a standard Brownian motion from the econometrician’s point of view.
2.2.2 Heterogeneous beliefs

We assume that all agents observe the following process:

\[ dS_t = dZ_S(t), \]

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, agents in the economy believe that \( S(t) \) is partially correlated with the fundamental shock to \( l_t \) and thus contains useful information.

There are two groups of agents, group A and group B. They have different interpretations of \( S(t) \) and so possess heterogeneous expectations. Before we formally introduce the beliefs of the two groups, it is useful to note that, in general, this heterogeneous economy could differ from a homogeneous economy through several channels. First, disagreement induces speculative trading between the two groups, therefore generating endogenous relative wealth fluctuation. 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, since each agent in the economy feels he has learned from the signal about \( l_t \), his posterior belief variance would be smaller than the econometrician’s and this leads to an underestimation of the risk in the economy. While the effects generated by the second and third channels are interesting in their own right, we are primarily interested in the first channel. In order to isolate the impact from the first channel, we adopt a specification that shuts down the other two channels.

To shut down the second channel, we assume that the two groups have exactly the opposite interpretations of \( S \). Agents in group A believe that the signal generating process is

\[ dS_t = \phi_S dZ_l(t) + \sqrt{1 - \phi_S^2} dZ_S(t), \]

where the parameter \( \phi_S \in [0, 1) \) measures the perceived correlation between the signal \( dS_t \) and the fundamental shock \( dZ_l(t) \). Symmetrically, agents in group B believe that the signal has a correlation of \(-\phi_S\) with the fundamental shock \( dZ_l(t)\):

\[ dS_t = -\phi_S dZ_l(t) + \sqrt{1 - \phi_S^2} dZ_S(t). \]

Since the two groups of agents have opposite perceptions about the correlation between \( dS_t \) and \( dZ_l(t) \), they would hold different beliefs and, as will be shown later, the average of
their beliefs coincides with the belief of the econometrician. This specification also includes the benchmark case with one rational agent as a special case of $\phi_S = 0$. Note also that it is difficult for these agents to detect the inconsistency between the perceived signal generating process and the actual process for two reasons. First, in both group-\(A\) and group-\(B\) agents’ minds, the perceived signal generating process has the same quadratic variation as the actual process. Second, agents cannot directly measure the correlation between \(dS_t\) and \(dZ_l(t)\) since \(dZ_l(t)\) is unobservable.

To turn off the third channel, we further assume that agents in both groups also perceive a higher volatility in the unobservable process of \(l_t\):

\[
dl_t = -\lambda_l(l_t - \bar{l})dt + k\sigma_l dZ_l(t),
\]

where

\[
k = \frac{1}{\sqrt{1 - \phi_S^2}} > 1.
\]

That is, agents exaggerate the volatility by a factor of \(k\). Note that because agents cannot observe \(l_t\), they cannot detect their exaggeration by quadratic variation. Moreover, we choose the magnitude of \(k\) so that, as we show next, it exactly offsets the underestimation of the uncertainty in \(l_t\).

We now derive group-\(A\) and group-\(B\) agents’ belief about \(l_t\). Agents’ information set at time \(t\) includes \(\{f_\tau, S(\tau)\}_{\tau=0}^t\). According to Theorem 12.7 of Liptser and Shiryaev (1977), agents’ posterior variance at the steady state is the positive root to the following quadratic equation of \(\gamma\):

\[
\frac{\lambda_l^2}{\sigma_l^2} \gamma^2 + 2\lambda_l \gamma - (1 - \phi_S^2) k^2 \sigma_l^2 = 0.
\]  

Since \((1 - \phi_S^2) k^2 = 1\), equations (4) and (7) imply that agents in both groups have the same posterior variance as the econometrician.

We denote group-\(i\) agents’ posterior distribution about \(l_t\) at time \(t\) by

\[
l_t|\{f_\tau, S(\tau)\}_{\tau=0}^t \sim N\left(\hat{l}_{it}, \hat{\gamma}\right), \quad i \in \{A, B\},
\]

where \(\hat{l}_{it}\) is the mean of group-\(i\) agents’ posterior distribution. We will refer to \(\hat{l}_{it}\) as their
belief hereafter. Theorem 12.7 of Liptser and Shiryaev (1977) implies

\[
d\hat{l}_i = -\lambda_i(\hat{l}_i - \bar{l})dt + \frac{\lambda_f}{\sigma_f}d\hat{Z}_i(t) + \begin{cases} \frac{\phi_S\sigma_l}{\sqrt{1-\phi^2_S}}dS_t & \text{if } i = A \\ -\frac{\phi_S\sigma_l}{\sqrt{1-\phi^2_S}}dS_t & \text{if } i = B \end{cases}
\] (8)

where

\[
d\hat{Z}_i = \frac{1}{\sigma_f}\left[df_t + \lambda_f(f_t - \hat{l}_i)dt\right]
\] (9)

is the information shock in $df_t$ to group-$i$ agents. $\hat{Z}_i$ is a standard Brownian motion from group-$i$ agents’ point of view.

As shown in (8), due to the difference in their prior knowledge about the correlation between $dS_t$ and $dZ_l(t)$, group-$A$ agents react positively to $dS_t$, while group-$B$ agents react negatively. Thus, the two groups hold heterogeneous beliefs about $l_t$ and the difference in their beliefs fluctuates with the signal flow. Furthermore, the two groups’ reactions to $dS_t$ have opposite signs but the same magnitude. As a result, if the average of the two groups’ prior beliefs about $l_0$ is the same as the econometrician’s prior belief, then in the future their average belief about $l_t$ would always keep track of the econometrician’s belief. We formally state these properties of agents’ beliefs in the following proposition and provide the proof in the Appendix.

**Proposition 1** The difference in the two groups’ beliefs has the following process:

\[
d\left(\hat{l}_i^A - \hat{l}_i^B\right) = -\left(\lambda_l + \frac{\lambda_f^2}{\sigma_f^2}\gamma\right)\left(\hat{l}_i^A - \hat{l}_i^B\right)dt + \frac{2\phi_S\sigma_l}{\sqrt{1-\phi^2_S}}dS_t,
\]

which fluctuates with $dS_t$ and mean-reverts to zero. Furthermore, if $\frac{1}{2}(\hat{l}_0^A + \hat{l}_0^B) = \hat{l}_R^0$, then the average of the two groups’ beliefs about $l_t$ always tracks the econometrician’s belief:

\[
\frac{1}{2}(\hat{l}_t^A + \hat{l}_t^B) = \hat{l}_t^R.
\]

This proposition results from the symmetric learning processes of the two groups. Unlike other models of heterogeneous beliefs in the literature, this model specification allows us to isolate belief-dispersion effects from other learning related effects, such as under-estimation of risk and erroneous average belief.
2.3 Capital markets equilibrium

The difference in agents’ beliefs causes speculative trading among themselves. Agents who are more optimistic about \( l_t \) would bet on interest rates rising against more pessimistic agents. Note that, from each group’s perspective, there are three sources of random shocks. For group-\( i \) agents, the shocks are \( dZ_{if} \), \( \dot{d}Z_{if} \), and \( dS_{it} \). Thus, the markets are complete if agents can trade a risk-free asset and three risky assets that span these three sources of random shocks. In reality, bond markets offer many securities, such as bonds with different maturities, for agents to construct their bets and to complete the markets. As a result, we analyze agents’ investment and consumption decisions, as well as their valuation of financial securities, in a complete-markets equilibrium.

We introduce a zero-net-supply risk-free asset and two zero-net-supply risky financial securities in the capital markets, in addition to the risky production technology.\(^3\) At time \( t \), the risk-free asset offers a short rate \( r_t \), which is endogenously determined in the equilibrium. The two risky financial securities offer the following return processes:

\[
\frac{dS_p}{S_p} = \mu_S(t)dt + dS_t, \quad (11)
\]

\[
\frac{df_p}{f_p} = \mu_f(t)dt + df_t, \quad (10)
\]

We refer to these securities as security \( f \) and security \( S \), respectively. Like futures contracts, these securities are continuously marked to the fluctuations of \( df_t \) and \( dS_t \), respectively. Since agents hold different views about the underlying innovation processes of these securities, they disagree about their expected returns. As a result, some agents want to take long positions, while others want to take short positions. Through trading, the contract terms \( \mu_f(t) \) and \( \mu_S(t) \) are continuously determined so that the aggregate demand for each of the securities is zero all the time. We could also view these financial securities as synthetic positions constructed by dynamically trading bonds. We choose to introduce these securities instead of specific bonds to simplify notations, and our specific choice of securities does not affect the equilibrium in complete markets.

\(^3\)We also allow agents to short-sell the risky technology. This can be implemented by offering a derivative contract on the return of the technology. The market clearing conditions, however, require that agents in aggregate hold a long position in the risky technology.
We assume that all agents have a logarithmic preference. Group-$i$ agents ($i \in \{A, B\}$) maximize, according to their beliefs, their lifetime utility from consumption by choosing their consumption $c^i_t$, and the fraction of their wealth invested in the risky technology and two financial securities ($x^i_I, x^i_f, x^i_S$):

$$\max_{\{c^i_t, x^i_I, x^i_f, x^i_S\}} E^i \int_0^\infty e^{-\beta t} u(c^i_t) dt,$$

where $E^i$ is the expectation operator under their probability measure, $\beta$ is their time-preference parameter, and

$$u(c^i_t) = \log(c^i_t)$$

is their utility function from consumption.

We solve each agent’s consumption and investment problem using the standard dynamic programming approach developed by Merton (1971). The results for logarithmic utility are well known: Each agent consumes wealth at a constant rate equal to his time preference parameter ($c^i_t = \beta W^i_t$) and invests in risky assets according to the assets’ instantaneous risk-return tradeoff.

In a competitive equilibrium, each agent chooses optimal consumption and investment strategies in accordance with his expectations and all markets clear. Market clearing conditions ensure: 1) the aggregate investment to the risk-free asset is zero; 2) the aggregate investment to each of the risky securities $f$ and $S$ is also zero; and 3) the aggregate investment to the risky technology is equal to the total wealth in the economy. We formally derive the equilibrium in Appendix A.2 and summarize three important properties of the equilibrium in the following theorem.

**Theorem 1** The equilibrium has the following properties:

1. The short rate is

   $$r_t = f_t - \sigma^2 f_t.$$

2. Define the wealth ratio of group-$A$ and group-$B$ agents by $\eta_t \equiv \frac{W^A_t}{W^B_t}$. Its logarithm follows a diffusion process:

   $$d \ln (\eta_t) = \frac{\lambda f}{\sigma_f} \left( \hat{i}^A_t - \hat{i}^B_t \right) d\hat{Z}^R_f (t).$$

12
3. The time $t$ price of an asset, which provides a single payoff $X_T$ at time $T$, is given by

$$P_t = \omega_t^A P_t^A + \omega_t^B P_t^B,$$

where $\omega_t^i$ is the wealth share of group-$i$ agents, and $P_t^i$ is the price of the asset in a hypothetical economy, in which only group-$i$ agents are present.

Theorem 1 shows that the equilibrium short rate is the expected instantaneous return of the risky technology adjusted for risk (equation (12)). This is because agents would demand a higher return from lending out capital when the expected return from the alternative option of investing in the risky technology is higher.

Theorem 1 also shows that when the two groups hold different beliefs, logarithm of their wealth ratio responds to the information shock $d\hat{Z}^R_f(t)$. This is a result of speculation among the two groups. For instance, if group-A agents hold a higher belief about $l_t$ ($\hat{l}^A_t > \hat{l}^B_t$), they would bet against group-B agents on future interest rate rising. Consequently, a positive information shock would favor group-A agents and cause wealth to flow from group B to group A. Also note that 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 neither group has a superior learning model.\footnote{See Kogan, et al. (2006) and Yan (2006) for more discussions on investors’ survival in the long run.}

Finally, Theorem 1 shows that the price of an asset is the wealth-weighted average of each group’s valuation of the asset in a corresponding homogeneous economy. This result allows us to represent asset prices in a heterogeneous economy using prices in homogeneous economies. Thus, asset pricing is remarkably simple even in a complex environment with heterogeneous agents. While this price representation depends on agents’ logarithmic preference and linear risky technology, it is independent of the specific information structure in our model. Several earlier models, e.g., Detemple and Murthy (1994) and Basak (2000), provide similar price representations under different information structures.

2.4 Bond price in a homogeneous economy

Theorem 1 allows us to express the price of a bond as the wealth-weighted average of each group’s bond valuation in a homogeneous economy. Thus, before analyzing the effects of
agents’ heterogeneous expectations on bond markets, we first derive bond prices in homogeneous economies in the following proposition 2, with a proof in Appendix A.3.

**Proposition 2** In a homogeneous economy with only group-i agents, the price of a zero-coupon bond with a face value of 1 and maturity $\tau$ is determined by

$$B^H(\tau, f_t, \tilde{l}_t) = e^{-a_f(\tau)f_t - a_l(\tau)\tilde{l}_t - b(\tau)},$$

where

$$a_f(\tau) = \frac{1}{\lambda_f} \left(1 - e^{-\lambda_f \tau}\right),$$

$$a_l(\tau) = \frac{1}{\lambda_l} \left(1 - e^{-\lambda_l \tau}\right) + \frac{1}{\lambda_f - \lambda_l} \left(e^{-\lambda_f \tau} - e^{-\lambda_l \tau}\right),$$

$$b(\tau) = \int_0^\tau \left[\lambda_l \tilde{l}_t a_l(s) - \frac{1}{2} \sigma_f^2 a_f^2(s) - \frac{1}{2} \left(\frac{1}{1 - \phi^2_{S,L}}\right) \sigma_f^2 - 2\lambda_l \gamma\right] a_l^2(s) - \lambda_f \gamma a_f(s) a_l(s) - \sigma_f^2 \right] ds.$$

Proposition 2 implies that the yield of a $\tau$-year bond in a homogeneous economy,

$$Y^H(\tau, f_t, \tilde{l}_t) = -\frac{1}{\tau} \log(B^H) = \frac{a_f(\tau)}{\tau} f_t + \frac{a_l(\tau)}{\tau} \tilde{l}_t + \frac{b(\tau)}{\tau},$$

is a linear function of two fundamental factors: $f_t$ and $\tilde{l}_t$. This specific form belongs to the general affine structure proposed by Duffie and Kan (1996).

The loading on factor $f_t$, $a_f(\tau)/\tau$, has a value of 1 when the bond maturity $\tau$ is zero and monotonically decreases to zero as the maturity increases, suggesting that short-term yields are more exposed to the fluctuations in $f_t$. The intuition of this pattern is as follows. $f_t$ is the expected instantaneous return from the risky technology, which serves as a close substitute for investing in short-term bonds. As a result, $f_t$ determines the short rate ($r_t = f_t - \sigma_f^2$ as in Theorem 1) and the fluctuation in $f_t$ has a greater impact on short-term yields. As bond maturity increases, the impact of $f_t$ becomes smaller.

Agents’ belief about $l_t$ determines their expectation of the future short rates, because $l_t$ is the level to which $f_t$ mean-reverts. The loading of the bond yield on $\tilde{l}_t$, $a_l(\tau)/\tau$, has a hump shape if $l_t$ mean-reverts ($\lambda_l > 0$).\footnote{In the case where mean reversion is not present ($\lambda_l = 0$), the factor loading $a_l(\tau)/\tau$ is a monotonically increasing function of bond maturity.} Since $l_t$ describes the long run productivity, as the
bond maturity increases from 0, the bond yield becomes more sensitive to the belief about \( l_t \), that is, as the bond maturity increases from 0 to an intermediate value, the loading \( a_l(\tau)/\tau \) increases. As the bond maturity increases further, the loading \( a_l(\tau)/\tau \) falls. This is caused by the mean reversion of \( l_t \), which causes any shock to \( l_t \) to eventually die out. This force causes the yields of very long-term bonds to have low exposure to agents’ belief about \( l_t \). This hump shape in the bond yield’s loading on \( l_t \) is important for understanding later results such as volatility amplification and bond return predictability.

2.5 Representative agent

As is well known, one can construct a representative agent to replicate the price dynamics in a complete-markets equilibrium with heterogeneous agents. Does this mean that we can simply focus on the representative agent’s belief process and ignore the heterogeneity between agents? To understand why the answer is no, we construct a representative agent for our model.\(^6\) If we restrict the representative agent to having the same logarithmic preference as the group-\( A \) and group-\( B \) agents, we obtain the same equilibrium as before by “twisting” the representative agent’s belief, as summarized in the following proposition with a proof in Appendix A.4.

**Proposition 3** To replicate the competitive equilibrium derived in Theorem 1, consider a representative agent who has the same logarithmic preference as agents in the heterogeneous economy. Then, at any point of time, the representative agent’s belief about \( l_t \), \( \hat{l}^N_t \), has to be the wealth-weighted average belief of group-\( A \) and group-\( B \) agents:

\[
\hat{l}^N_t = \omega^A_t \hat{l}^A_t + \omega^B_t \hat{l}^B_t.
\]

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 would change in response not only to the belief fluctuation of each group, but also to the relative wealth fluctuation. According to Theorem 1, heterogeneous beliefs cause the two groups to trade against each other and consequently

---

\(^6\)See Jouini and Napp (2005) for a recent analysis of the existence of a “consensus” belief for the representative agent in an exchange economy with agents holding heterogeneous beliefs.
cause their relative wealth to fluctuate with information shock to the market. Thus, although the two groups’ beliefs are symmetrically distributed around the econometrician’s belief, through the relative wealth fluctuation channel, their belief dispersion could still affect the representative agent’s belief and therefore equilibrium asset prices. We analyze these effects in the next section.

### 3 Effects of Heterogeneous Expectations

Combining Proposition 2 with Theorem 1, we can express the price of a $\tau$-year zero-coupon bond at time $t$ as

$$B_t = \omega_t^A B^H(\tau, f_t, \hat{l}_t^A) + \omega_t^B B^H(\tau, f_t, \hat{l}_t^B),$$

(17)

where $\omega_t^A$ and $\omega_t^B$ are the two groups’ wealth shares in the economy, and $B^H(\tau, f_t, \hat{l}_t)$, given in Proposition 2, is the bond price in a homogeneous economy in which only group-$i$ agents are present. The implied bond yield in this heterogeneous economy is

$$Y_t(\tau) = \frac{a f(\tau)}{\tau} f_t + \frac{b(\tau)}{\tau} - \frac{1}{\tau} \log \left[ \omega_t^A e^{-a_l(\tau)\hat{l}_t^A} + \omega_t^B e^{-a_l(\tau)\hat{l}_t^B} \right].$$

(18)

Note that $Y_t$ is not a linear function of agents’ beliefs $\hat{l}_t^A$ and $\hat{l}_t^B$. That is, bond yields in this heterogeneous economy have a non-affine structure. This structure derives from the market aggregation of agents’ heterogeneous valuations of the bond.

#### 3.1 Trading volume

Heterogeneous expectations cause agents to take speculative positions against each other in bond markets. These speculative positions can cause fluctuations in agents’ wealth shares upon the arrival of new information. Agents 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 agents’ wealth and therefore a larger trading volume in the bond markets. We use the volatility of one group’s position changes as a measure of trading volume. This measure corresponds to the conventional volume measure in a discrete-time set up. We summarize the effect of agents’ belief dispersion on trading volume in Proposition 4, and provide a formal derivation and further discussion on our volume measure in Appendix A.5.
Proposition 4 Trading volume (the fluctuation in agents’ speculative positions) increases with the belief dispersion between the two groups of investors.

There is now a growing literature analyzing trading volume caused by heterogeneous beliefs, e.g., Harris and Raviv (1993) and Scheinkman and Xiong (2003). While these models demonstrate that heterogeneous beliefs lead to trading, trading typically occurs when agents’ beliefs flip, that is, \( \hat{l}^A_t - \hat{l}^B_t \) changes its sign. Thus, trading volume of this type only increases with the frequency with which agents’ beliefs flip. Our model adds to this literature by showing that even without agents’ beliefs flipping, the wealth fluctuation caused by their speculative positions already leads to trading.

3.2 Volatility amplification

The wealth fluctuation caused by agents’ speculative positions not only leads to trading in bond markets, but also amplifies bond yield volatility. Loosely speaking, bond yields are determined by agents’ wealth-weighted average belief about future interest rates (equation (18)). Since agents who are more optimistic about future rates bet on these rates rising against more pessimistic agents, any positive news about future rates would cause wealth to flow from pessimistic agents to optimistic agents, making the optimistic belief carry 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 and so increases the bond yield volatility. We summarize this intuition in the following proposition, and provide a formal proof in Appendix A.6.

Proposition 5 Bond yield volatility increases with belief dispersion.

This volatility amplification mechanism can help explain the “excess volatility puzzle” for 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 asset pricing model with recursive utility preferences and exogenous consumption growth and inflation,
the model predicts less volatility for long yields relative to short yields. Relating to this literature, Proposition 5 shows that extending standard representative-agent models with heterogeneous expectations can help 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 agents’ belief dispersion and relative wealth also cause risk premia to vary over time. To examine the time variation in risk premia, we first derive the dynamics of the stochastic discount factor, with a proof in Appendix A.7.

**Proposition 6** From the viewpoint of the econometrician, the stochastic discount factor has the following process:

\[
\frac{dM_t}{M_t} = -(f_t - \sigma_f^2)dt - \sigma_f dZ_I - \frac{\lambda_f}{\sigma_f} \left( i_t^R - \sum_{i=A}^{B} \omega_i \hat{l}_t^R \right) d\hat{Z}_R^f, \tag{19}
\]

where \( \hat{l}_t^R \) is the econometrician’s belief about \( l_t \), and \( d\hat{Z}_R^f \) the information shock defined in equation (6).

Proposition 6 shows that from the viewpoint of the econometrician the market price of risk (risk premium per unit of risk) for the aggregate production shock \( dZ_I \) is \( \sigma_I \), while the market price of risk for the information shock \( d\hat{Z}_R^f \) is proportional to \( i_t^R - \sum_{i=A}^{B} \omega_i \hat{l}_t^R \), the difference between the econometrician’s belief about \( l_t \) and the wealth-weighted average belief.

In the benchmark case where agents are homogeneous and have the same belief as the econometrician (\( \hat{l}_t = \hat{l}_t^R \)), the risk premium for the information shock \( d\hat{Z}_R^f \) is zero and the market only offers a constant price of risk for the exposure to the aggregate production shock \( dZ_I \). When the two groups’ beliefs are divergent, however, there is a non-zero risk premium for the information shock \( d\hat{Z}_R^f \). Moreover, this premium varies over time depending on the wealth fluctuation among agents.

The intuition is as follows. Suppose the two groups’ wealth-weighted average belief about \( l_t \) is above the econometrician’s belief. Then, relative to the econometrician, agents are more optimistic about the rise of \( f_t \) in the future, and so more optimistic about assets that are
positively exposed to $d\tilde{Z}_R^f$ (i.e., those prices are positively correlated with $f_t$). 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. As a result, the wealth fluctuation affects the difference between the wealth-weighted average belief and the econometrician’s belief and so 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 section illustrates the impact of agents’ heterogeneous expectations on bond markets by simulating the heterogeneous economy based on a set of calibrated model parameters. Theorem 1 implies that the short rate follows

$$dr_t = -\lambda_f [r_t - (l_t - \sigma^2_t)] dt + \sigma_f dZ_f.$$  

The short rate mean-reverts to a time-varying long-run mean $l_t - \sigma^2_t$. Balduzzi, Das and Foresi (1998) estimate two-factor interest rate models with this structure and find that the long-run mean of the short rate moves slowly with a mean-reversion parameter of 0.07 in their sample. Since the mean-reversion parameter of this long-run mean process corresponds to $\lambda_l$, we choose $\lambda_l$ to be 0.07, which implies that it takes $\ln(2)/\lambda_l = 9.9$ 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 also show that the mean-reversion parameter of the short rate ($\lambda_f$ in our model) ranges from 0.2 to 3 in different sample periods. We choose a value of 1 for $\lambda_f$ and this implies that it takes $\ln(2)/\lambda_f = 0.69$ year for the difference between the short rate and its long-run mean to die out by half.\footnote{These two mean-reversion parameters affect the magnitude of agents’ belief dispersion effect. Intuitively, a larger $\lambda_l$ parameter causes $l_t$ to revert faster to its long run mean $\bar{l}$, therefore making agents’ belief dispersion about $l_t$ less important for bond prices; a larger $\lambda_f$ parameter causes $f_t$ to revert faster to $l_t$, therefore making agents’ belief dispersion about $l_t$ more important for bond prices.}

19
We choose $\sigma_f = 1.25\%$ to match the short rate volatility in the data, and set $\sigma_l = 1.2\%$ so that the volatility of $l_t$ 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, since $\sigma_l$ measures agents’ aggregate consumption volatility (Theorem 1), we choose $\sigma_l = 2\%$ to match the aggregate consumption volatility in the data.\footnote{One could also choose $\sigma_l$ to match the volatility of the aggregate production. This would have little or no impact on the volatility amplification effect and the bond-yield regression results.}

Parameter $\phi_S$ directly affects the amount of belief dispersion between the two groups. We choose $\phi_S = 0.75$ to generate some modest belief dispersion: In our simulated data, the average dispersion between the two groups, $|\hat{l}_t^A - \hat{l}_t^B|$, is only 1.70\%. This amount is rather modest compared with the typical dispersions in surveys of future inflation and GDP growth rates (see footnote 1). 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 $f_0$ and $l_0$ start from their steady state value $\bar{l}$, which we set at 5\%; And, the two groups also share an identical prior belief about $l_0$ equal to the steady value $\bar{l}$; $\hat{l}_0^A = \hat{l}_0^B = \bar{l}$. All the model parameters are summarized below:

$$
\lambda_l = 0.07, \ \lambda_f = 1, \ \sigma_f = 1.25\%, \ \sigma_l = 1.2\%, \ \sigma_I = 2\%, \\
\phi_S = 0.75, \ \omega_0^A = \omega_0^B = 0.5, \ f_0 = l_0 = \hat{l}_0^A = \hat{l}_0^B = \bar{l} = 5\%.
$$

(20)

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.

### 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 10 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 percentile 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 increases...
then starts to fall slightly. The magnitude and shape of this volatility curve is similar to those estimated in Dai and Singleton (2003).

Figure 1: The term structure of bond yield volatility. Using parameters specified in equation (20), 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 zero to 10 years. The simulation is iterated 10,000 times and the figure plots the average (solid line), 95th 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 and 5th percentile (dashed lines) of the estimated volatility across the 10,000 paths in the homogeneous economy.

To illustrate the volatility amplification effect discussed in Section 3.2, we compute the volatility curve in a hypothetical homogeneous economy in which all agents hold the equal weighted average belief of the two groups in the above simulated heterogeneous economy (this average belief 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.
fluctuation. The lower solid line in Figure 1 plots the 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 10 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 homogenous case, the bond yield is a linear combination of two factors:

\[ Y_t(\tau) = \frac{a_f(\tau)}{\tau} f_t + \frac{a_l(\tau)}{\tau} i_t^R + \frac{b(\tau)}{\tau}. \]

From our earlier discussion, the loading on random factor \( f_t, \frac{a_f(\tau)}{\tau} \), decreases monotonically with \( \tau \), while the loading on random factor \( i_t^R, \frac{a_l(\tau)}{\tau} \), has a hump shape. The monotonically decreasing shape of the volatility curve reflects that the contribution of the first factor to the bond yield volatility dominates the contribution of the second factor.

To simplify our discussion of the heterogenous case, we approximate equation (18) by a linear form:

\[ Y_t(\tau) \approx \frac{a_f(\tau)}{\tau} f_t + \frac{a_l(\tau)}{\tau} \left( \omega^A_t \hat{i}^A_t + \omega^B_t \hat{i}^B_t \right) + \frac{b(\tau)}{\tau}. \]

Note that the second factor now becomes the wealth-weighted average belief \( \omega^A_t \hat{i}^A_t + \omega^B_t \hat{i}^B_t \), which is more volatile than the second factor in the homogeneous economy, \( i_t^R \). 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 a long-term bond or in the short rate over the same period. Despite its intuitive appeal, this prediction
is rejected by many empirical studies, e.g., Fama and Bliss (1987) and Campbell and Shiller (1991).

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

\[ Y_{t+1}(n-1) - Y_t(n) = \alpha_n + \beta_n \frac{Y_t(n) - Y_t(1)}{n-1}, \]  

(21)

where \( Y_t(n) \) is the \( n \)-month yield at month \( t \), \( \alpha_n \) is the regression constant, and \( \beta_n \) is the regression coefficient. They show that the expectations hypothesis is equivalent to the following null hypothesis for regression (21):

\[ \beta_n = 1. \]

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 agent 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 1952-1991. It shows that \( \beta_n \) starts with a value of 0.003 for 2-month yield, and then monotonically decreases as the bond maturity increases. \( \beta_n \) eventually takes a value of -4.226 for 10-year yield. All these coefficients are significantly different from 1 (the null), and the coefficient of 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 the wealth-weighted average belief about the future short rates is higher than the econometrician’s belief. On the one hand, it implies that agents discount long term bonds more heavily, which leads to higher long term bond yields and so larger yield spreads; on the other, 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 (20). For each
<table>
<thead>
<tr>
<th>n</th>
<th>2</th>
<th>3</th>
<th>6</th>
<th>12</th>
<th>24</th>
<th>48</th>
<th>120</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_n$</td>
<td>0.003</td>
<td>-0.145</td>
<td>-0.835</td>
<td>-1.435</td>
<td>-1.448</td>
<td>-2.262</td>
<td>-4.226</td>
</tr>
<tr>
<td>s.e.</td>
<td>(0.191)</td>
<td>(0.282)</td>
<td>(0.442)</td>
<td>(0.599)</td>
<td>(1.004)</td>
<td>(1.458)</td>
<td>(2.076)</td>
</tr>
</tbody>
</table>

| $\beta_n$ | -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) |

Table 1: The coefficients of yield change regressions. This table reports the estimates of $\beta_n$ in (21) 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 $\beta_n$ across the 10,000 simulated paths of the heterogeneous economy with parameters from (20).

simulated path, we run regression (21) 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 2 months to 10 years, with a similar trend and magnitude to that in Panel A. These coefficients are also significantly lower than 0 based on the standard errors across the 10,000 sample paths. Note that the null hypothesis holds in a homogeneous economy with each agent 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. Wachter (2006) argues for time-varying risk preference of the representative agent, while our model proposes a new mechanism based on agents’ heterogeneous expectations.

### 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
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), which regresses $n$-year bond excess returns on $n$-year forward spreads. Can our model explain this interesting phenomenon? To examine this question, we run the regressions in Cochrane and Piazzesi (2005) and Fama and Bliss (1987) using our simulated data.

Following Cochrane and Piazzesi, for each of the 10,000 simulated paths of our heterogeneous economy, we regress bond excess returns on one-year bond yield and three- and five-year forward rates:\footnote{See Dai, Singleton and Yang (2004), Cochrane and Piazzesi (2004) for more discussions on this result.}

\[
rx_{t+1}(n) = \beta_0(n) + \beta_1(n) Y_t(1) + \beta_3(n) F_t(3) + \beta_5(n) F_t(5) + \varepsilon_{t+1}(n), \quad n = 2, 3, 4, 5, \tag{22}
\]

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

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

$B_t(n)$ is the time−$t$ price of an $n$−year zero coupon bond, $Y_t(1)$ is 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 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: A tent-shaped function 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 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 in Cochrane and Piazzesi (2005) to describe bond premia of all maturities by a single factor. First, we regress the average (across maturity)
Figure 2: Coefficients in Cochrane-Piazzesi bond return regression. Using parameters specified in equation (20), we simulate the heterogeneous economy for 50 years. For each simulated path, we run Cochrane-Piazzesi regressions (22) and (24). We iterate the simulation and regressions 10,000 times and this figure plots the 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 (22) for two through 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 (24).

The excess return on forward rates:

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

to identify the tent shaped factor $TF_t$,

$$TF_t = \gamma_0 + \gamma_1 Y_t (1) + \gamma_3 F_t (3) + \gamma_5 F_t (5). \tag{23}$$

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

$$r_{x_{t+1}}(n) = b(n) TF_t + \varepsilon_{t+1} (n), \quad n = 2, 3, 4, 5. \tag{24}$$
Table 2: Coefficients in Cochrane-Piazzesi bond return regression. Using parameters specified in equation (20), we simulate the heterogeneous economy for 50 years. For each simulated path, we run Cochrane-Piazzesi regressions (22) and (24). 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 (22) for two through 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 (24).

<table>
<thead>
<tr>
<th></th>
<th>A. Unrestricted</th>
<th></th>
<th>B. Restricted</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(\beta_1(n))</td>
<td>(\beta_3(n))</td>
<td>(\beta_5(n))</td>
<td>(R^2)</td>
</tr>
<tr>
<td>(rx(2))</td>
<td>-1.11</td>
<td>2.63</td>
<td>-1.42</td>
<td>20.8%</td>
</tr>
<tr>
<td>s.e.</td>
<td>(0.05)</td>
<td>(0.38)</td>
<td>(0.38)</td>
<td></td>
</tr>
<tr>
<td>(rx(3))</td>
<td>-2.18</td>
<td>4.91</td>
<td>-2.48</td>
<td>20.1%</td>
</tr>
<tr>
<td>s.e.</td>
<td>(0.09)</td>
<td>(0.79)</td>
<td>(0.78)</td>
<td></td>
</tr>
<tr>
<td>(rx(4))</td>
<td>-3.15</td>
<td>6.63</td>
<td>-3.05</td>
<td>19.5%</td>
</tr>
<tr>
<td>s.e.</td>
<td>(0.14)</td>
<td>(1.18)</td>
<td>(1.18)</td>
<td></td>
</tr>
<tr>
<td>(rx(5))</td>
<td>-4.02</td>
<td>7.82</td>
<td>-3.15</td>
<td>19.1%</td>
</tr>
<tr>
<td>s.e.</td>
<td>(0.18)</td>
<td>(1.56)</td>
<td>(1.56)</td>
<td></td>
</tr>
</tbody>
</table>

To separately identify the values of \(\gamma_i\) \((i = 0, 1, 3, 5)\) and \(b(n)\) \((n = 2, 3, 4, 5)\), we impose \(1/4 \sum_{n=2}^{5} b(n) = 1\). This two-stage regression puts the following restrictions on the slope coefficients of regression (22):

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

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 and Bliss (1987) regressions: regressing n-year bond excess returns on n-year forward spreads. While 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 weighted average belief about the future short rates leads to higher future bond returns. Moreover, as will be elaborated next, a higher 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 weighted average belief means a larger value of $TF_t$, we first use equation (17) to derive the $\tau$-year instantaneous forward rate at time $t$, $F_t(\tau)$:

$$F_t(\tau) = a'_f(\tau)f_t + b'(\tau) + a'_l(\tau)\tilde{l}_t,$$

where $\tilde{l}_t$ is a weighted average of the two groups’ beliefs

$$\tilde{l}_t = \frac{B^H(\tau, f_t, \hat{i}_A)}{B_t} \omega^A_l A_t^A + \frac{B^H(\tau, f_t, \hat{i}_B)}{B_t} \omega^B_l B_t^B.$$

Note that $a'_l(\tau)$, the forward rate’s loading on $\tilde{l}_t$, has a hump shape with respect to $\tau$. That is, the instantaneous forward rates for the intermediate future are more sensitive to the belief about $l_t$ than the forward rates for the near and very distant future. This is a direct implication from the result, noted earlier in Section 2.4, that the intermediate term bond yields are most sensitive to the belief about $l_t$.\footnote{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 $l_t$ than the short term and very long term yields implies that the intermediate forward rates have higher exposure to the beliefs about $l_t$ than the forward rates in the near and distant future.}

Under the parameters specified in equation (20), $a'_l(\tau)$ attains the maximum at around $\tau = 3$. That is, the three year forward rate is more sensitive to $\tilde{l}_t$ than the one year and five year forward rates. Note that $\tilde{l}_t$ fluctuates with the wealth distribution of the two groups. As the optimistic group’s wealth share goes up, $\tilde{l}_t$ increases and so the hump-shape of $a'_l(\tau)$ 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 (23) from our simulated data, \( TF_t = -2.1 - 2.6Y_t(1) + 5.3F_t(3) - 2.3F_t(5). \)\(^{12}\)

\[ \text{(1)} \]

\[ \text{(3)} \]

\[ \text{(5)} \]

4 Conclusion

We have presented 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 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. We show that a modest amount of heterogeneous expectation 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.

\(^{12}\)The estimates of \( \gamma_0 \) through \( \gamma_5 \) are the average of the estimates across the 10,000 simulated paths, and are all significantly different from zero.
A Proofs

A.1 Proof of Proposition 1

Using equation (8), we take the difference of \( \hat{d}^{A}_t \) and \( \hat{d}^{B}_t \):

\[
\begin{align*}
\hat{d}^{A}_t - \hat{d}^{B}_t &= -\lambda_f(\hat{l}^{A}_t - \hat{l}^{B}_t)dt + \frac{\lambda_f}{\sigma_f} \left[ d\hat{Z}^A_f(t) - d\hat{Z}^B_f(t) \right] + \frac{2\phi_S \sigma_f}{\sqrt{1 - \phi_S^2}} dS_t \\
&= - \left( \lambda_f + \frac{\lambda_f^2}{\sigma_f^2} \right) (\hat{l}^{A}_t - \hat{l}^{B}_t) dt + \frac{2\phi_S \sigma_f}{\sqrt{1 - \phi_S^2}} dS_t.
\end{align*}
\]

We define

\[
\bar{\hat{l}}_t^M \equiv (\hat{l}^{A}_t + \hat{l}^{B}_t)/2
\]

as the average belief of group-A and group-B agents. Then, by substituting in their belief dynamics in equation (8), we have

\[
\begin{align*}
d\bar{\hat{l}}^M_t &= \frac{d\hat{l}^{A}_t + d\hat{l}^{B}_t}{2} \\
&= -\lambda_f(\bar{\hat{l}}_t^M - \bar{\hat{l}}_t) dt + \frac{\lambda_f}{\sigma_f} \left\{ \frac{1}{\sigma_f} \left[ df_t + \lambda_f(f_t - \bar{\hat{l}}_t^M)dt \right] \right\}
\end{align*}
\]

Comparing the equation above with equation (5) shows that the dynamics of \( \bar{\hat{l}}_t^M \) is identical to the dynamics of \( \hat{l}_t^R \). Thus, if \( \bar{\hat{l}}_0^M = \hat{l}_0^R \), then \( \bar{\hat{l}}_t^M = \hat{l}_t^R \).

A.2 Derivation of the equilibrium and proof of Theorem 1

We start by deriving group-i agents’ optimal investment and consumption strategies. To simplify notation, we put the return processes of securities \( f \) and \( S \) in a column vector:

\[
d\vec{R}_t = \begin{pmatrix} \frac{dp_f}{p_f} \\ \frac{dp_s}{p_s} \end{pmatrix}^t,
\]

where \(^t\) is the transpose operator. By rewriting equations (10) and (11) in group-i agents’ probability measure, we obtain

\[
d\vec{R}_t = \vec{\mu}_i dt + \Sigma \cdot d\vec{Z}_i(t),
\]

where the vector of expected returns is given by

\[
\vec{\mu}_i = \begin{pmatrix} \hat{\mu}_i^f(t) \\ \hat{\mu}_i^s(t) \end{pmatrix} = \begin{pmatrix} \mu_f(t) - \lambda_f(f_t - \hat{l}_t) \\ \mu_s(t) \end{pmatrix}
\]
and the volatility matrix $\Sigma$ and the diffusion vector $d\tilde{Z}^i(t)$ are given by

$$
\Sigma = \begin{pmatrix} \sigma_f & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad d\tilde{Z}^i(t) = \begin{pmatrix} d\tilde{Z}^i_f(t) \\ dS_t \end{pmatrix}.
$$

Note that group-$A$ and group-$B$ agents hold different expectations about security $f$'s return, but the same expectation about security $S$.

We use a vector

$$
\vec{X}^i = (x^i_f, x^i_S)'
$$

to denote the fractions of group-$i$ agents' wealth invested in securities $f$ and $S$, and $c^i_t$ as their consumption, and $x^i_f$ as the fraction of their wealth invested in the risky technology. Their wealth process follows

$$
\frac{dW^i_t}{W^i_t} = \left[ r_t - c^i_t/W^i_t + x^i_f (f_t - r_t) + x^i_f (\mu^i_t - r_t) \right] dt + \vec{X}^i \cdot \Sigma \cdot d\tilde{Z}^i_t + x^i_I dZ^i_I(t).
$$

(28)

We follow Merton (1971) to derive their optimal consumption and investment strategies:

$$
c^i_t = \beta W^i_t, \quad x^i_f = \frac{f_t - r_t}{\sigma_I^2}, \quad \text{and} \quad \vec{X}^i = (\mu^i_t - r_t)' \cdot \Sigma^{-2}.
$$

(29)

We now derive the equilibrium short rate and security returns from market clearing. The market clearing conditions require that the aggregate investment to the risky technology is equal to the total wealth in the economy:

$$
\sum_{i=A}^{B} x^i_I(t) W^i_t = W_t.
$$

By substituting in agents' investment strategy in equation (29) and dividing both sides by $W_t$, we obtain that

$$
\frac{f_t - r_t}{\sigma_I^2} \sum_{i=A}^{B} \omega^i_t = 1.
$$

Since $\sum_{i=A}^{B} \omega^i_t = 1$, we obtain the equilibrium short rate:

$$
r_t = f_t - \sigma_I^2.
$$

The market clearing conditions also require that the aggregate investment to the security $f$ is zero:

$$
\sum_{i=A}^{B} x^i_f(t) W^i_t = 0.
$$

31
By substituting in agents’ investment strategy in equation (29) and dividing both sides by \( W_t \), we obtain that

\[
\sum_{i=A}^{B} \omega_i^j \frac{\hat{\mu}^j_i - \hat{r}_i}{\sigma^j_i} = \sum_{i=A}^{B} \omega_i^j \frac{\mu_f(t) - \lambda_f(f_t - \hat{l}_i)}{\sigma^j_i} - \hat{r}_i = 0.
\]

which further implies that

\[
\mu_f(t) = r_t + \sum_{i=A}^{B} \omega_i^j \lambda_f(f_t - \hat{l}_i) = r_t + \lambda_f f_t - \lambda_f \sum_{i=A}^{B} \omega_i^j \hat{l}_i.
\]

This result shows that the contract term of security \( f \) is determined by the short rate, \( r_t \), minus the wealth-weighted average of agents’ beliefs about the drift rate of \( df_t \). Following a similar procedure, we can also derive the contract term of security \( S \):

\[
\mu_S(t) = r_t,
\]

which is equal to the short rate because all agents share the same belief that the drift rate of \( dS_t \) is zero.

Since the risky technology is the only storage tool in the economy and each agent consumes wealth at a constant rate \( \beta \), the aggregate wealth fluctuates according to

\[
\frac{dW_t}{W_t} = \frac{dI_t}{I_t} - \beta dt = (f_t - \beta) dt + \sigma_f dZ_f(t).
\]

We denote the difference in their beliefs by \( g_t \equiv \hat{l}_t^A - \hat{l}_t^B \). The following lemma provides a useful property of the wealth ratio process—\( \eta_t \) can act as the Randon-Nikodyn derivative of group-\( A \) agents’ probability measure with respect to group-\( B \) agents’ measure.

**Lemma 1** In group-\( B \) agents’ measure, the wealth ratio process fluctuates according to

\[
\frac{d\eta_t}{\eta_t} = \frac{\lambda_f}{\sigma_f}g_t \hat{Z}_f^B(t).
\]

If \( X_T \) is a random variable to be realized at time \( T > t \) and \( E^A[X_T] < \infty \), then group-\( A \) agents’ expectation of \( X_T \) at time \( t \) can be transformed into group-\( B \) agents’ expectation through the wealth ratio process between the two groups:

\[
E_t^A[X_T] = E_t^B \left[ \frac{\eta_T}{\eta_t} X_T \right].
\]
Proof: Applying Ito’s lemma to $\eta_t$ in group-$B$ agents’ probability measure, we obtain

$$\frac{d\eta_t}{\eta_t} = \frac{dW^A_t}{W^A_t} - \frac{dW^B_t}{W^B_t} + \left( \frac{dW^B_t}{W^B_t} \right) - \left( \frac{dW^A_t}{W^A_t} \right) \cdot \left( \frac{dW^B_t}{W^B_t} \right).$$

(34)

By substituting group-$B$ agents’ consumption and investment strategies into equation (28), we obtain their wealth process:

$$\frac{dW^B_t}{W^B_t} = \left[ r_t - \beta + \left( \frac{f_t - r_t}{\sigma_I} \right)^2 + (\bar{\mu}_B^t - r_t) \right] \cdot \Sigma^{-2} \cdot (\bar{\mu}_B^t - r_t) \right] \cdot dt + (\bar{\mu}_B^t - r_t) \cdot \Sigma^{-1} \cdot d\bar{Z}^B (t).$$

By following a similar procedure, we obtain group-$A$ agents’ wealth process:

$$\frac{dW^A_t}{W^A_t} = \left[ r_t - \beta + \left( \frac{f_t - r_t}{\sigma_I} \right)^2 + (\bar{\mu}_A^t - r_t) \right] \cdot \Sigma^{-2} \cdot (\bar{\mu}_A^t - r_t) \right] \cdot dt + (\bar{\mu}_A^t - r_t) \cdot \Sigma^{-1} \cdot d\bar{Z}^B (t).$$

By substituting $\frac{dW^B_t}{W^B_t}$ and $\frac{dW^A_t}{W^A_t}$ into equation (34), we obtain

$$\frac{d\eta_t}{\eta_t} = (\bar{\mu}_A^t - \bar{\mu}_B^t) \cdot \Sigma^{-1} \cdot d\bar{Z}^B (t).$$

Equation (27) implies that

$$(\bar{\mu}_A^t - \bar{\mu}_B^t) \cdot \Sigma^{-1}.$$
Equation (33) shows that the volatility of the wealth ratio is proportional to belief dispersion $|g_t|$. Intuitively, higher belief dispersion induce agents to take more aggressive speculative and so their wealth ratio becomes more volatile. Lemma 1 also shows that the wealth ratio process between agents in groups $A$ and $B$ acts as the Randon-Nikodym derivative of group-$A$ agents’ probability measure with respect to group-$B$ agents’ measure. The intuition is as follows. If group-$A$ agents assign a higher probability to a future state than group-$B$ agents, it is natural for these agents to trade in such a way that the wealth ratio between them, $W_t^A/W_t^B$, is also higher in that state. Lemma 1 implies that, as a consequence of logarithmic preference, the ratio of probabilities assigned by these groups to different states is perfectly correlated with their wealth ratio.

Next, we derive the wealth share process of group-$B$ agents in the econometrician’s probability measure. We use equations (6) and (9) to express $d\hat{Z}_f^B(t)$ in the econometrician’s probability measure:

$$d\hat{Z}_f^B(t) = d\hat{Z}_R^f(t) - \frac{\lambda_f}{\sigma_f} \left( \hat{i}_t^B - \hat{i}_t^R \right) dt.$$

Then, equation (33) implies that

$$\frac{d\eta_t}{\eta_t} = -\frac{\lambda_f^2}{\sigma_f^2} g_t \left( \hat{i}_t^B - \hat{i}_t^R \right) dt + \frac{\lambda_f}{\sigma_f} g_t d\hat{Z}_R^f(t).$$

Since $\hat{i}_t^R = \frac{1}{2} \left( \hat{i}_t^A + \hat{i}_t^B \right)$ (as in Proposition 1),

$$\frac{d\eta_t}{\eta_t} = \frac{1}{2} \frac{\lambda_f^2}{\sigma_f^2} g_t^2 dt + \frac{\lambda_f}{\sigma_f} g_t d\hat{Z}_R^f(t).$$

Then, Ito’s lemma implies that

$$d \ln (\eta_t) = \frac{\lambda_f}{\sigma_f} g_t d\hat{Z}_R^f(t).$$

To derive asset prices, we start with agents’ stochastic discount factor. When agents are homogeneous, they share the same stochastic discount factor, which is determined by their marginal utility of consumption. With a logarithmic preference, agents consume a fixed fraction of their wealth and the stochastic discount factor is inversely related to their aggregate wealth. More specifically, the stochastic discount factor, which we denote by $M_t^H$, is

$$\frac{M_t^H}{M_0^H} = e^{-\beta t} \frac{u'(c_t)}{u'(c_0)} = e^{-\beta t} \frac{c_0}{c_t} = e^{-\beta t} \frac{W_0}{W_t}.$$  

(35)
When agents have heterogeneous beliefs about the probabilities of future states, they have different stochastic discount factors. However, in the absence of arbitrage, they have to share the same security valuations. For our derivation, we will use the probability measure and the stochastic discount factor of group-\( B \) agents. Group-\( B \) agents’ consumption is

\[ c_t^B = \beta W_t^B = \frac{\omega_t^B}{\omega_t^A + \omega_t^B} \beta W_t = \frac{1}{\eta_t + 1} \beta W_t. \]

The implied stochastic discount factor is \( M_t = e^{-\beta u'(c_t^B)/u(c_t^B)} \). Substituting in agent-\( B \)’s consumption, after some algebra, we obtain

\[
\frac{M_t}{M_0} = \left( \omega_0^A \eta_t + \omega_0^B \right) \frac{M_t^H}{M_0^H}.
\]

Thus, at time \( t \), the price of a financial security that pays off \( X_T \) at time \( T \) is

\[
P_t = E_t^B \left[ \frac{M_T}{M_t} X_T \right] = E_t^B \left[ e^{-\beta(T-t)} \frac{W_t \eta_T}{W_T \eta_t} X_T \right] = \omega_t^A E_t^A \left[ e^{-\beta(T-t)} \frac{W_t \eta_T}{W_T \eta_t} X_T \right] + \omega_t^B E_t^B \left[ e^{-\beta(T-t)} \frac{W_t \eta_T}{W_T \eta_t} X_T \right].
\]

Since \( \frac{\eta_T}{\eta_t} \) is the Randon-Nikodym derivative of group-\( A \) agents’ probability measure with respect to the measure of group-\( B \) agents (Lemma 1),

\[
E_t^B \left[ e^{-\beta(T-t)} \frac{W_t \eta_T}{W_T \eta_t} X_T \right] = E_t^B \left[ e^{-\beta(T-t)} \frac{W_t \eta_T}{W_T \eta_t} X_T \right].
\]

Thus,

\[
P_t = \omega_t^A E_t^A \left[ e^{-\beta(T-t)} \frac{W_t \eta_T}{W_T \eta_t} X_T \right] + \omega_t^B E_t^B \left[ e^{-\beta(T-t)} \frac{W_t \eta_T}{W_T \eta_t} X_T \right] = \omega_t^A E_t^A \left[ \frac{M_T}{M_t} X_T \right] + \omega_t^B E_t^B \left[ \frac{M_T}{M_t} X_T \right],
\]

where \( E_t^i \left[ \frac{M_T}{M_t} X_T \right] \) is the price of the security in a homogeneous economy where only group-\( i \) agents are present.

A.3 Proof of Proposition 2

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

\[
B_t^i = B^H \left( \tau, f_t, \hat{\eta}_t \right).
\]
The bond’s return has to satisfy the following relationship with the stochastic discount factor in the homogeneous economy:

\[ E_t \left( \frac{dB^H}{B^H} \right) + E_t \left( \frac{dM_t^H}{M_t^H} \right) + E_t \left( \frac{dB^H dM_t^H}{M_t^H} \right) = 0. \tag{37} \]

Applying Ito’s lemma to equations (35) and (36) provides

\[ \frac{dM_t^H}{M_t^H} = (-f_t + \sigma_t^2) dt - \sigma_t dZ_t, \]

and

\[
\frac{dB^H}{B^H} = \left\{ -\frac{B_t^H}{B^H} - \lambda_f (f_t - \hat{f}_t) \frac{B_t^H}{B^H} - \lambda_l (\hat{l}_t - \bar{l}) \frac{B_t^H}{B^H} + \frac{1}{2} \sigma_f^2 \frac{B_{tt}^H}{B^H} \right\} dt + \\
+ \left( \sigma_f^2 \frac{B_{tt}^H}{B^H} + \frac{\lambda_f \sigma_l^2}{\sigma_f^2} \right) \frac{dB_t^H}{B^H} dZ_t(t) + \frac{\phi_{f^2}}{1 - \phi_{f^2}} \sigma_f B_t^H dS.
\]

By substituting \( \frac{dB^H}{B^H} \) and \( \frac{dM_t^H}{M_t^H} \) into equation (37), we obtain the following equation:

\[ 0 = -\frac{B_t^H}{B^H} - \lambda_f (f_t - \hat{f}_t) \frac{B_t^H}{B^H} - \lambda_l (\hat{l}_t - \bar{l}) \frac{B_t^H}{B^H} + \frac{1}{2} \sigma_f^2 \frac{B_{tt}^H}{B^H} + \frac{1}{2} \left( \frac{1}{1 - \phi_{f^2}} \sigma_l^2 - 2 \lambda_l \gamma \right) \frac{B_{tt}^H}{B^H} + \lambda_f \gamma \frac{B_{tt}^H}{B^H} - f_t + \sigma_t^2 \]

We conjecture the following solution

\[ B^H (\tau, f_t, \hat{l}_t) = e^{-a_f(\tau)f_t - a_l(\tau)\hat{l}_t - b(\tau)}. \]

By substituting the conjectured solution into the differential equation in (38) and collecting common terms, we obtain the following algebra equation:

\[ 0 = [a_f'(\tau) + \lambda_f a_f(\tau) - 1] f_t + \left[ a_l'(\tau) - \lambda_f a_f(\tau) + \lambda_l a_l(\tau) \right] \hat{l}_t \]

\[ + [b'(\tau) - \lambda_l a_l(\tau) + \frac{1}{2} \sigma_f^2 a_f(\tau)^2 + \frac{1}{2} (\sigma_l^2 - 2 \lambda_l \gamma) a_l(\tau)^2 + \lambda_f \gamma a_f(\tau) a_l(\tau) + \sigma_t^2]. \]

Since this equation has to hold for any values of \( f_t \) and \( \hat{l}_t \), their coefficients must be zero.

Thus, \( a_f(\tau), a_l(\tau), \) and \( b(\tau) \) satisfy the following differential equations

\[
a_f'(\tau) + \lambda_f a_f(\tau) - 1 = 0, \\
a_l'(\tau) - \lambda_f a_f(\tau) + \lambda_l a_l(\tau) = 0, \\
b'(\tau) - \lambda_l a_l(\tau) + \frac{1}{2} \sigma_f^2 a_f(\tau)^2 + \frac{1}{2} \left( \frac{1}{1 - \phi_{f^2}} \sigma_l^2 - 2 \lambda_l \gamma \right) a_l^2(\tau) + \lambda_f \gamma a_f(\tau) a_l(\tau) + \sigma_t^2 = 0,
\]

36
subject to the boundary conditions

\[ a_f(0) = a_i(0) = b(0) = 0. \]

Solving these equations provides the bond price formula given in Proposition 2.

### A.4 Proof of Proposition 3

To replicate the price dynamics in the heterogeneous-agent economy, we need to make the representative agent’s stochastic discount factor the same as group-B agents’ after adjusting for the difference in their probability measures. That is, the representative agent’s marginal utility have the following property in any future state:

\[ u'(c_t^B) = \eta_t^N u'(c_t^N), \]

where \( u'(c_t^B) \) is group-B agents’ marginal utility from consumption, \( u'(c_t^N) \) is the representative agent’s marginal utility, and \( \eta_t^N \) is the change of measure from the representative agent’s measure to group-B agents’ measure. Therefore, \( \eta_t^N \) has the following property

\[
\frac{d\eta_t^N}{\eta_t^N} = (\bar{\mu}_t^N - \bar{\mu}_t^B)' \Sigma^{-1} d\bar{Z}^B (t),
\]

(39)

where \( \bar{\mu}_t^N \) is the representative agent’s expected returns

\[ \bar{\mu}_t^N = \begin{pmatrix} \mu_f(t) - \lambda_f & 0 \\ \lambda_f & 0 \end{pmatrix} = \eta_t + 1. \]

Note that agents with a logarithmic preference always consume a fixed fraction of their wealth over time: \( c_t^B = \beta W_t^B \) and \( c_t^N = \beta (W_t^A + W_t^B) \). Thus, we can derive the difference in the probability measures of group-B agents and the representative agents:

\[ \eta_t^N = c_t^N c_t^B = \frac{W_t^A + W_t^B}{W_t^B} = \eta_t + 1. \]

This further implies that

\[ d\eta_t^N = d\eta_t, \]

and

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

Substituting in the dynamics of \( \frac{d\eta_t}{\eta_t} \) (Lemma 1), we obtain

\[
\frac{d\eta_t^N}{\eta_t^N} = \left( \lambda_f \frac{\eta_t}{1 + \eta_t} \right) \Sigma^{-1} d\bar{Z}^B (t).
\]

(40)
Comparing (39) and (40), we obtain
\[ \hat{l}_N^t = \hat{l}_B^t + \frac{\eta_t}{1 + \eta_t} g_t = \frac{\eta_t}{1 + \eta_t} \hat{l}_A^t + \frac{1}{1 + \eta_t} \hat{l}_B^t. \]

Note that \( \frac{\eta_t}{1 + \eta_t} \) and \( \frac{1}{1 + \eta_t} \) are the wealth shares of group-A and group-B agents.

A.5 Proof of Proposition 4

Agents’ belief dispersion about \( l_t \) could lead to speculative positions in risky securities \( f \) and \( S \). We can directly compute group-B agents’ positions in these securities. Equation (29) shows that their position in security \( f \) is
\[ n_f(t) = W_t B \hat{\mu}_f(t) - r_t \]
\[ = W_t B \hat{\mu}_f(t) - \frac{\mu_f(t) - \hat{l}_B^t}{\sigma_f^2} - r_t. \]

By substituting in \( \mu_f(t) \) from Theorem 1, we obtain that
\[ n_f(t) = \frac{\lambda_f}{\sigma_f^2} W_t \frac{\eta_t}{(\eta_t + 1)^2} g_t. \]

Note that group-B agents’ position in security \( f \) is proportional to \( g_t \). This implies as the belief dispersion \( |g_t| \) widens, group-B agents take a larger position in security \( f \).

Similarly, we can derive group-B agents’ position in securities \( S \): \( n_S(t) = 0 \). Agents do not trade security \( S \) because their disagreement in the value of \( l_t \) does not lead to a disagreement about the expected return of security \( S \). Thus, we only need to consider trading volume in security \( f \).

Since group-B agents have to trade with group-A agents to change their position, the absolute value of the change in group-B agents’ position determines trading volume in the bond markets. In our model, the change in agents’ 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 agents’ 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-B agents’ position in security \( f \), \( dn_f(t) \), whose diffusion terms are
\[ \frac{\lambda_f}{\sigma_f^2} \left[ \frac{\eta_t}{(\eta_t + 1)^2} g_t dW_t - W_t g_t \frac{\eta_t - 1}{(\eta_t + 1)^3} d\eta_t + W_t \frac{\eta_t}{(\eta_t + 1)^2} dg_t \right]. \]
The fluctuation in the position is determined by the fluctuations in the aggregate wealth, in the wealth ratio between the two groups, and in the difference in agents' beliefs. 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 \(\hat{g_t}^2\). Thus, trading volume of security \(f\) increases with agents’ belief dispersion.

### A.6 Proof of Proposition 5

By the definition of bond yield \(Y_t(\tau) = -\frac{1}{2} \log(B_t)\), its volatility is proportional to that of the bond return:

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

Applying Ito’s lemma to equation (17) in the econometrician’s probability measure provides the following diffusion terms of \(\frac{dB_t}{B_t}\):

\[
- \left[ a_f(\tau)\sigma_f + a_l(\tau)\lambda_f \sigma_f^{-1}\gamma - \frac{\lambda_f}{\sigma_f} \frac{\eta_t}{(1 + \eta_t)^2} \gamma \frac{g_t e^{-a_l(\tau)g_t/2} - e^{a_l(\tau)g_t/2}}{\eta_t e^{-a_l(\tau)g_t/2} + e^{a_l(\tau)g_t/2}} \right] d\tilde{Z}_f^B
\]

\[- a_l(\tau) \frac{\phi_S \sigma_l}{\sqrt{1 - \phi_S^2}} \frac{\eta_t e^{-a_l(\tau)g_t/2} - e^{a_l(\tau)g_t/2}}{\eta_t e^{-a_l(\tau)g_t/2} + e^{a_l(\tau)g_t/2}} dZ_S(t) \cdot \]

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

\[
\left( \frac{dB_t}{B_t} \right)^2 = \left[ a_f(\tau)\sigma_f + a_l(\tau)\lambda_f \sigma_f^{-1}\gamma + \frac{\lambda_f}{\sigma_f} \frac{\eta_t}{(1 + \eta_t)^2} K_1(g_t) \right]^2 dt + a_l^2(\tau) \frac{\phi_S^2 \sigma_l^2}{(1 - \phi_S^2)} K_2(g_t) dt
\]

where

\[
K_1 (g_t) = -g_t \frac{e^{-a_l(\tau)g_t/2} - e^{a_l(\tau)g_t/2}}{\eta_t e^{-a_l(\tau)g_t/2} + e^{a_l(\tau)g_t/2}},
\]

and

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

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.

### A.7 Proof of Proposition 6

As noted in the proof of Theorem 1, agent B’s consumption is \(c_t^B = \beta \frac{1}{\lambda_t} W_t\), and hence his marginal utility is \(e^{-\beta t} \frac{1}{\lambda_t} W_t\). Applying Ito’s lemma to it and substituting (29) and (32), we obtain (19).
References

Acemoglu, Daron, Victor Chernozhukov, and Muhamet Yildiz (2007), Learning and disagreement in an uncertain world, Working Paper, MIT.


