Speculation and the Term Structure of Interest Rates

Francisco Barillas
Goizueta Business School at Emory University

Kristoffer P. Nimark
Cornell University

We develop and estimate a tractable equilibrium term structure model populated with rational but heterogeneously informed traders that take on speculative positions to exploit what they perceive to be inaccurate market expectations about future bond prices. The speculative motive is an important driver of trading volume. Yield dynamics due to speculation are (1) statistically distinct from classical term structure components due to risk premiums and expectations about future short rates and are orthogonal to public information available to traders in real time and (2) quantitatively important, accounting for a substantial fraction of the variation of long maturity U.S. bond yields. (*JEL* E43, G12, G14)

Received May 12, 2014; editorial decision May 10, 2016 by Editor Pietro Veronesi.

A fundamental question in finance is what the economic forces are that explain variation in asset prices and returns. This paper demonstrates that allowing for heterogeneous information sets among rational traders introduces a speculative component in bond yields that is absent in models in which all traders share the same information. The speculative term is empirically important and statistically distinct from both risk premiums and terms reflecting expectations about future risk-free short rates.

Many bonds, and U.S. Treasury bonds in particular, are traded in very liquid secondary markets. In such a market, the price an individual trader will pay for a long maturity bond depends on how much he or she thinks other traders will pay for the same bond in the future. If traders have access to different information, this price may differ from what an individual trader would be willing to pay for

The authors thank David Backus, Toni Braun, Bernard Dumas, Michael Gallmeyer, Jerome Henry, Bryce Little, Elmar Mertens, Francisco Peneranda, Pontus Rendahl, Helen Rey, Ken Singleton, Thomas Sargent, Chris Telmer and seminar and conference participants at New York University, Bank of Norway, Sveriges Riksbank, EABCN conference on *Uncertainty Over the Business Cycle*, ESSIM 2009, 3rd European Workshop on General Equilibrium Theory, University of Alicante, SED 2009 and NORMAC 2009, Federal Reserve Bank of New York, Cornell University, UPF Finance Lunch, Bank of England, Queen Mary College, UC Berkeley, Federal Reserve Bank of San Francisco, 2010 NBER Summer Institute, UC San Diego, and the Western Finance Association 2011 for useful comments and suggestions. Send correspondence to Kristoffer P. Nimark, Economics Department, Cornell University, 404 Uris Hall, Ithaca, NY 14853; telephone: (607) 255-2818. E-mail: pkn8@cornell.edu.

doi:10.1093/rfs/hhx059 Advance Access publication June 14, 2017

the bond if he or she had to hold it until maturity. The possibility of reselling a bond then changes its equilibrium price as traders take speculative positions to exploit their private information.

This paper presents and structurally estimates an equilibrium model of the term structure of interest rates that is populated with traders that engage in this type of speculative behavior. In the model, individual traders can identify bonds that, conditional on their own information sets, have a positive expected excess return. In the absence of arbitrage, expected returns in excess of the risk-free rate must be compensation for risk. Traders will hold more of the bonds with a higher expected return in their portfolios. In equilibrium, the increased riskiness of a less balanced portfolio is exactly offset by the higher expected return. We show formally that heterogeneous information introduces a source of time-varying expected excess returns that, unlike the excess returns documented by, for instance, Fama and Bliss (1987) and Campbell and Shiller (1991), cannot be predicted conditional on past bond yields.

When aggregated, the speculative behavior of individual traders introduces new dynamics to bond prices. We demonstrate that when traders have heterogeneous information sets, bond yields are partly determined by a speculative component that reflects traders' expectations about the error in the average, or market, expectations of future risk-free interest rates. Since it is not possible for individual traders to predict the errors that other traders make based on information available to everybody, the speculative component in bond prices must be orthogonal to publicly available information. Heterogeneous information thus introduces a third term in bond yields that is statistically distinct from the classical components of yield curve decompositions, that is, terms due to risk-premiums and terms reflecting expectations about future risk-free short rates.

Despite the fact that the speculative component in bond yields must be orthogonal to public information, it is possible to quantify its importance using only publicly available data on bond yields. This is so because we as econometricians have access to the full sample of data, and the speculative term is orthogonal only to public information available to traders in real time. That is, we can use public information available in period $t+1, t+2, \ldots$ and so on, to back out an estimate of the speculative term in period t. The estimated model suggests that speculative dynamics are quantitatively important and can explain a substantial fraction of the variation in U.S. bond yields.

Bond prices are derived from the optimal portfolio decisions of traders, and because traders change their portfolios in response to new information, the model generates trading volume. Turnover, as measured by trading volume over average supply of bonds, is on the order of 50% for medium maturity bonds. Traders' portfolio holdings of bonds are motivated both by fundamental and speculative purposes. We find that for short and long maturity bonds, the standard deviation of a trader's speculative holding of bonds is about 80% of the standard deviation of his total bond holdings. The speculative motive appears to be least important for traders' holdings of bonds of maturities just below one year, for which the variation of the speculative holdings is about 35% of the total variation.

A necessary condition for traders to have any relevant private information about bond returns is that current bond prices do not perfectly reveal the state of the economy. There are both a priori reasons and empirical evidence to support the view that bond prices do not reveal all information relevant for predicting future bond returns. Grossman and Stiglitz (1980) argue that if it is costly to gather information and prices are observed costlessly, prices cannot fully reveal all information relevant for predicting future returns. For the bond market, the most important variable to forecast is the short interest rate. In most developed countries, the short interest rate is set by a central bank that responds to macroeconomic developments. If it is costly to gather information about the macro economy, Grossman and Stiglitz's argument implies that bond prices cannot reveal all information relevant for predicting bond returns.

In practice, there is a vast amount of financial and macroeconomic data available that central banks may use to inform policy decisions, for example, Bernanke and Boivin (2003). Mönch (2008) demonstrates within a no-arbitrage FAVAR term structure model that using a large number of macroeconomic time series does help predict future yields at intermediate to long horizons. In a few closely related papers, Joslin, Priebsch, and Singleton (2014) and Duffee (2011) present evidence suggesting that the factors that can be found by inverting yields are not sufficient to optimally predict future bond returns. They find that while the usual level, slope, and curvature factors explain virtually all of the cross-sectional variation in yields, additional factors are needed to forecast excess returns. Ludvigson and Ng (2009) provide more evidence that current bond yields are not sufficient to optimally forecast bond returns. They show that compared to using only yield data, drawing on a very large panel of macroeconomic data helps predict excess returns. Stated another way, these statistical models all suggest that current bond yields are not sufficient to optimally predict future bond yields.

Empirically, we find that the speculative component in our model contains substantial additional information about future bond returns that is not spanned by either the macro variables used by Joslin, Priebsch, and Singleton (2014) nor the macro factors in Ludvigson and Ng (2009). This result is based on an ex post smoothed estimate of the latent state and is not helpful for real-time prediction of returns. However, it demonstrates that the latent state of our model is not simply proxying for unspanned, but observed, macro variables.

If bond prices do not reveal all the information that is relevant for predicting bond returns, and if the number of potentially useful time series is very large, it becomes more probable that different traders will use different subsets of the available information to make decisions. Here, we model this by endowing traders with partly private information that they can exploit when trading. This setup also accords well with the casual observation that at least one motive for trade in assets is possession of information that is not, or at least is not believed to be, already reflected in prices. Formally, our setup is similar to the information structure in Diamond and Verecchia (1981), Admati (1985), Singleton (1987), Allen, Morris, and Shin (2006), and Bacchetta and van Wincoop (2006).

There exists a very large theoretical literature that studies asset pricing with heterogeneously informed agents. Hellwig (1980), Diamond and Verecchia (1981), Admati (1985), and Singleton (1987) are some of the early references. More recent examples include papers by Allen, Morris, and Shin (2006), Kasa, Walker, and Whiteman (2014), Bacchetta and van Wincoop (2006, 2008), Cespa and Vives (2012), and Makarov and Rytchkov (2012). These papers either present purely theoretical models or models calibrated to explain some feature of the data. The model presented here is directly estimated using likelihood-based methods. To the best of our knowledge, this is the first paper to empirically quantify the importance of heterogeneous information sets for asset prices and returns in an internally consistent structural model.

The traders that populate our model, use the information in both their private signals and in bond prices efficiently to predict bond returns. This makes our setup different from existing alternative approaches to model disagreement, or differences in beliefs, such as the term structure model in Xiong and Yan (2010). Xiong and Yan study a model in which two groups of agents misinterpret a common, but uninformative, signal about future inflation in different ways. This generates beliefs about the real return of bonds that differ across the two groups of traders. Xiong and Yan show that in such a setup, bond prices are determined by a wealth-weighted average of the two groups' beliefs and argue that differences in beliefs can explain both the failure of the expectations hypothesis and the excess volatility of long bond yields. An important difference between their setup and ours is that the traders in Xiong and Yan's model do not attempt to use the information contained in bond prices to make better predictions about future bond returns, even though this would be possible in their model. In the model of Xiong and Yan, an econometrician outside the model can predict excess returns from past bond prices, and could in principle trade on this information and would then make larger profits than the agents inside their model. To convincingly estimate the quantitative importance of speculative dynamics, we think it is important to make sure that there are no unexploited profit opportunities from running simple predictive regressions. In our model, because agents use the information contained in bond prices efficiently, an outside econometrician conditioning only on past bond prices would do worse than the agents inside the model.

Another difference-in-beliefs-based alternative to model expectation heterogeneity is to let traders learn rationally from prices but starting from heterogenous priors, like in Buraschi and Jiltsov (2006). However, rational learning from common signals implies that the beliefs of different traders will converge over time. Such an approach is thus not suitable for modeling and estimating phenomena that do not subside over time. Because agents in our model observe partly private signals about a state that is time-varying, the beliefs of our traders do not converge over time. In the term structure model of Buraschi and Whelan (2016), two agents learn about consumption growth rates from common signals but use different models to do so. In their set up, the two agents also disagree asymptotically, but in expectations, one agent is always more optimistic than the other, raising questions regarding why agents do not update their models in face of evidence that it is misspecified.

Based on these considerations, we think modeling heterogeneous expectations as arising from individual traders observing different signals while also, using the information contained in bond prices efficiently is a more suitable approach for empirical work.

1. A Bond Pricing Model

This section presents an equilibrium bond-pricing model. Traders are risk averse, rational, and ex ante identical but observe different signals relevant for predicting future bond prices. They choose a portfolio of risky bonds to maximize next-period wealth. Traders that have observed signals that make them more optimistic about the return of a given bond will hold relatively more of that bond in their portfolios. The equilibrium price of a bond is a function of the average expectations of the price of the same bond in the next period, discounted by the risk-free short interest rate. Bond prices are also affected by supply shocks that prevent equilibrium prices from revealing the average expectation of future bond prices.

The model is relatively tractable and in Section 3 we use it to draw out the consequences for term structure dynamics of relaxing the assumption that traders all have access to the same information.

1.1 Demand for long maturity bonds

Time is discrete and indexed by t. Like in Allen, Morris, and Shin (2006), there are overlapping generations of agents who each live for two periods. Each generation consists of a continuum of households with unit mass. Each household is endowed with one unit of wealth that it invests when young. When old, households unwind their asset positions and use the proceeds to consume. Unlike in the model of Allen, Morris, and Shin, the owners of wealth, that is, the households, do not trade assets themselves. Instead, a continuum of traders, indexed by $j \in (0,1)$, trade on behalf of the households, with households diversifying their funds across the continuum of traders. While not modeled explicitly here, this setup can be motivated as a perfectly competitive limit case of the mutual funds model of Garcia and Vanden (2009) that allows uninformed households to benefit from mutual funds private information, while diversifying away idiosyncratic risk associated with individual funds. More importantly, the assumption that the ownership of the assets is separated from the privately informed traders keeps the model tractable by abstracting from information-induced wealth heterogeneity.

The formal structure of the model is as follows. Trader j invests one unit of wealth in period t on behalf of households born in period t. In period $t+1$ trader j unwinds the position of the now-old generation of households who then use the proceeds to consume. Traders are infinitely lived and perform the same service for the next generation of households.

There are two types of assets: a risk-free one period bond with (log) return r_t and risky zero-coupon bonds of maturities $2,3,...,\overline{n}$ periods. Trader j chooses a vector of portfolio weights $\boldsymbol{\alpha}_t^j$ to maximize the expected log of wealth under management W_{t+1}^j in period $t+1$. That is, trader j solves the problem

$$\max_{\boldsymbol{\alpha}_t^j} E\left[\log W_{t+1}^j \mid \Omega_t^j\right], \tag{1}$$

subject to

$$W_{t+1}^j = 1 + r_{t,j}^p, \tag{2}$$

where Ω_t^j denotes trader j's information set and $r_{t,j}^p$ is the log return of the portfolio chosen by trader j in period t. All traders observe the short risk-free rate r_t as well as the price of all bonds.

In equilibrium, log returns of individual bonds will be normally distributed. However, the log return on a portfolio of assets with individual lognormal returns is not normally distributed. Following Campbell and Viceira (2002a, 2002b), we therefore use a second-order Taylor expansion to approximate the log excess portfolio return as

$$r_{t,j}^p - r_t = \boldsymbol{\alpha}_t^{\prime j}\mathbf{rx}_{t+1} + \frac{1}{2}\boldsymbol{\alpha}_t^{\prime j}\, diag\left[\Sigma_{rx,t}^j\right] - \frac{1}{2}\boldsymbol{\alpha}_t^{\prime j}\,\Sigma_{rx,t}^j \boldsymbol{\alpha}_t^j, \tag{3}$$

where $\mathbf{rx}_{t+1}$ is a vector of period $t+1$ excess returns on bonds defined as

$$\mathbf{rx}_{t+1} \equiv \begin{bmatrix} p_{t+1}^1 - p_t^2 - r_t \\ p_{t+1}^2 - p_t^3 - r_t \\ \vdots \\ p_{t+1}^{\overline{n}-1} - p_t^{\overline{n}} - r_t, \end{bmatrix}, \tag{4}$$

and p_t^n is the log price of a bond with n periods to maturity. The matrix $\Sigma_{rx,t}^j$ is the covariance of log bond returns conditional on trader j's information set. In equilibrium, conditional returns of individual bonds are normally distributed, with time-invariant conditional covariances that are common across all traders. We can thus suppress the time subscripts and trader indices on the conditional bond-return covariance matrix and write Σ_{rx} instead of $\Sigma_{rx,t}^j$ for all t and j. Maximizing the expected log wealth in Equation (2) with respect to $\boldsymbol{\alpha}_t^j$ then gives the optimal portfolio weights

$$\boldsymbol{\alpha}_t^j = \Sigma_{rx}^{-1} E\left[\mathbf{rx}_{t+1} \mid \Omega_t^j\right] + \frac{1}{2}\Sigma_{rx}^{-1} diag[\Sigma_{rx}]. \tag{5}$$

The higher return a trader expects to earn on a bond, the more will he or she hold of it in his portfolio. Because traders with different return expectations will hold different portfolios, the conditional variance of portfolio returns will differ across traders. Risk aversion prevents the most optimistic trader from demanding all of the available supply. Since each trader j has one unit of wealth to invest, integrating the portfolio weights in Equation (5) across traders yields the aggregate demand for bonds.

1.2 Bond supply

The vector of bond supply $\mathbf{s}_t$ is stochastic and distributed according to

$$\mathbf{s}_t = \boldsymbol{\mu} + \Sigma_{rx}^{-1}\mathbf{v}_t : \mathbf{v}_t \sim N\left(\mathbf{0}, VV'\right). \tag{6}$$

To simplify notation, the vector of supply shocks $\mathbf{v}_t$ are normalized by the inverse of the conditional variance of bond prices Σ_{rx}^{-1}. The supply shocks $\mathbf{v}_t$ play a similar role here as the noise traders in Admati (1985). That is, they prevent equilibrium prices from fully revealing the information held by other traders. While there may be some uncertainty about the total number of bonds outstanding, a more appealing interpretation of the supply shocks is in terms of effective supply, as argued by Easley and O'Hara (2004). They define the "float" of an asset as the actual number of assets available for trade in a given period.

The effective float may vary for several reasons that are not directly driven by expected bond returns. For instance, Treasury bonds are a preferred form of collateral that can be used as such in transactions otherwise unrelated to the bond market, for example, Bartolini et al. (2011). Many mutual funds have investment rules that specify that a fixed proportion of their assets must be held in U.S. Treasuries, see Gomes, Kotlikoff, and Viceira (2008). Such rules may force funds to buy or sell bonds when the value of other asset classes changes, affecting the available supply of Treasuries. Other financial intermediaries such as insurance companies and pension funds have strong preferences for particular maturities; that is, some financial institutions may have preferred (maturity) habitats like in Vayanos and Vila (2009) and Greenwood and Vayanos (2014). Such entities take positions that are determined by factors partly outside the bond market and may thus influence the effective supply of bonds of different maturities. The Federal Reserve also perform open-market operations by buying and selling Treasuries, again affecting the amount of Treasuries available to trade in secondary markets.

1.3 Equilibrium bond prices

Equating aggregate demand $\int \boldsymbol{\alpha}_t^j dj$ with supply $\mathbf{s}_t$ and solving for the log price p_t^n gives

$$p_t^n = \frac{1}{2}\sigma_n^2 - r_t + \int E\left[p_{t+1}^{n-1} \mid \Omega_t^j\right] dj - \Sigma_{rx}^n \boldsymbol{\mu} - v_t^n, \tag{7}$$

where $\frac{1}{2}\sigma_n^2$ and v_t^n are the relevant elements of $\frac{1}{2}diag[\Sigma_{rx}]$ and $\mathbf{v}_t$ respectively and Σ_{rx}^n is the n^{th} row of Σ_{rx}. The price of an n-period bond in period t thus depends on the average expectation in period t of the price of an $(n-1)$-period bond in period $t+1$.

1.4 The term structure of interest rates and higher order expectations

The log price of a one-period risk-free bond is the inverse of the short interest rate, that is,

$$p_t^1 = -r_t. \tag{8}$$

Taking this as the starting point, we can apply Equation (7) recursively to find the price of long maturity bonds. The log price of a two-period bond is then given by

$$p_t^2 = \frac{1}{2}\sigma_2^2 - r_t - \int E\left[r_{t+1} \mid \Omega_t^j\right] dj - \Sigma_{rx}^2 \boldsymbol{\mu} - v_t^2, \tag{9}$$

that is, p_t^2 is a function of the average first-order expectations about the next period risk-free rate r_t.

Continuing with the same logic, the price of a three-period bond is the average period t expectation of the price of a two-period bond in period $t+1$, discounted by the short rate r_t. Leading Equation (9) by 1 period and substituting into Equation(7) with $n=3$ gives

$$\begin{aligned} p_t^3 = & \frac{1}{2}\left(\sigma_2^2 + \sigma_3^2\right) - \Sigma_p^2 \boldsymbol{\mu} - \Sigma_p^3 \boldsymbol{\mu} \\ & - r_t - \int E\left[r_{t+1} \mid \Omega_t^j\right] dj \\ & - \int E\left[\int E\left[r_{t+2} \mid \Omega_{t+1}^{j'}\right] dj' \mid \Omega_t^j\right] dj \\ & - v_t^3. \end{aligned} \tag{10}$$

The expression Equation (10) demonstrates that the period t price of a three-period bond is a function not only of the average expectation of future risk-free interest rates but also of higher-order expectations. That is, the price partly depends on the average expectation in period t of the average expectation in period $t+1$ of the risk-free rate in period $t+2$. In general, second- and higher-order expectations do not coincide with first-order expectations when traders have heterogeneous information sets. The price of a three-period bond will then deviate from the "consensus value" of the bond, that is, the price the bond would have if it reflected only the average (first-order) period t expectation about risk-free interest rates in period $t+1$ and $t+2$.

Higher-order expectations will matter for the price of all bonds of maturity $n>2$. Recursive forward substitution of, (7) can be used to find a general expression for the price of an n-period bond as

$$p_t^n = \sum_{i=2}^{n} \left(\frac{1}{2}\sigma_i^2 - \Sigma_{rx}^i \boldsymbol{\mu} \right) - \sum_{k=0}^{n-1} r_{t:t+k}^{(k)} - v_t^n, \tag{11}$$

where we define the more compact notation

$$r_{t:t+k}^{(k)} \equiv \int E\left[\int E\left[\ldots \int E\left[r_{t+k} \,|\, \Omega_{t+k-1}^{j''}\right] dj'' \ldots \,|\, \Omega_{t+1}^{j'}\right] dj' \,|\, \Omega_t^j\right] dj, \tag{12}$$

for a k order expectation of r_{t+k}. The price of an n-period bond thus depends on average expectations of future short rates of order up to $n-1$. As usual, the yield y_t^n of an n-period bond can be computed as $y_t^n = -n^{-1} p_t^n$.

1.5 Unconditional bond yields

Traders form model-consistent expectations, which implies that the unconditional mean of the higher order expectations of the risk-free rate in the bond price Equation (11), coincide with the true unconditional mean. The unconditional yield of an n-period bond is thus given by

$$E\left[y_t^n\right] = E\left[r_t\right] + n^{-1} \sum_{i=2}^{n} \left(\Sigma_{rx}^i \boldsymbol{\mu} - \frac{1}{2}\sigma_i^2 \right). \tag{13}$$

The term $E[r_t]$ in (13) reflects how the average risk-free short rate affects long maturity yields. The second term, $n^{-1}\sum_{i=2}^{n} \Sigma_{rx}^i \boldsymbol{\mu}$, captures both risk premiums via the covariances in Σ_{rx}^i and supply effects from the vector $\boldsymbol{\mu}$. Risk premiums will be high if the conditional variances are large or if conditional excess returns are positively correlated. The average supply of bonds may increase or decrease bond yields depending on whether the conditional returns are positively or negatively correlated. The last component on the right-hand side of Equation (13) is a Jensen's inequality term due to the log transformation.

Unconditional yields depend on the conditional covariance of bond returns and will thus be influenced by traders' information sets. However, the unconditional yields are known to the traders in the model and do not influence their filtering problem.

2. Heterogeneous Information, Excess Returns, and Speculation

In this section we derive the main theoretical implications of relaxing the assumption that all traders have access to the same information. First, we will demonstrate that heterogeneous information introduces trader-specific risk premiums. We prove formally that, unlike classical bond risk premiums, risk premiums due to information heterogeneity must be orthogonal to publicly available information. Second, we define the speculative portfolio as the component of a trader's portfolio held to exploit what he or she perceives to be inaccurate market expectations about next-period bond prices. Third, we derive the speculative component of bond yields and prove that, just like the trader-specific component in risk premiums, it must be orthogonal to publicly available information.

2.1 Heterogeneous information and expected excess returns

The holding period return on a zero-coupon bond depends on how its price changes over time. To the extent that different traders have different expectations about future bond prices, they will also have different expectations about bond returns. In our model, this can be seen most clearly from the definition of the realized excess return on an n-period bond

$$rx_{t+1}^{n} \equiv p_{t+1}^{n-1} - p_t^n - r_t. \tag{14}$$

The excess return that trader j expects to earn on an n-period bond is thus given by

$$E\left[rx_{t+1}^{n} \mid \Omega_t^j\right] = E\left[p_{t+1}^{n-1} \mid \Omega_t^j\right] - p_t^n - r_t. \tag{15}$$

since current bond prices and short rates are directly observed by all traders. By substituting out the current bond price p_t^n using the expression in Equation (7), the excess return that trader j expects to earn on an n-period bond can be expressed as a sum of a trader-specific and common components

$$E\left[rx_{t+1}^{n} \mid \Omega_t^j\right] = \underbrace{E\left[p_{t+1}^{n-1} \mid \Omega_t^j\right] - \int E\left[p_{t+1}^{n-1} \mid \Omega_t^{j'}\right] dj'}_{trader\ specific} - \underbrace{\frac{1}{2}\sigma_n^2 + \Sigma_{rx}^n \boldsymbol{\mu} + v_t^n}_{common}. \tag{16}$$

In equilibrium, a positive expected excess return can only be earned as compensation for risk. Since individual portfolios are determined by expected excess returns and because traders are risk averse, a trader who is more optimistic than the average trader about the return of an n-period bond will hold more of it in his portfolio and have a larger conditional portfolio return variance. The risk that a more optimistic trader is compensated for is thus the risk associated with holding a portfolio with a higher conditional variance of returns.

In the absence of information heterogeneity, the expected excess return would be determined by the constants $\frac{1}{2}\sigma_n^2 + \Sigma_{rx}^n \boldsymbol{\mu}$ and the supply shock v_t^n. There is thus a time-varying component in risk premiums common to all traders. However, the component of excess return that is due to information heterogeneity is statistically distinct from the common component since it must be orthogonal to public information in real time. Before proving this statement formally, we first define the relevant information set.

Definition 1. The public information set Ω_t at time t is the intersection of the period t information sets of all traders

$$\Omega_t \equiv \bigcap_{j \in (0,1)} \Omega_t^j. \tag{17}$$

Proposition 1. The trader-specific component in the expected excess return

$$E\left[rx_{t+1}^n \mid \Omega_t^j\right] - \int E\left[rx_{t+1}^n \mid \Omega_t^{j'}\right] dj' \tag{18}$$

is orthogonal to public information in real time.

Proof. For any random variable X, the law of iterated expectations (e.g., Brockwell and Davis 2006) states that

$$E\left(E\left[X \mid \Omega'\right] \mid \Omega\right) = E(X \mid \Omega) \tag{19}$$

if and only if $\Omega \subseteq \Omega'$. Take expectations of the left-hand side of Equation (18) with respect to the public information set that is Equation (17) and use that $\Omega_t \subseteq \Omega_t^j$ to get

$$E\left[\left(E\left[rx_{t+1}^n \mid \Omega_t^j\right] - \int E\left[rx_{t+1}^n \mid \Omega_t^{j'}\right] dj'\right) \mid \Omega_t\right] = E\left[rx_{t+1}^n - rx_{t+1}^n \mid \Omega_t\right], \tag{20}$$

$$= 0 \tag{21}$$

which completes the proof. ■

In two influential papers, Fama and Bliss (1987) and Campbell and Shiller (1991) argue that excess returns on bonds can be predicted using current yields. One implication of Proposition 1 is thus that the trader-specific component in expected excess return is statistically distinct from the classic predictable excess returns documented in these papers.

2.2 The speculative portfolio

Traders that have different return expectations will hold different portfolios. We define the speculative component of trader j's portfolio as the bonds that trader j holds because he or she believes average return expectations are inaccurate. That is, the speculative component in trader j's portfolio is the difference between trader j's actual portfolio and the portfolio trader j believes the average trader holds and it is given by

$$\boldsymbol{\alpha}_t^j - E\left(\int \boldsymbol{\alpha}_t^i di \mid \Omega_t^j\right) = \Sigma_{rx}^{-1} E\left[\left(\mathbf{rx}_{t+1} - \int E\left(\mathbf{rx}_{t+1} \mid \Omega_t^i\right) di\right) \mid \Omega_t^j\right]. \tag{22}$$

The speculative component in trader j's portfolio is thus the (covariance-weighted) difference between trader j's expected returns and the returns that trader j believes the average trader expects to earn on bonds. If all other traders shared trader j's expectations, bond prices would adjust until all traders, including trader j, would hold the average portfolio. Trader j thus owns some bonds only because he or she believes that the average, or market, expectations about bond returns are incorrect.

2.3 Speculation, bond prices, and public information

When aggregated, the speculative behavior of individual traders affects the demand for bonds, and in extension, bond prices. Above, we defined the speculative portfolio in terms of differences in one-period return expectations which depend on the expected next period price. Of course, the next-period price will also be partly determined by speculative behavior, and expectations about the price further into the future, and so on. In order to take into account the total effect of speculation on a bond's price, it is helpful to first define a useful counterfactual price.

2.3.1 The consensus price. Following Allen, Morris, and Shin (2006) we define the "consensus price" $\overline{p}_t^n$ of an n-period bond as the price that would "reflect the 'average opinion' of the fundamental value of the asset properly discounted." The consensus price is thus the counterfactual price a bond would have, if by chance, all traders happened to share the average trader's period t expectations about the risk-free interest rates between period t and $t+n-1$ and this fact was common knowledge. It can be found by replacing the higher-order expectations of the risk-free rate in (11) with the average trader's first-order expectations

$$\overline{p}_t^n \equiv \frac{1}{2}\sum_{i=2}^{n}\left(\sigma_i^2 - \Sigma_{rx}^i \boldsymbol{\mu}\right) - \int \sum_{k=0}^{n-1} E\left[r_{t+k} \mid \Omega_t^j\right] dj - v_t^n. \tag{23}$$

We use the counterfactual consensus price $\overline{p}_t^n$ to define the speculative component in actual bond prices.

2.3.2 The speculative component in bond prices. The speculative component in bond prices is the difference between the actual price and the counterfactual consensus price. Taking the difference between (11) and (23), we get

$$p_t^n - \overline{p}_t^n = \sum_{k=0}^{n-1}\left(\int E\left[r_{t+k} \mid \Omega_t^j\right] dj - r_{t:t+k}^{(k)}\right). \tag{24}$$

The speculative component in an n-period bond price can thus be expressed as the difference between first- and higher-order expectations about future short interest rates.[1]

[1] In a different context, Bacchetta and van Wincoop (2006) show that a similar term (which they label the "higher order wedge") can be expressed as an average expectation error of the innovations to the fundamental process in their model.

It is straightforward to show that the speculative component in bond prices must be orthogonal to public information.

Proposition 2. The speculative term $p_t^n - \overline{p}_t^n$ is orthogonal to public information in real time, that is,

$$E\left[p_t^n - \overline{p}_t^n \mid \Omega_t^p\right] = 0. \tag{25}$$

Proof. In the Appendix. ■

While the formal proof of Proposition 2 is given in the Appendix, the logic is simple and intuitive. The speculative component in Equation (24) consists of higher-order expectation errors about the risk-free interest rate, that is, predictions about other traders' prediction errors. By definition, the public information set is available to all traders. Clearly, it is not possible for an individual trader to predict the errors that other traders are making by using information that is also available to them. The speculative component in a bond's price must therefore be orthogonal to public information available in real time.

Allen, Morris, and Shin (2006) argue that with privately informed traders, asset prices may display "drift"; that is, slow adjustment to shocks with several small price changes in the same direction. While this is true if one conditions on the actual value of the fundamental, Banerjee, Kaniel, and Kremer (2009) show in a three-period model that heterogeneous information is not sufficient to generate price drift when drift is defined as future price changes being predictable based on past price changes. Their result is a manifestation of a more general implication of rational expectations that is used in the proof of Proposition 2. Heterogeneous information cannot generate *any* any predictable phenomena conditional on information that is publicly available in real time, such as a past prices. If it were possible to predict such phenomena using public information, it would imply that agents could predict other agents' mistakes based on information that is available also to all other agents.

That the speculative component in bond prices must be orthogonal to public information is thus a consequence of the fact that traders form rational, model-consistent expectations. This makes the speculative component derived here different from the speculative component in the difference-in-beliefs model of Xiong and Yan (2010). In their model, traders are boundedly rational and do not condition on bond prices when they form expectations about future bond prices. To an outside econometrician, the speculative component in their model looks like classical risk premiums; that is, it makes excess returns predictable based on current bond yields. Indeed, Xiong and Yan proposes differences-in-beliefs as an alternative explanation to classical risk-premiums for why excess bond returns are predictable based on past bond prices. In their model, there may thus be money left on the table that an agent using a simple forecasting model could exploit.

While a setup with boundedly rational agents who agree to disagree can capture qualitative implications of heterogeneous beliefs, it is a less appealing approach to use for empirical work. Arguably, speculation would be an implausible explanation of bond yield dynamics if it could only be quantitatively important if it also implied large, but unexploited, profit opportunities from running predictability regressions. Because the traders in our model use information in bond prices efficiently, an outside econometrician conditioning only on bond prices would make smaller profits than the agents in our model. That agents form model-consistent expectations and use the information in bond prices efficiently thus allows us to quantify the importance of speculative dynamics, while making sure that there are no arbitrage opportunities from conditioning on prices left unexploited.

2.4 Decomposing bond prices

There exists a very large empirical term structure literature that implicitly or explicitly decomposes long-term interest rates into expectations about future risk-free short interest rates and risk premiums, for example, Cochrane and Piazzesi (2008) and Joslin, Singleton, and Zhu (2011). The premise for these type of two-way decompositions is that risk premiums and expectations about future risk-free interest rates are sufficient to completely account for the yield-to-maturity of a bond. However, heterogeneous information introduces a third component to bond yields due to speculative behavior by traders.

Add and subtract the consensus price of Equation (23) from the right-hand side of the price of an n-period bond of Equation (11) and rearrange to get

$$\begin{aligned} p_t^n = & \underbrace{\sum_{k=0}^{n-1} \int E\left[r_{t+k} \mid \Omega_t^j\right] dj}_{\text{Average first-order short rate expectations}} \\ & + \sum_{k=0}^{n-1} \underbrace{\left(\int E\left[r_{t+k} \mid \Omega_t^j\right] dj - r_{t:t+k}^{(k)} \right)}_{\text{Speculative component}} \\ & \underbrace{+ \frac{1}{2} \sum_{i=2}^{n} \left(\sigma_i^2 - \Sigma_{rx}^i \boldsymbol{\mu}\right) - v_t^n}_{\text{Common risk premiums}} \end{aligned} \tag{26}$$

The price of a long maturity bond can thus be expressed as the sum of average first-order expectations about future risk-free short rates, a speculative component due to higher-order prediction errors, and a risk premiums component common to all traders. From Proposition 2 we know that the speculative component must be orthogonal to public information in real time.

The speculative component is thus statistically distinct from both common risk premiums and first-order expectations about future risk-free rates.

In a model with perfect or common information, the speculative component would be zero at all times and bond prices would then be a function only of common short rate expectations and risk premiums. The speculative component would also be zero if there were no secondary markets for trading bonds. In the absence of secondary markets, bonds can only be purchased when they are issued and must then be held until maturity. In such a setting, the expectation of other traders' expectations would not matter for the equilibrium price, since the price of a zero-coupon bond at maturity is simply its face value, which is known to all traders and does not depend on the expectations of other traders. The new dynamics introduced to the term structure by heterogeneous information sets are thus dependent on the fact that long maturity bonds can be traded in secondary markets.

This ends the theoretical part of the paper. Before turning to the data, we can summarize our findings so far. With heterogeneous information sets, individual traders can identify and take advantage of predictable excess returns that would be absent in a model with only common information. We also demonstrated that the new bond-price dynamics introduced by speculative behavior must be orthogonal to public information. This has an interesting empirical implication: speculative dynamics cannot be detected using real-time public data. However, as econometricians we can use public information from periods $t+s : s>0$ to extract an estimate of the speculative component in bond yields in period t. To do so, we need to specify explicit processes for the risk-free short rate, bond supply and traders' information sets.

3. Empirical Specification

Above, bond prices were derived as functions of higher-order expectations of future short rates. To have an operational model that we can use to quantify the implications of heterogeneous information, we here specify explicit processes for the short rate, the supply of long maturity bonds and the information sets of the traders. In this section we also describe how the model can be solved and estimated.

3.1 The short rate and the exogenous factors

The short interest rate r_t is an affine function of a vector of exogenous latent factors $\mathbf{x}_t$

$$r_t = \delta_0 + \delta_x \mathbf{x}_t, \tag{27}$$

where the factors follow the vector autoregressive process

$$\mathbf{x}_t = F\mathbf{x}_{t-1} + C\mathbf{u}_t : \mathbf{u}_t \sim N(0, I). \tag{28}$$

We will normalize the short rate and factor processes by assuming that δ_x is a vector of ones, F is a diagonal matrix with the ith diagonal element represented by f_i and C is a lower triangular matrix with typical element c_{ij}. Normalizing F and C to be diagonal and lower triangular does not restrict the dynamics of r_t.

In the estimated model, $\mathbf{x}_t$ is a four-dimensional vector. This gives a sufficiently high-dimensional latent state to make the filtering problem of traders nontrivial, while keeping the model computationally tractable.

3.2 Parameterizing bond supply

The bond supply distribution in Equation (6) is parameterized as follows. The mean supply vector $\boldsymbol{\mu}$ has a typical element μ_n given by $\mu\lambda^n$. The parameter λ governs how the average supply of bonds changes with maturity n. With $\lambda > 1$, supply increases with maturity, and conversely, $\lambda < 1$ implies that average supply decreases with maturity.

The matrix V, that is, the square root of the covariance of the supply shocks $\mathbf{v}_t$, is diagonal with the nth diagonal element given by σn^{-1}. This parameterization implies that the standard deviation of the direct effect of supply shocks on bond yields is constant across maturities and that supply shocks are independent across maturities. These restrictions are imposed to economize on the number of free parameters but imply small costs in terms of fit.

3.3 Traders' information sets

All traders observe a vector of public signals containing the current short rate r_t and bond yields of maturity $2, 3, ..., \overline{n}$ collected in the vector $\mathbf{y}_t$. Heterogeneous information is introduced through trader-specific signals about the latent factors $\mathbf{x}_t$. The vector of private signals $\mathbf{z}_t^j$ observed by trader j is specified as

$$\mathbf{z}_t^j = \mathbf{x}_t + Q\zeta_t^j : \zeta_t^j \sim N(0, I_4), \tag{29}$$

where Q is a diagonal matrix with the ith diagonal element denoted q_i. Each element in the signal vector is thus the sum of the true factor and an idiosyncratic noise component. The noise is uncorrelated across signals and time.

The vector $\mathbf{s}_t^j$ defined as

$$\mathbf{s}_t^j = \begin{bmatrix} \mathbf{z}_t^{j\prime} & r_t & \mathbf{y}_t' \end{bmatrix}', \tag{30}$$

then contains all the signals that trader j observes in period t. Trader j's information set in period t also includes all previous signals

$$\Omega_t^j = \left\{ \mathbf{s}_t^j, \Omega_{t-1}^j \right\}, \tag{31}$$

and traders thus condition their expectations on the entire history of observed signals.

3.4 The law of motion of the state

When traders have heterogeneous information sets, it becomes optimal for them to form expectations about other traders' expectations. Natural representations of the state in this class of models tend to be infinite.[2] The model is solved using the method proposed in Nimark (2017). The equilibrium law of motion for the (finite dimensional) state X_t is of the form

$$X_t = MX_{t-1} + N\mathbf{e}_t, \tag{32}$$

where the state vector X_t is given by the hierarchy of higher-order expectations of the exogenous factors $\mathbf{x}_t$

$$X_t \equiv \left[\begin{array}{cccc} \mathbf{x}_t^{(0)\prime} & \mathbf{x}_t^{(1)\prime} & \cdots & \mathbf{x}_t^{(\bar{k})\prime} \end{array} \right]', \tag{33}$$

where the k order expectations is defined recursively as

$$\mathbf{x}_t^{(k)} \equiv \int E\left[\mathbf{x}_t^{(k-1)} \mid \Omega_t^j\right] dj \tag{34}$$

starting from $\mathbf{x}_t^{(0)} \equiv \mathbf{x}_t$. The solution method in Nimark (2017) relies on the fact that the impact of higher order expectations on bond prices decreases "fast enough" in the order of expectation, and that the variance of higher-order expectations is bounded by the variance of the true factors. Together, these facts imply that the equilibrium representation can be approximated with a state vector that contains only a finite number of higher-order expectations of the factors. The integer $\bar{k}$ is the maximum order of expectation considered and can be chosen to achieve an arbitrarily close approximation in the limit as $\bar{k} \to \infty$. In the estimated model, $\bar{k} = 40$.

The vector $\mathbf{e}_t$ contains all the aggregate shocks that affect the extended state X_t and includes both the factor shocks $\mathbf{u}_t$ and the supply shocks $\mathbf{v}_t$. The supply shocks do not directly affect the factors $\mathbf{x}_t$ but they do affect traders' (higher-order) expectations about $\mathbf{x}_t$ since traders use bond yields to extract information about $\mathbf{x}_t$.

Common knowledge of the model among traders is used to pin down the law of motion for X_t, that is, to find M and N in Equation (32). The logic is as follows: As usual in rational expectations models, first-order expectations $\mathbf{x}_t^{(1)}$ are optimal, that is, model consistent estimates of the actual factors $\mathbf{x}_t$. The knowledge that other traders have model-consistent expectations allow traders to treat average first-order expectations as a stochastic process with known properties when they form second-order expectations. Common knowledge of the model thus implies that second-order expectations $\mathbf{x}_t^{(2)}$ are optimal estimates of $\mathbf{x}_t^{(1)}$ given the law of motion for $\mathbf{x}_t^{(1)}$. Imposing this structure on all orders of expectations allows us to find the law of motion for the complete hierarchy

[2] See Townsend (1983), Sargent (1991), and Makarov and Rytchkov (2012).

of expectations as functions of the structural parameters of the model. The Appendix describes how to find the law of motion for the state in practice.

The state vector X_t is high dimensional, but by itself, this does not increase our degrees of freedom in terms of fitting bond yields. In fact, because the endogenous state variables $\mathbf{x}_t^{(k)}$ are rational expectations of the lower-order expectations in $\mathbf{x}_t^{(k-1)}$, the matrices M and N in the law of motion in Equation (32) are completely pinned down by the parameters of the process governing the true exogenous factors $\mathbf{x}_t$ and the precision of traders' information sets.

3.5 Bond prices and the state

For a given law of motion in Equation (32), bond prices can be derived using the average-expectation operator $H:\mathbb{R}^{\overline{k}+1}\rightarrow\mathbb{R}^{\overline{k}+1}$ that annihilates the lowest-order expectation of a hierarchy so that

$$\begin{bmatrix} \mathbf{x}_t^{(1)} \\ \mathbf{x}_t^{(2)} \\ \vdots \\ \mathbf{x}_t^{(\overline{k}+1)} \end{bmatrix} = H \begin{bmatrix} \mathbf{x}_t^{(0)} \\ \mathbf{x}_t^{(1)} \\ \vdots \\ \mathbf{x}_t^{(\overline{k})} \end{bmatrix}, \tag{35}$$

where $\mathbf{x}_t^{(k)}=0:k>\overline{k}$. The average (first-order) expectation about the state in period t is thus given by HX_t. The average expectation in period t of what the state will be in period $t+1$ is thus given by MHX_t. Combining the operator H that increases the order of expectations by one step with the matrix M from the law of motion (32) that moves expectations one step forward in time, allows us to compute the k-order expectation of the short rate in period $t+k$ as

$$r_{t:t+k}^{(k)}=\begin{bmatrix}\delta_x & \mathbf{0}\end{bmatrix}(MH)^{n-1}X_t. \tag{36}$$

Substituting Equation (36) into the bond-pricing equation that is Equation (11) then gives

$$p_t^n=\frac{1}{2}\sum_{i=2}^{n}\left(\sigma_i^2-\Sigma_{rx}^i\boldsymbol{\mu}\right)-n\delta_0-\sum_{s=0}^{n-1}\begin{bmatrix}\delta_x & \mathbf{0}\end{bmatrix}(MH)^s X_t-v_t^n. \tag{37}$$

The matrix M governs the actual dynamics of r_t while bonds are priced as if X_t were observed by all traders and followed a process governed by MH. The matrices M and MH are thus analogous to the "physical" and "risk-neutral" dynamics in a standard no-arbitrage model, though the interpretation is different.

3.6 The estimated state-space system

The state in Equation (32) and the bond price in Equation (37) can be combined into a state-space system of the form

$$X_t=MX_{t-1}+N\mathbf{e}_t \tag{38}$$

$$\overline{\mathbf{y}}_t=\overline{A}+\overline{B}X_t+\overline{R}\mathbf{v}_t. \tag{39}$$

Combining the fact that $y_t^n = -n^{-1} p_t^n$ with (37) implies that the rows of $\overline{A}$ and $\overline{B}$ that correspond to the n-period bond yield in the measurement equation are given by

$$A_n = -n^{-1}\frac{1}{2}\sum_{i=2}^{n}\left(\sigma_i^2 - \Sigma_p^i \boldsymbol{\mu}\right) + \delta_0 \tag{40}$$

$$B_n = n^{-1}\sum_{s=0}^{n-1}\left[\delta_x \quad \mathbf{0}\right](MH)^s. \tag{41}$$

The vector of parameters to be estimated is denoted $\theta \equiv \{F, C, Q, \delta_0, \boldsymbol{\mu}, \lambda, \sigma_v\}$ and consists of a total of 22 parameters. Evaluating the log likelihood function for the state-space system in Equations (38) and (39) allows us to form a posterior estimate for θ. The yields used for estimation are the 1-, 2-, 3-, 4-, and 5-year interest rates on U.S. Treasuries taken from the CRSP database. The sample period runs from July 1952 to January 2013 and contains 727 monthly observations.

We use uniform priors on all model parameters. The *Survey of Professional Forecasters* documents substantial disagreement of interest rate forecasts among survey respondents and the average cross-sectional dispersion of responses of one-quarter-ahead forecasts of the Federal Funds Rate is about 40 basis points (bps) in the 1980 to 2014 sample. To take this evidence into account, an informative prior is used on the model implied one-month-ahead forecast dispersion. The prior distribution of the standard deviation of the cross-sectional forecast dispersion is centered around 20 bps with a standard deviation of 5 bps. This ensures that a low posterior probability is associated with parameterizations that imply either counterfactually small or implausibly large degrees of forecast dispersion among the traders in the model.

The posterior parameter distributions was generated from 200,000 draws from an adaptive Metropolis algorithm (see Haario, Saksman, and Tamminen 2001), initialized from a parameter vector found by maximizing the posterior using the simulated annealing maximizer of Goffe (1996). The results reported in the next section are based on the last 100,000 draws.

4. Empirical Results

Table 1 reports the posterior estimates of the model parameters. The mode $\widehat{\theta}$ is the parameter vector from the Markov chain that achieves the highest posterior likelihood. All parameters appear to be well identified.

4.1 Model fit

Before discussing the empirical importance of the speculative motive, we assess how the model fits bond yields as well as how well it fits the cross-sectional dispersion of expectations documented in the *Survey of Professional Forecasters*.

Table 1
Posterior parameter estimates, 1952:M7-2013:M1

θ	Mode $\widehat{\theta}$	Prior dist.	Posterior 2.5%–97.5%
		Short rate process	
δ_0	0.059	$U(0,\infty)$	0.058–0.061
f_1	0.99	$U(0,0.999)$	0.98–0.99
f_2	0.96	$U(0,0.999)$	0.95–0.98
f_3	0.70	$U(0,0.999)$	0.67–0.72
f_4	0.022	$U(0,0.999)$	0.020–0.027
c_{11}	0.012	$U(0,\infty)$	0.011–0.013
c_{22}	0.0037	$U(0,\infty)$	0.0032–0.0038
c_{33}	0.012	$U(0,\infty)$	0.010–0.013
c_{44}	0.010	$U(-\infty,\infty)$	0.0094–0.013
c_{21}	−0.0015	$U(-\infty,\infty)$	(−0.0017)–(−0.0013)
c_{31}	−0.0019	$U(-\infty,\infty)$	(−0.0020)–(−0.0090)
c_{32}	−0.0045	$U(-\infty,\infty)$	(−0.0049)–(−0.0043)
c_{41}	0.00026	$U(-\infty,\infty)$	0.00024–0.00034
c_{42}	−0.020	$U(-\infty,\infty)$	(−0.021)−(−0.0019)
c_{43}	0.0067	$U(-\infty,\infty)$	0.0066–0.0068
		Noise in private signals	
q_1	0.011	$U(0,\infty)$	0.010–00.012
q_2	0.015	$U(0,\infty)$	0.013–0.017
q_3	0.0027	$U(0,\infty)$	0.0025–0.0028
q_4	0.0014	$U(0,\infty)$	0.0013–0.0016
		Bond supply	
μ	0.50	$U(0,\infty)$	0.49–0.51
λ	0.94	$U(0,\infty)$	0.94–0.96
σ_v	0.0086	$U(0,\infty)$	0.0082–0.0088

Prior and posterior distributions of model parameters estimated on a monthly sample of 1-, 2-, 3-, 4-, and 5-year bond yields from 1952:M7-2013:M1.

Figure 1 illustrates the model implied unconditional yields together with the sample counterparts. The model fits unconditional yields well. At the posterior mode, the unconditional one-year yield is 4.8% and the unconditional five-year yield is 6.3%, to be compared with the respective sample means of 5.1 and 5.7%. At the short end, the sample mean is well within the 95% posterior probability interval, but the model slightly overpredicts unconditional long yields. The upward slope is driven by the covariance structure of conditional returns. The posterior mode of λ is 0.94. A value of λ less than 1 implies that the average supply of bonds is decreasing in maturity which is consistent with the data.[3]

To evaluate the cross-sectional fit of term structure models, it is common to compute the standard deviation of pricing errors, that is, the standard deviation of the difference between the arbitrage-free yields implied by the model and the observed yields in the sample. For instance, Dai and Singleton (2000) report the standard deviation of the pricing errors for various affine term structure models ranging from 9.6 to 16.5 bps across models and yields. In models with a lower number of factors than the number of yields used in estimation, such pricing errors are necessary to avoid stochastic singularity. Our bond-pricing model

[3] Data on outstanding maturities of U.S. Treasuries is available at https://www.treasury.gov/resource-center/data-chart-center/quarterly-refunding/Pages/Latest.aspx

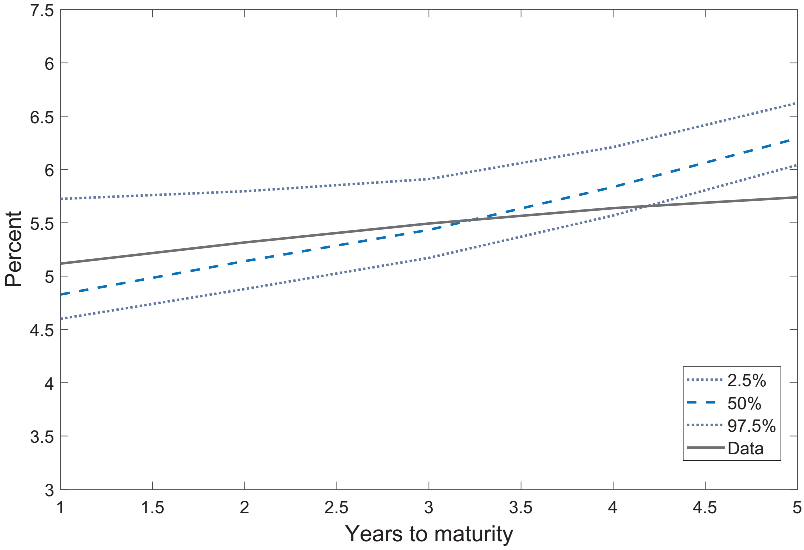

Figure 1
Unconditional yields from model (posterior median, dashed line, and 95% probability interval, dotted lines) and data (solid gray line).

provides a probabilistic description of the entire term structure, and observed yields are thus completely accounted for by Equation (11). In our framework, and unlike in many existing papers (e.g., Joslin, Singleton, and Zhu 2011; Duffee 2011), there are thus no pricing or measurement errors that drive a wedge between the observed and the model-implied arbitrage-free yields. The conditional fit of our model, according to such measures, is thus exact.[4]

The bond supply shocks in our model allow us to estimate the model with a larger number of yields than the number of latent factors. However, unlike the pricing errors used in Duffee (2011), the supply shocks in our model are priced. The supply shocks affect bond prices through 2 distinct channels. First, because traders require higher expected returns as compensation for absorbing more risky bonds in their portfolios, prices fall in response to an increase in supply. Second, the supply shocks contribute to the conditional variance of returns in Equation (7), making average bond prices lower than they would otherwise be.

Another difference between our supply shocks and the commonly used pricing errors arises because traders in our model update their expectations about the state X_t after observing bond yields. Since the supply shocks affect bond yields, supply shocks affect traders' (higher-order) expectations about the

---

[4] Note that an exact conditional fit, as defined here, does not necessarily imply a good fit in the statistical sense. Instead, it simply means that the measurement equation used in estimation coincides with the model-implied pricing equation.

persistent latent factors $\mathbf{x}_t$. So while supply shocks are independent across time and maturities, a supply shock to a single maturity bond has persistent effects on the entire cross-section of bond yields through its effect on agents' state estimates. The supply shocks thus have quite different observable implications compared to classical white-noise pricing errors. Hamilton and Wu (2011) argue that the statistical assumption of i.i.d. white-noise measurement errors in affine term structure models can be rejected. Because a single supply shock has persistent effects on yields across all maturities, our model is not subject to Hamilton and Wu (2011) critique.

Because we allow for heterogeneously informed traders our model can, in addition to fitting bond yields, potentially also fit the measured dispersion of interest rate forecasts. In the difference-in-beliefs models of Xiong and Yan (2010), Buraschi and Jiltsov (2006), and Buraschi and Whelan (2016), there are two distinct groups of agents. While the beliefs across the two groups differ, within a group all agents hold the same beliefs. This structure implies that the cross-sectional distributions of forecasts in these models have positive mass only at two points of the support. In our model, the cross-sectional distribution of traders' expectations is Gaussian and therefore more closely resembles the distribution of survey responses. This makes it straightforward to compare our model's predictions for traders' forecast disagreement to survey data, such as the *Survey of Professional Forecasters*.

At the posterior mode, the model-implied cross-sectional dispersion of forecasts across the traders is approximately 10 bps. The dispersion is neither too large or too small to appear a priori unreasonable and it is somewhat lower than the dispersion in the *Survey of Professional Forecasters*. One possible interpretation of this result is that traders in reality are better and more uniformly informed than survey respondents and this may be inferable from bond yield dynamics.

4.2 Historical decomposition of bond yields

Proposition 2 above established that the speculative term in the price of a bond can be expressed as a higher-order prediction error that is orthogonal to public information. Nevertheless, as econometricians, we can quantify this term using public bond price data since the period t higher-order prediction error is only orthogonal to information known to all traders up to period t. Since ex post, we can use the full sample and exploit information for $t+s:s>0$ to back out information about the higher-order prediction error in period t, we can form an estimate of the speculative component.

The procedure is as follows. For a given parameter vector θ, the Kalman simulation smoother can be used to draw from the smoothed-state distribution $p\left(X^T \mid \mathbf{y}^T, \theta\right)$ (e.g., Durbin and Koopman 2002). To construct the posterior distribution of the state X^T, draw repeatedly from the posterior parameter distribution $p\left(\theta \mid \mathbf{y}^T\right)$, and for each draw of θ, generate a draw from the conditional state distribution $p\left(X^T \mid \mathbf{y}^T, \theta\right)$. The speculative term in an n-period

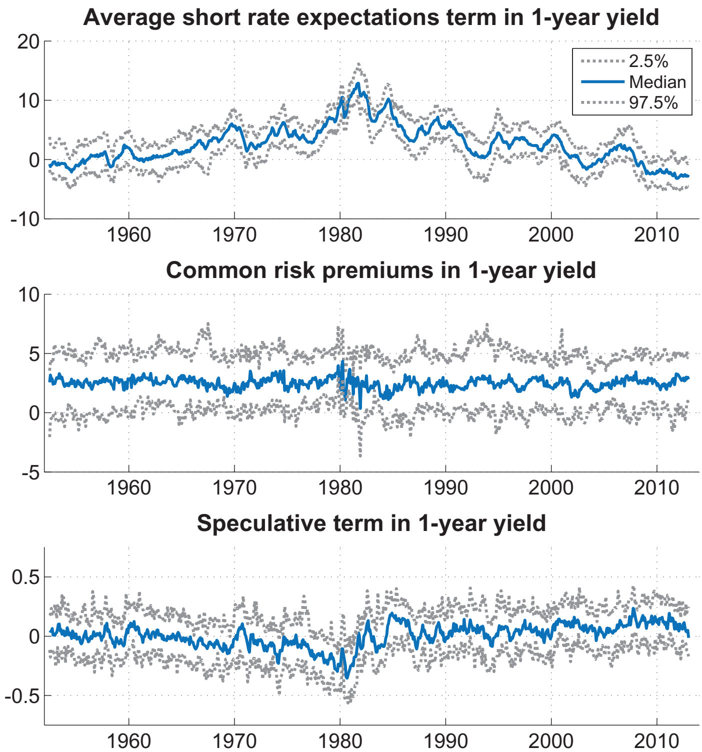

Figure 2
Historical decomposition of five-year yield, median (solid), and 95% probability interval (dotted).

bond yield, defined as $-n^{-1}\left(p_t^n - \overline{p}_t^n\right)$, can be expressed as a linear function of the state X_t. The simulated distribution of the state can then be used to compute the implied posterior distribution of the speculative component in bond yields.

Agents' average first-order expectations of future risk-free rates are also linear functions of the state. Once we have a posterior distribution of the state we can thus construct a posterior distribution of the decomposition in Equation (26) and quantify how much the terms due to average first-order expectations about future risk-free rates, common risk premiums, and the speculation each contributed to bond yields over the sample period. Figures 2 and 3 illustrate this decomposition for one- and five-year bond yields.

Like in standard models, most of the variation in bond yields is explained by variation in expected future risk-free rates. Risk premiums are positive on average and most volatile around the early 1980s for both the one- and five-year yields.

The speculative terms are positively correlated across maturities and more volatile in the five-year bond than in the one-year bond. The largest variations in the speculative term occur around 1980, when the speculative component

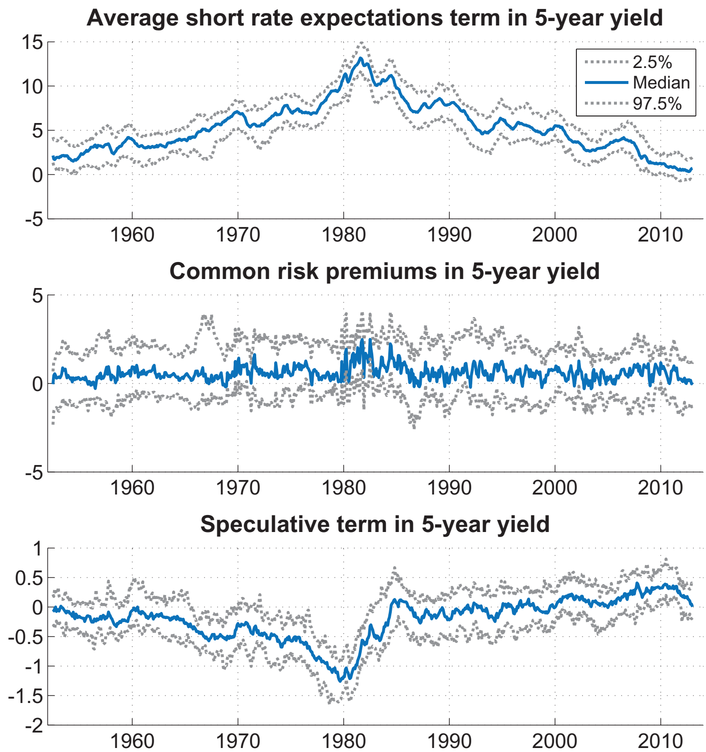

Figure 3
Historical decomposition of five-year yield, median (solid), and 95% probability interval (dotted).

for the five-year bond contributes 120 bps to the five-year yield. This period coincides with the so-called "Volcker disinflation" when the then Federal Reserve chairman Paul Volcker sharply raised interest rates to bring inflation under control, causing a recession; see for instance the account in Goodfriend and King (2005). A negative speculative component indicates that traders believed that other traders underestimated future short rates. Stated differently, the episode around 1980 during which the speculative component is large and negative was a period when individual traders perhaps believed that other traders attached too much credibility to chairman Volcker's disinflation policy and were individually more skeptical about the probability of his eventual success.

In absolute terms, the speculative term was largest in the early 1980s. However, as a fraction of the total yield, speculation appears to have been more important in the last decade. In 2011, the mode of the estimate of the speculative term reached 40 basis points at a time when five-year yields were around 3%.

4.3 Speculation and the expectations hypothesis

One way to think of the well-known failure of the expectations hypothesis is in terms of a yield decomposition: If expectations of future short rates are not sufficient to explain the variation in bond yields, the expectations hypothesis fails. In this sense, the speculative component helps to explain the failure of the expectations hypothesis since it provides a second wedge, in addition to classical risk premiums, between bond yields and expectations of future short rates.

A second way to think about the failure of the expectations hypothesis is in terms of excess returns being predictable, which is another way of stating that expectations of short rates are not enough to explain bond yields. In this sense, the speculative component does not help to explain the well-documented empirical regularity that excess returns can be predicted conditional on the current yield curve. This is so because the speculative component must be orthogonal to publicly available information in real time.

Singleton (2006) points out that violations of the expectation hypothesis in U.S. data are most pronounced when the period 1979–1983 is included in the sample. Risk-premiums-based explanations of this episode emphasize that the early 1980s was a period when either traders demanded more compensation to hold a given amount of risk because of the recession or the amount of risk was perceived to be higher than usual because of more volatile interest rates. This is also the case for our model, though it abstracts from persistent variation in common risk premiums, which may be a source of misspecification. One concern might be that because of the restrictive way that common risk premiums is introduced, the model here simply relabels some of what in reality is risk premiums as speculation. However, while the early 1980s are associated with large movements in both risk premiums and speculation, the two are not observationally equivalent. That speculative dynamics must be orthogonal to public information in real time makes it econometrically distinct from other sources of time variation in bond yields.

Barillas and Nimark (2017) presents an empirically more flexible, but more reduced form SDF-based no-arbitrage model of the term structure that also allows for heterogeneous information. That model abstracts from agents' portfolio decisions but allows for richer dynamics of risk-premiums and nests a standard three-factor affine model. The empirical importance of speculative dynamics according to that model is quantitatively similar to what we find here. That the models agree on the quantitative importance of speculative dynamics, despite being derived from very different principles, should increase our confidence in the results.

4.4 Speculation and yield volatility

One way to illustrate the quantitative importance of the speculative term in bond yields of different maturities is to compute its standard deviation relative to the standard deviation of yields. Figure 4 displays the ratio of

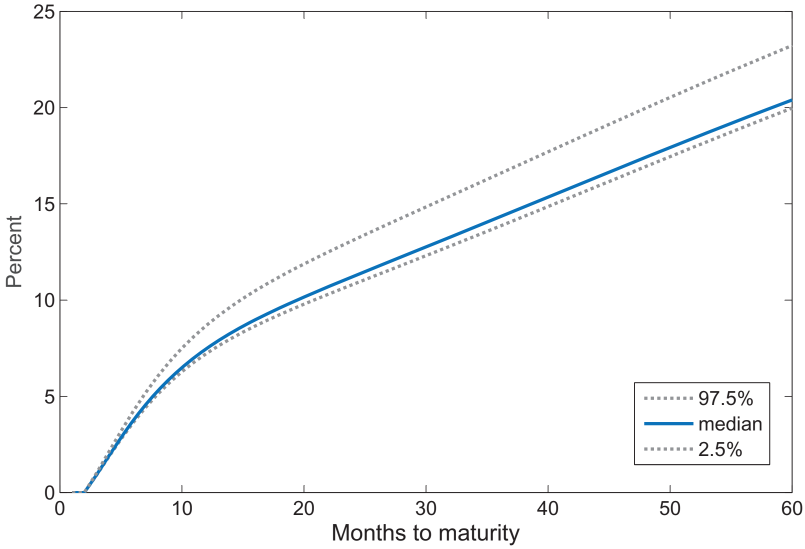

Figure 4
Relative standard deviation of speculative term and yields across maturities. Median (solid) and 95% probability interval (dotted).

the standard deviation of the speculative component and bond yields across the yield curve. At the median, the standard deviation of the speculative component relative to the standard deviation of the yield increases from just below 10% for one-year bonds to more than 20% for the five-year bond. The speculative term thus accounts for a substantial fraction of the variation in long bond yields. The fact that speculative dynamics appear to be relatively more important for longer maturities may also help explain the evidence in Gürkaynak, Sack, and Swanson (2005), who argue that current macro models of the term structure have trouble explaining the "excess" variability of long bond yields. Embedding a heterogeneous information structure in a macro model may improve these models' ability to match the variance of long-term yields.

4.5 Trading volume and speculative portfolios

Unlike existing empirical term structure models that rely on no-arbitrage and representative agent assumptions, our model can be used to study trading volume. Trading volume in our model occurs for 2 distinct reasons. First, different traders receive different information and therefore have different expectations about future returns. Traders that become relatively more pessimistic will then sell bonds to traders that become relatively more optimistic. Second, random changes in supply require traders to adjust their portfolios for markets to clear.

We define trading volume V_t^n of an n-period bond like in He and Wang (1995)

$$V_t^n = \frac{1}{2}\int \left|\alpha_t^{n,j} - \alpha_{t-1}^{n+1,j}\right| dj + \frac{1}{2}\left|v_t^n - v_{t-1}^{n+1}\right|, \tag{42}$$

where $\alpha_t^{n,j}$ is the n^{th} element of $\boldsymbol{\alpha}_t^j$, that is, the share of n-period bonds in trader j's portfolio. v_t^n is the random component of the supply of n-period bonds in period t. We can use Equation (42) to compute the model-implied trading volume at the posterior mode.

The average trading volume is u-shaped in maturities, with higher trading volume for very short and for medium to long maturity bonds. The bond maturity with the lowest trading volume is the 9-month bond. To get a sense of the magnitude of trading volume relative to supply, we can divide average volume with the average supply of bonds to get a measure of turnover. Turnover is increasing in maturity, starting from about 18% for one-year bonds and reaching about 50% for three-year bonds. Turnover is substantially higher for five-year bonds, with a trading volume of more than double the average supply of bonds.

SIFMA (2016) provides data on primary dealer trading volume of Treasuries from 2005 to 2016. The data available is not disaggregated enough to compute turnover for specific maturities, but the average monthly trading volume of Treasuries with a maturity of fewer than five years is on the order of about 80% of the total (value-weighted) amount of bonds outstanding. The turnover implied by the model is thus similar in magnitude to that in the data if we average turnover across maturities, but there are several reasons to be cautious when comparing the quantitative volume predictions of the model to the volume data from SIFMA.

First, while there are good data available on the daily trading volume among primary dealers, the data do not include trading among nonprimary dealers. The SIFMA data thus likely understate the total trading volume of Treasuries. Second, the available data refer to daily trading volumes, and the monthly averages are computed by summing daily volumes over trading days, making them not directly comparable to the predictions from a monthly frequency model. Third, bond supply in the model corresponds more closely to the effective float of bonds that are available for trade rather than the total number of bonds outstanding, which would then tend to make the model overpredict turnover.

Because bond supply is random, and in equilibrium must be absorbed into traders' portfolios, trading volume would be positive in the model even if all traders had access to the same information. To get a measure of the model's predictions about the relative importance of the speculative motive for explaining trading volume, we can compare the standard deviation of a typical trader's actual holdings of bonds given by Equation (5) with the standard deviation of the amount he or she holds for speculative purposes as defined by the speculative portfolio in Equation (22). For short and long maturity bonds, the standard deviation of the speculative holding is about 80% of the standard deviation of the total holdings. The speculative motive appears to be least important for traders' holdings of bonds of maturities just below one year, for which the variation of the speculative component is about 35% of the total variation in portfolios.

4.6 The speculative component and return predictability

In this section we present evidence from predictive regression of excess returns on the speculative components in bond yields. The smoothed estimates of the speculative components in period t is by construction a function of future changes in bond yields that could not be predicted based on period t information. It would thus not be possible for the traders in the model, nor for us as econometricians, to run these regressions in real time. However, if traders could observe the speculative component in real time, it should help them predict excess returns. Running predictive return regressions thus allows us to test whether the persistent latent structure of the model is consistent with the data and a failure to find a correlation between the speculative component and future excess returns would be a strong suggestion of a misspecified model.

Joslin, Priebsch, and Singleton (2014) find that macro factors associated with inflation expectations and industrial production, but unspanned by bond yields, help forecast excess returns. By including appropriate controls, we here also test whether the persistent latent state of our model is simply proxying for these observed, but unspanned, macro factors.

To these ends, we run predictive regressions of one-year holding excess returns of a portfolio of equally weighted 2-, 3- , 4-, and 5-year bonds on the speculative component in one-year and five-year bond yields. As controls, we include up to the first five principal components of bond yields, the *Chicago Fed National Activity Index*, and the median expected price change over the next 12 months from the *University of Michigan Survey of Consumers*. The data sample is from January 1978 to January 2013. Table 2 reports the results.

We start by confirming some well-documented regularities: First, using only principal components results in an adjusted R^2 of excess returns of about 0.2. When we also include the macro variables, that is, the *Chicago Fed National Activity Index* and inflation expectations from the *Michigan Survey*, the R^2 increases to 0.39. This is in line with the results in Joslin, Priebsch, and Singleton (2014), who document that these variables contain information that helps predict excess returns which is not spanned by the principal components of bond yields.

The R^2 increases substantially when we include the speculative components as predictors. The speculative components together with the first 5 principal components achieves an R^2 of 0.66. Consistent with the model, the latent factors would thus add substantial return forecasting power, were they to be observed in real time.

Table 2
Excess return predictability and the speculative term

Predictors	Adjusted R^2
PC123	0.14
PC12345	0.20
Macro	0.22
PC12345 + Macro	0.39
Spec1	0.44
Spec5	0.31
Spec1 + Spec5 + PC12345	0.66
Spec1 + Spec5 + Macro	0.45
Spec1 + Spec5 + PC12345 + Macro	0.70

Predictive regressions, dependent variable is the excess holding one-year return of an equally weighted portfolio of 2-, 3-, 4-, and 5-year bonds. P123 and P12345 refers to, respectively, the first three and first five principal components of yields. Macro contains the *Chicago Fed National Activity Index* and the median expected price change over the next 12 months from the *University of Michigan Survey of Consumers*. Spec1 and Spec5 are the one- and five-year speculative components defined as $-n^{-1}(p_t^n - \bar{p}_t^n)$. The data sample is from January 1978 to January 2013.

Perhaps more importantly, the predictive regressions also provide strong evidence that our latent state is not simply proxying for observed, but unspanned, macro factors. The increase in R^2 from 0.39 to 0.70 when the speculative components are added to the principal components and macro factors demonstrates that the speculative components contain information about future returns that is orthogonal to both the PCs and macro factors. The incremental increase in forecasting power from including the speculative components is almost as large as the total combined forecasting power of the PCs and macro factors. The large R^2 achieved when we include the speculative components in the conditioning set also imply that our state is not spanned by the macro factors in Ludvigson and Ng (2009), who report a maximum adjusted R^2 of 0.45.

5. Conclusions

A fundamental question in finance is which economic forces account for the variation in asset prices and returns. In this paper, we have argued that if agents have access to different information, expectations about future risk-free short rates and risk premiums may not be sufficient to explain bond yields. Instead, we have proposed that a novel speculative term, driven by heterogeneously informed traders, can account for a substantial fraction of the variation in historical U.S. bond yields along with the classic terms. We also have found that the speculative motive is an important driver of trading volume.

As a theoretical contribution, we have developed a tractable term structure model and used it to demonstrate that trader-specific excess returns, as well as the component in bond prices that is due to heterogeneous information, must be orthogonal to public information in real time. This property makes the speculative component in bond prices identified here different from the excess returns that can be predicted conditional on past bond yields, for example, Fama and Bliss (1987) and Campbell and Shiller (1991). The speculative component is thus statistically distinct from the two classical components of the yield curve: risk premiums and terms reflecting expectations about future short rates.

Because traders in our model form rational expectations, the speculative component estimated here is distinct not only from the classical components of the term structure but also from the speculative component in the difference-in-beliefs model of Xiong and Yan (2010). They propose that speculation among boundedly rational traders can provide an alternative explanation for the widely documented time variation in predictable excess returns. In their model, the speculative term would to an outside econometrician looks like classic risk premiums. From an empirical perspective, we think one appealing feature of our setup is that speculation in our model has statistical properties that makes it distinct from classical risk premiums and thus easier to identify in the data. Another appealing feature of our set-up is that traders use the information contained in prices efficiently. This ensures that in our model, there are no unexploited profit opportunities from running predictive regressions of excess returns on bond prices.

In this paper, we have estimated a relatively structural model that makes explicit assumptions about traders' utility functions and information sets. That the speculative term must be orthogonal to public information available to traders in real time makes it difficult, or perhaps impossible, to use less model-dependent, regression-based strategies to quantify the importance of speculation among rational traders. One methodological contribution of the paper is thus to demonstrate how a structural approach can be used to estimate a historical time series of the effect of speculation on bond yields using publicly available bond yield data.

Appendix

A. Proof of Proposition 2

Proposition 2 The speculative term $p_t^n - \overline{p}_t^n$ is orthogonal to public information in real time, that is,

$$E\left[p_t^n - \overline{p}_t^n \mid \Omega_t\right] = 0. \tag{A1}$$

Proof. A typical element in the sum of higher order prediction errors

$$p_t^n - \overline{p}_t^n = \sum_{k=0}^{n-1} \left(\int E\left[r_{t+k} \mid \Omega_t^j\right] dj - r_{t:t+k}^{(k)} \right) \tag{A2}$$

is can be written as

$$\int E\left[r_{t+k} \mid \Omega_t^j\right] dj - \int E\left[\int E\left[\ldots \int E\left[r_{t+k} \mid \Omega_{t+k-1}^{j''}\right] dj'' \ldots \mid \Omega_{t+1}^{j'}\right] dj' \mid \Omega_t^j\right] dj \tag{A3}$$

Taking expectations of both terms conditional on Ω_t^p gives

$$E\left(\int E\left[r_{t+k} \mid \Omega_t^j\right] dj \mid \Omega_t\right) - \tag{A4}$$

$$E\left(\int E\left[\int E\left[\ldots \int E\left[r_{t+k} \mid \Omega_{t+k-1}^{j''}\right] dj'' \ldots \mid \Omega_{t+1}^{j'}\right] dj' \mid \Omega_t^j\right] dj \mid \Omega_t\right)$$

By Definition 1 the public information set is a subset of each trader's information set, that is, that $\Omega_t \subseteq \Omega_t^j$ for each j. This fact, together with the law of iterated expectations implies that

$$E\left(\int E\left[r_{t+k} \mid \Omega_t^j\right] dj \mid \Omega_t\right) - \tag{A5}$$

$$E\left(\int E\left[\int E\left[\ldots \int E\left[r_{t+k} \mid \Omega_{t+k-1}^{j''}\right] dj'' \ldots \mid \Omega_{t+1}^{j'}\right] dj' \mid \Omega_t^j\right] dj \mid \Omega_t\right)$$

$$= E(r_{t+k} \mid \Omega_t) - E(r_{t+k} \mid \Omega_t)$$

$$= 0,$$

which completes the proof. ■

B. Solving the Model

Solving the model implies finding a law of motion for the higher-order expectations of x_t of the form

$$X_{t+1} = M X_{t-1} + N \mathbf{e}_t \tag{A6}$$

where

$$X_t \equiv \begin{bmatrix} x_t^{(0)} \\ x_t^{(1)} \\ \vdots \\ x_t^{(\bar{k})} \end{bmatrix}, \mathbf{e}_t = \begin{bmatrix} \mathbf{u}_t \\ \mathbf{v}_t \end{bmatrix}$$

That is, to solve the model, we need to find the matrices M and N as functions of the parameters governing the short rate process, the supply of long maturity bonds and the idiosyncratic noise shocks. The integer $\bar{k}$ is the maximum order of expectation considered and can be chosen to achieve an arbitrarily close approximation to the limit as $\bar{k} \to \infty$. Here, a brief overview of the method is given, but the reader is referred to Nimark (2017) for more details on the solution method.

First, common knowledge of the model can be used to pin down the law of motion for the vector X_t containing the hierarchy of higher-order expectations of x_t. Rational, that is, model-consistent, expectations of x_t thus imply a law of motion for average expectations $\mathbf{x}_t^{(1)}$ which can then be treated as a new stochastic process. Knowledge that other traders are rational means that second-order expectations $\mathbf{x}_t^{(2)}$ are determined by the average across traders of the rational expectations of the stochastic process $\mathbf{x}_t^{(1)}$. The average third-order expectation $\mathbf{x}_t^{(3)}$ is then the average of the rational expectations of the process $\mathbf{x}_t^{(2)}$, and so on. Imposing this structure on all orders of expectations allows us to find the matrices M and N. How this is implemented in practice is described below.

Second, the method exploits that the importance of higher-order expectations is decreasing in the order of expectations. This result has 2 components:

(i) The variances of higher order expectations of the factors $\mathbf{x}_t$ are bounded by the variance of the true process. More generally, the variance of $k+1$ order expectation cannot be larger than the variance of a k order expectation

$$cov\left(\mathbf{x}_t^{(k+1)}\right) \leq cov\left(x_t^{(k)}\right) \tag{A7}$$

To see why, first define the average $k+1$ order expectation error $\zeta_t^{(k+1)}$

$$\mathbf{x}_t^{(k)} \equiv \mathbf{x}_t^{(k+1)} + \zeta_t^{(k+1)} \tag{A8}$$

Since $\mathbf{x}_t^{(k+1)}$ is the average of an optimal estimate of $\mathbf{x}_t^{(k)}$ the error $\zeta_t^{(k+1)}$ must be orthogonal to $\mathbf{x}_t^{(k+1)}$ so that

$$cov\left(\mathbf{x}_t^{(k)}\right) = cov\left(\mathbf{x}_t^{(k+1)}\right) + cov\left(\zeta_t^{(k+1)}\right). \tag{A9}$$

Now, since covariances are positive semidefinite we have that

$$cov\left(\zeta_t^{(k+1)}\right) \geq \mathbf{0} \tag{A10}$$

and the inequality (A7) follows immediately. (This is an abbreviated description of a more formal proof available in Nimark (2017).

That the variances of higher order expectations of the factors are bounded is not sufficient for an accurate finite dimensional solution. We also need (ii) that the impact of the expectations of the factors on bond yields decreases "fast enough" in the order of expectation. The proof of this result is somewhat involved, and interested readers are referred to the original reference. That is, to solve the model, we need to find the matrices M, N, and B as functions of the parameters governing the short rate process, the stochastic supply shocks, and the idiosyncratic noise shocks. The integer $\overline{k}$ is the maximum order of expectation considered and can be chosen to achieve an arbitrarily close approximation to the limit as $\overline{k} \to \infty$. Here, a brief overview of the method is given, but the reader is referred to Nimark (2017) for more details on the solution method.

B.1 The law of motion of the state

To find the law of motion for the hierarchy of expectations X_t we use the following strategy. For a given M, N, and B in Equations (38) to (39) we will derive the law of motion for trader j's expectations of X_t, denoted $X_{t|t}^j \equiv E\left[X_t \mid \Omega_t^j\right]$. First, write the vector of signals S_t^j as a function of the state, the aggregate shocks and the idiosyncratic shocks

$$S_t^j = \begin{bmatrix} \mathbf{z}_\mathbf{t}^{\prime\mathbf{j}} & \widetilde{r}_t & \mathbf{y}_\mathbf{t}^{\prime} \end{bmatrix}^{\prime} \tag{A11}$$

$$= \mu_S + DX_t + R\begin{bmatrix} \zeta_t^j \\ \mathbf{u}_t \\ \widehat{\mathbf{v}}_t \end{bmatrix} \tag{A12}$$

where by (27), (29), (40) and (41) μ_S and D are given by

$$\mu_S = \begin{bmatrix} \mathbf{0}_{1\times 4} & \delta_0 & A_2 & \cdots & A_{\overline{n}} \end{bmatrix}^{\prime} \tag{A13}$$

and

$$D = \begin{bmatrix} I_4 & \mathbf{0} \\ \mathbf{1}_{1\times 4} & \mathbf{0} \\ \multicolumn{2}{c}{B} \end{bmatrix}, B \equiv \begin{bmatrix} B_1^{\prime} & B_2^{\prime} & \cdots & B_{\overline{n}}^{\prime} \end{bmatrix}^{\prime}. \tag{A14}$$

The matrix R can be partitioned conformably to the idiosyncratic and aggregate shocks

$$R = \begin{bmatrix} R_j & R_A \end{bmatrix}.$$

Trader j's updating equation of the state $X^j_{t|t}$ estimate will then follow

$$X^j_{t|t} = MX^j_{t|t-1} + K\left(S^j_t - \mu_S - DMX^j_{t|t-1}\right) \tag{A15}$$

Rewriting the observable vector S^j_t as a function of the lagged state and taking averages across traders using that $\int \zeta^j_t dj = 0$ yields

$$X_{t|t} = MX_{t|t-1} + K\left(DMX_{t-1} + (DN + R_A)\mathbf{e}_t + -DMX^j_{t|t-1}\right) \tag{A16}$$

$$= (M - KDM)X_{t|t-1} + KDMX_{t-1} + K(DN + R_A)\mathbf{e}_t \tag{A17}$$

Appending the average updating equation to the exogenous state gives us the conjectured form of the law of motion of $\mathbf{x}^{(0:\bar{k})}_t$

$$\begin{bmatrix} \mathbf{x}_t \\ X_{t|t} \end{bmatrix} = M \begin{bmatrix} \mathbf{x}_{t-1} \\ X_{t-1|t-1} \end{bmatrix} + N\mathbf{e}_t$$

where M and N are given by

$$M = \begin{bmatrix} F & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{0}_{3\times 3} & \mathbf{0} \\ \mathbf{0} & [M - KDM]_- \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ [KDM]_- \end{bmatrix} \tag{A18}$$

$$N = \begin{bmatrix} C & 0 \\ \mathbf{0} & \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ [K(DN + R_A)]_- \end{bmatrix} \tag{A19}$$

where $[\cdot]$ indicates that the a last row or column has been canceled to make a the matrix $[\cdot]$ conformable, that is, implementing that $\mathbf{x}^{(k)}_t = 0 : k > \bar{k}$. The Kalman gain K in (A15) is given by

$$K = \left(PD' + NR'\right)(DPD' + RR')^{-1} \tag{A20}$$

$$P = M\left(P - \left(PD' + NR'\right)(DPD' + RR')^{-1}\left(PD' + NR'\right)'\right)M' + NN' \tag{A21}$$

The model is solved by finding a fixed point that satisfies Equations (39), (A18), (A19), (A20), and (A21).

References

Admati, A. R. 1985. A noisy rational expectations equilibrium for multi-asset securities markets. *Econometrica* 53:629–57.

Allen, F., S. Morris, and H. S. Shin. 2006. Beauty contests and iterated expectations in asset markets. *Review of Financial Studies* 19:719–52.

Bacchetta, P., and E. van Wincoop. 2006. Can information heterogeneity explain the exchange rate determination puzzle? *American Economic Review* 96:552–76.

———. 2008. Higher order expectations in asset pricing. *Journal of Money, Credit and Banking* 40:837–66.

Banerjee, S., R. Kaniel, and I. Kremer. 2009. Price drift as an outcome of differences in higher-order beliefs. *Review of Financial Studies* 22:3707–34.

Bartolini, L., S. Hilton, S. Sundaresan, and C. Tonetti. 2011. Collateral values by asset class: Evidence from primary securities dealers. *American Economic Review* 96:552–76.

Barillas, F., and K. Nimark. 2017. Speculation and the bond market: An empirical no-arbitrage framework. Working Paper, Cornell University.

Bernanke, B. S., and J. Boivin. 2003. Monetary policy in a data-rich environment. *Journal of Monetary Economics* 50:525–46.

Brockwell, P. J., and R. A. Davis. 2006. *Time Series: theory and methods*, New York: Springer.

Buraschi, A., and A. Jiltsov. 2006. Model uncertainty and option markets with heterogeneous beliefs. *Journal of Finance* 61:2841–97.

Buraschi, A., and P. Whelan. 2016. Speculation, hedging, and interest rates. Working Paper, Imperial College London.

Campbell, J. Y., and R. J. Shiller. 1991. Yield spreads and interest rate movements: A bird's eye view. *Review of Economic Studies* 58:495–514.

Campbell, J. Y., and L. M. Viceira. 2002a. *Strategic asset allocation*. Cambridge: Cambridge University Press.

———. (2002b), On-line appendix to *Strategic asset allocation*, Cambridge: Cambridge University Press.

Cespa, G., and X. Vives. 2012. Dynamic trading and asset prices: Keynes vs. Hayek. *Review of Economic Studies* 79:539–80.

Cochrane, J., and M. Piazzesi. 2005. Bond risk premia. *American Economic Review* 94:138–60.

———. 2008. Decomposing the yield curve. Mimeo, University of Chicago.

Dai, Q., and K. J. Singleton. 2000. Specification analysis of affine term structure models. *Journal of Finance* 55:1943–78.

Diamond, D. W., and R. E. Verecchia. 1981. Aggregation in a noisy, rational expectations economy. *Journal of Financial Economics* 9:221–35.

Duffee, G. 2011. Information in (and not in) the term structure of interest rates. *Review of Financial Studies* 24:2895–934.

Durbin, J., and S. J. Koopman. 2002. A simple and efficient simulation smoother for state space time series analysis. *Biometrica* 89:603–15.

Easley, D., and M. O'Hara. 2004. Information and the cost of capital. *Journal of Finance* 59:1553–83.

Fama, E. F., and R. R. Bliss. 1987. The information in long-maturity forward rates. *American Economic Review* 77:680–92.

Garcia, D., and J. Vanden. 2009. Information acquisition and mutual funds. *Journal of Economic Theory* 144:1965–95.

Goffe, B. 1996. SIMANN: A global optimization algorithm using simulated annealing. *Studies in Nonlinear Dynamics and Econometrics* 1:169–76.

Gomes, F. J., L. J. Kotlikoff, and L. M. Viceira. 2008. Optimal life-cycle investing with flexible labor supply: A welfare analysis of life-cycle funds. Working Paper, NBER.

Goodfriend, M., and R. King. 2005. The Incredible Volcker Disinflation. *Journal of Monetary Economics* 52:981–1015.

Greenwood, R., and D. Vayanos. 2014. Bond supply and excess bond returns. *Review of Financial Studies* 27:663–713.

Grossman, S., and J. Stiglitz. 1980. On the impossibility of informationally efficient markets. *American Economic Review* 70:393–408.

Gürkaynak, R. S., B. Sack, and E. Swanson. 2005. The sensitivity of long-term interest rates to economic news: Evidence and implications for macroeconomic models. *American Economic Review* 95: 425–36.

Haario, H., E. Saksman, and J. Tamminen. 2001. An adaptive metropolis algorithm. *Bernoulli* 7:223–42.

Hamilton, J. D., and C. Wu. 2011. Testable implications of affine term structure models. *Journal of Econometrics* 178:231–42.

He, H., and J. Wang. 1995. Differential information and dynamic behavior of stock trading volume. *Review of Financial Studies* 8:919–72.

Hellwig, M. 1980. On the aggregation of information in competitive markets. *Journal of Economic Theory* 22:477–98.

Joslin, S., M. Priebsch, and K. Singleton. 2014. Risk premiums in dynamic term structure models with unspanned macro risks. *Journal of Finance* 3:1197–233.

Joslin, S., K. Singleton, and H. Zhu. 2011. A new perspective on Gaussian DTSMs. *Review of Financial Studies* 24:926–70.

Kasa, K., T. Walker, and C. Whiteman. 2014. Asset prices in a time series model with perpetually disparately informed, competitive traders. *Review of Economic Studies* 81:1137–63.

Ludvigson, S., and S. Ng. 2009. Macro factors in bond risk premia. *Review of Financial Studies* 22:5027–67.

Makarov, I., and O. Rytchkov. 2012. Forecasting the forecasts of others: Implications for asset pricing. *Journal of Economic Theory* 147:941–66.

Mönch, E. 2008. Forecasting the yield curve in a data-rich environment: A no-arbitrage factor-augmented VAR approach. *Journal of Econometrics* 146:26–43.

Nimark, K. 2017. Dynamic higher order expectations. Working Paper, Cornell University.

Townsend, R. M. 1983. Forecasting the forecasts of others. *Journal of Political Economy* 91:546–88.

Sargent, T. J. 1991. Equilibrium with signal extraction from endogenous variables. *Journal of Economic Dynamics and Control* 15:245–73.

Sifma. 2016. US bond market trading volume. http://www.sifma.org/research/statistics.aspx.

Singleton, K. J. 1987. Asset prices in a time series model with disparately informed, competitive traders. In *New approaches to monetary economics*, Eds. W. Burnett and K. Singleton. Cambridge: Cambridge University Press.

———. 2006. *Empirical dynamic asset pricing*. Princeton, NJ: Princeton University Press.

Vayanos, D., and J. L. Vila. 2009. A preferred-habitat model of the term structure of interest rates. Working Paper, NBER.

Xiong, W., and H. Yan. 2010. Heterogenous expectations and bond markets. *Review of Finacial Studies* 23: 1433–66.