Speculation and the Bond Market:
An Empirical No-arbitrage Framework

FRANCISCO BARILLAS and KRISTOFFER NIMARK*

Abstract

An affine no-arbitrage asset pricing framework is developed that allows for agents to have rational but heterogeneous expectations. The framework can match both bond yields and the observed dispersion of yield expectations in survey data. heterogeneous information introduces a speculative component in bond prices that (i) is statistically distinct from classical components such as risk-premia and expectations about future short rates and (ii) quantitatively important, at times accounting for up to 125 basis points of US yields. Allowing for heterogeneous expectations also changes the estimated relative importance of risk-premia and expectations about future short rates in historical bond yields compared to a standard affine model. The framework imposes weaker restrictions than existing heterogeneous information asset pricing models and is thus well-suited to empirically quantify the importance of relaxing the common information assumption.

*April 13, 2017. Barillas is with the Goizueta Business School at Emory University. *E-mail*: francisco.barillas@emory.edu. Nimark is with the Economics Department, Cornell University. *E-mail*: pkn8@cornell.edu. The authors thank Tarun Chordia, Thomas Dangl, Alex Kohlhas, Christian Matthes, Joonki Noh, Stefan Pitschner, Cesare Robotti, Jay Shanken, Chris Telmer, Yimei Zou and seminar participants at the NBER Summer Institute 2012, Duke University, Chicago Finance Junior Symposium, Down Under Finance Conference at Melbourne University, European Winter Finance Summit, Universitat Pompeu Fabra, the Toulouse TIGER forum 2013, Banque de France, Cornell University and Federal Reserve Bank of Chicago for useful comments and suggestions.

A fundamental question in financial economics is what the forces are that can explain the observed variation in asset prices. That expectations about future asset prices need to be an important component in any such explanation is uncontroversial. However, in spite of strong empirical evidence pointing towards economic agents having heterogeneous expectations about future returns, most empirical asset pricing models abstract away from this heterogeneity. One reason for this state of affairs may be that the literature has not yet developed methods to credibly quantify the importance of expectations heterogeneity for asset prices. This paper aims to help fill this gap. Specifically, we propose an affine no-arbitrage framework that can be used to estimate the importance of heterogeneous expectations while imposing only a minimum of structure on the data. When applied to US term structure and forecast survey data, we find that allowing for heterogeneous expectations gives rise to an empirically important speculative component in bond yields that explain a substantial fraction of variation in bond yields. Allowing for heterogeneous expectations also changes the cyclical properties of risk premia, compared to those estimated using a standard model. Together, these results suggest that the standard models that decompose bond yields into terms explained by a common market expectations about future short rates and risk premia may be inadequate to fully account for the variation in bond yields and may lead to biased estimates of these classical yield curve terms.

It is a well-documented fact that variation in expectations about future risk-free rates can explain most of the observed variation in bond yields. However, reasonable expectations about future risk-free interest rates are not sufficiently volatile to explain all of the variation in long bond yields.[1] To explain the remaining variation, classical term structure models introduce time-varying risk premia. These models thus decomposes bond yields into expectations about future interest rates and risk premia. However, both casual observation and survey evidence suggest that there is a lot of disagreement about future interest rates. For instance, the *Survey of Professional Forecasters* documents that the average cross-sectional standard deviation of the one-year-ahead forecasts of the Federal Funds Rate over the period 1981 to 2012 is approximately 40 basis points. In this paper we propose an affine no-arbitrage asset pricing framework that is consistent also with this evidence.

There exists a large theoretical literature that analyzes the consequences of heterogeneous expectations for asset pricing.[2] This literature has typically used highly stylized models to derive sharp theoretical in-

[1] E.g. Joslin, Singleton and Zhu (2011), Duffee (2002), Cochrane and Piazzesi (2008), Bauer, Rudebusch and Wu (2012), Joslin, Priebsch and Singleton (2014).

[2] Some early examples of papers studying the theoretical implications of heterogeneous information on asset prices in a rational setting are Grossman (1976), Hellwig (1980), Grossman and Stiglitz (1980), Admati (1985), Singleton (1987).

sights, but little has been done in terms of linking these results to the data. Both Allen, Morris and Shin (2006) and Bacchetta and van Wincoop (2006) demonstrate that in markets where assets are traded among agents who may not want to hold an asset until it is liquidated, expectations about the resale value of an asset will matter for its current price. The resale value of a bond depends on how much other agents will pay for it in the future, and in extension, on how much other agents think that the bond can be resold for even further into the future. In the words of Townsend (1983), agents thus need to "forecast the forecasts of others". When agents have access to different information about future fundamentals, the price of the asset may then deviate systematically from the "consensus value" defined as the hypothetical price that would reflect the average opinion of the (appropriately discounted) fundamental value of the asset. These deviations from the consensus price occur because an individual agent's expectation about the resale value of an asset can with heterogeneous information sets be different from what the individual agent would be willing to pay for the asset if he were to hold it until maturity. Heterogeneous expectations then give rise to speculative "beauty contest"-type behavior of the form described by Keynes (1936).

In this paper we propose a flexible empirical framework that can be used to quantify the empirical importance of the type of speculative behaviour described above. We do this by extending the standard affine no-arbitrage asset pricing framework to allow for heterogeneous information. The approach we take here imposes as little structure on the data as possible, while still ensuring that bond prices are arbitrage-free. The advantage of imposing only no-arbitrage restrictions instead of a full equilibrium structure is the same with heterogeneous information as in a full information setting: By imposing weaker restrictions that should hold across a wide class of models, we let the data speak louder. The downside of not writing down a complete equilibrium model is that the results may be more difficult to interpret. However, using a less restrictive but empirically more flexible approach should allow us to robustly identify the role of speculation in bond yields. The empirical results from the present paper can then be used to guide the search for more structural, and perhaps more restrictive, equilibrium models that can match these facts.[3]

The model developed here stays as close as possible to the large empirical literature that uses affine models to study asset prices, while extending it to allow for heterogeneous expectations. In the standard full information affine no-arbitrage framework, variation across time in expected excess returns is explained

[3]An analogy is how Structural Vector Auto Regressions (SVARs) have been used in the macroeconomic literature to identify the effect of monetary policy shocks while imposing minimal restrictions on the dynamics of the data, e.g. Sims (1980) and Leeper, Sims and Zha (1996). The empirical models have been followed by more restrictive general equilibrium models developed and parameterized to match the evidence from the SVARs, e.g. Christiano, Eichenbaum and Evans (2005).

by variation across time in either the amount of risk or in the required compensation for a given amount of risk. Standard affine models of Dai and Singleton (2000) and Joslin, Singleton and Zhu (2011) focus on the latter and identify the price of risk as an affine function of a small number of factors that also determine the dynamics of the risk-free short rate. Here, we specify a model in which variation in expected excess returns across individual agents, in the absence of arbitrage, must be accompanied by variation across agents in the required compensation for risk. The framework is flexible and nests a standard affine Gaussian term structure model if the signals observed by agents are perfectly informative about the state. This facilitates comparison of our results to the large existing literature on affine term structure models. However, the framework is general and can also be used to price other classes of assets.

The agents in our proposed model have heterogeneous expectations about future bond yields which makes it possible to use individual survey responses of interest rate forecasts in combination with likelihood based methods to estimate the parameters of the model. When implementing the framework empirically, we treat the individual responses in the *Survey of Professional Forecasters* as being representative of the bond yield expectations of agents randomly drawn from the population of agents in the model. Unlike papers that treat survey data as a noisy measure of a single representative agent's forecast, we use the full cross-section of survey responses to estimate the precision of the heterogeneous information sets available to the agents in the model. We believe our paper is the first to use survey data to estimate a term structure model that is consistent with the observed dispersion of survey forecasts.[4]

We document several empirical results. First, we perform a novel three-way decomposition of historical bond yields and show that in addition to the classic components due to risk premia and expectations about future risk-free short rates, heterogeneous information introduces a third term due to speculation. The speculative component in bond yields is quantitatively important, accounting for up to 125 basis points in the early 1990s and up to 100 basis point of yields in the low nominal yield environment of the last decade.

The speculative term arises when individual agents exploit their private information to predict other agents' prediction errors. In the model, all agents form rational, or model consistent, expectations and use all available information efficiently. While it is possible for agents to use their private information to predict other agents' prediction errors, it is not possible to do so conditional only on information that is also available to all other agents. The speculative component must therefore be orthogonal to all public

[4] D'Amico, Kim and Wei (2008), Chun (2011) and Piazzesi and Schneider (2011) also use survey data to estimate term structure dynamics. However, these papers use only the mean, and not the individual responses, of the surveys to estimate their models.

information available in real time, and we prove this formally. This feature also makes the speculative component statistically distinct from traditional risk premia, which can be predicted conditional on publicly available information such as bond prices. The speculative term that we identify in the data is thus not simply a relabelling of the classical terms explained by expectations about future risk-free interest rates and risk-premia.

Second, allowing for heterogeneous information changes the cyclical properties of the common component of risk premia, as compared to a full information model. Risk premia estimated from the model with heterogeneous information is less volatile than and imperfectly correlated with risk premia extracted using the nested full information model of Joslin, Singleton and Zhu (2011). In particular, the heterogeneous information model attributes much less of the high long maturity yields during the period of the Volcker disinflation in the early 1980s to large risk premia than standard models, e.g. Cochrane and Piazzesi (2008). Instead, our model attributes most of the high long maturity bond yields of that period to expectations about future short interest rates.

In the model presented here, agents use both their private signals and the information in current bond yields to form expectations about future bond yields. This has two important implications. First, and as argued above, the speculative term must be orthogonal to bond prices in real time. Second, agents' information cannot be too precise if the model is to fit the cross-sectional dispersion of survey forecasts. The information in the cross-section of survey forecasts thus clearly disciplines the model parameters that directly govern the precision of agents' information. If the agent-specific signals are too precise or so noisy that they will be disregarded, the model will fail to fit the cross-sectional dispersion of forecasts in the Survey of Professional Forecasters. Less obviously, using the full cross-section of individual survey responses also restricts the dynamics of bond yields. If observing bond yields reveals the latent factors perfectly, agents will also disregard their agent-specific signals. Parameterizations of the model that make the latent factors an invertible function of bond yields will thus be rejected by the data, since too informative bond prices would imply a counter-factually degenerate cross-sectional distribution of expectations. That the model is forced to match the observed dispersion of yield forecasts thus empirically imposes restrictions that are similar to the theoretical restrictions imposed in models with unspanned factors, e.g. Duffee (2011), Joslin, Priebsch and Singleton (2014) and Barillas (2013). These papers propose models in which by construction, the state is not an invertible function of the cross-section of bond yields and show that this helps explain excess returns. Our framework demonstrates that this type of restrictions are a natural implication of agents disagreeing about future bond returns.

The restrictions on the speculative component and the informativeness of bond prices discussed above are consequences of the fact that the agents in our model use the information in bond yields to form rational expectations. In models employing alternative assumptions to generate expectations heterogeneity these restrictions may be absent. For instance, in the difference-in-beliefs term structure models of Xiong and Yan (2010) and Chen, Joslin and Tran (2010, 2012) agents' beliefs are posited to follow exogenous processes and agents do not use the information contained in current bond prices to update their expectations about future bond prices. In the model of Xiong and Yan (2010), the speculative component in bond yields would to an outside econometrician be indistinguishable from traditional time-varying risk premia implying that an outside econometrician conditioning only on current bond prices would do better than the agents inside the model in terms of predicting bond returns. We think that the fact there are no such opportunities for an outside econometrician implied by our framework makes it more suitable for empirically quantifying the importance of the speculative motive.

I. An affine term structure model with heterogeneous information

This section describes a framework for arbitrage-free asset pricing where agents have heterogeneous information relevant for predicting future bond returns. The basic set-up follows as closely as possible the large existing affine term structure literature (see Duffie and Kan 1996 and Dai and Singleton 2000). However, allowing for heterogeneously informed agents necessitates three changes in terms of how the model is specified and solved relative to the standard full information set-up.

First, in the absence of arbitrage, heterogeneity in expected returns implies heterogeneity in required compensation for risk. We therefore need to specify a functional form for the individual agents' stochastic discount factors that is consistent with heterogeneity in agents' required compensation for risk. Below we propose a form that is analogous to the standard formulation under full information. This strategy allows for a flexible empirical specification while nesting as special cases both standard full information affine model such as Joslin, Singleton and Zhu (2011) and more restrictive equilibrium models such as Barillas and Nimark (2016).

Second, for the model to be able to match the dispersion of expectations as measured by survey data, equilibrium prices must not reveal the state perfectly. In order to prevent prices from being too informative, we therefore introduce maturity specific shocks. Formally, these take the same form as supply by noise traders in equilibrium models of asset pricing with heterogeneous information such as Admati (1985). Similarly to the stochastic supply by noise traders in equilibrium models, the maturity specific shocks in our model are priced factors. They contribute both to the risks that need to be priced as well as changes how much compensation agents require to hold a given amount of risk. Here we show how to incorporate this type of shocks directly into the specification of the stochastic discount factor so that the "noisy" prices are arbitrage-free. This may be of independent interest to some readers as an alternative to the common strategy of assuming that the arbitrage-free bond yields are observed with (unpriced) measurement errors. In our approach, there is no wedge between the arbitrage-free and the observed yields.

Third, heterogeneous information sets make it necessary for agents to "forecast the forecasts of others", e.g. Townsend (1983). The reason is that the price of a bond today partly depends on what agents think other agents will pay for the bond in the future. With heterogeneously informed traders, agents may expect others to pay more or less for a bond in the future than they would be willing to pay themselves, were they to hold on to the bond until it matures. Speculation in our model is driven by agents trying to exploit that markets may, from the perspective of an individual trader, misprice bonds in the future. To predict mispricing by other agents, individual agents need to form higher-order expectations, i.e. expectations about other agents expectations. These higher-order expectations then need to be included in the state vector. The law of motion of the state is then endogenous and depends partly on agents' information sets. Because agents observe bond prices, and because bond prices therefore affect the expectations that make up the state vector, bond prices have to be determined jointly with the law of motion of the state. Heterogeneous information thus introduces an additional step in deriving the process for bond prices that is not present in a full information set-up with only exogenous state variables.

A. *Stochastic discount factors and heterogeneous information*

An important implication of information heterogeneity is that in equilibrium, agents can only disagree about asset returns if they also require different compensation for risk. This argument is quite general and can be understood without reference to a fully specified model. To see how, consider the fundamental pricing equation of a standard common information model. There, the price P_t^n of a zero-coupon, no-default bond with n periods to maturity is given by

$$P_t^n = E\left[M_{t+1}P_{t+1}^{n-1} \mid \Omega_t\right] \tag{1}$$

where Ω_t is the common information set in period t and M_{t+1} is the stochastic discount factor. In the absence of arbitrage, this relationship has to hold for all maturities n. In a model with heterogeneous information a similar relation holds, except that the stochastic discount factor is now agent-specific so that for all agents $j \in (0,1)$ and all maturities n, the relationship

$$P_t^n = E\left[M_{t+1}^j P_{t+1}^{n-1} \mid \Omega_t^j\right] \tag{2}$$

must hold.[5] Here Ω_t^j is the information set of agent j in period t. All agents observe the current price of bonds so the left hand side of (2) is common to all agents. However, agent-specific information sets introduce heterogeneity in expectations of P_{t+1}^{n-1}. Note that, for (2) to continue to hold when expectations about P_{t+1}^{n-1} differ across agents, the stochastic discount factor M_{t+1}^j must also be agent-specific. If in period t agent j is more optimistic than the average agent about bond prices in period $t+1$, he must then also require more compensation for risk than the average agent. The agent specific components in the expectations of P_{t+1}^{n-1} and M_{t+1}^j must therefore move in opposite directions. There is thus a close relationship between heterogeneity in expectations about future prices and heterogeneity in stochastic discount factors. Any stochastic discount factor based framework that incorporates heterogeneity in expected returns must therefore allow for heterogeneity in stochastic discount factors as well. As a consequence, the empirical framework presented below features agent-specific state variables. These state variables are a sufficient statistic both for an individual agent's expectations about future bond prices as well as for his required compensation for risk.

B. Bond prices discount factors and higher-order expectations

heterogeneous expectations not only implies that stochastic discount factors must be heterogeneous across agents, but also that higher-order expectations affect the price of long maturity bonds. This can be demonstrated by first taking the log of the no arbitrage condition (2) to get

$$p_t^n = E\left[m_{t+1}^j \mid \Omega_t^j\right] + E\left[p_{t+1}^{n-1} \mid \Omega_t^j\right] + \frac{1}{2}Var\left(m_{t+1}^j + p_{t+1}^{n-1} \mid \Omega_t^j\right). \tag{3}$$

[5]Papers that feature heterogeneous discount factors include Mankiw (1986) and Constantinides and Duffie (1996).

where lower case letters denote the log of the variable denoted by the corresponding upper case letter. We can iterate (3) forward in time to express the log price of a bond as the sum of higher-order expectations about future stochastic discount factors so that

$$\begin{aligned} p_t^n &= E\left[m_{t+1}^j \mid \Omega_t^j\right] \\ &+E\left[\int E\left[m_{t+2}^i \mid \Omega_{t+1}^i\right] di \mid \Omega_t^j\right] + ... \\ &... + E\left[\int E\left[... \int E\left[m_{t+n}^{i'} \mid \Omega_{t+n-1}^{i'}\right] di' ...\right] di \mid \Omega_t^j\right] \\ &+\frac{1}{2}\sum_{s=0}^{n-1} Var\left(m_{t+1+s}^j + p_{t+1+s}^{n-1-s} \mid \Omega_{t+s}^j\right). \end{aligned} \tag{4}$$

To arrive at the expression (4) we also used the assumption that individual agents are price takers so that we replace the expectation of the next period price p_{t+1}^{n-1} by the average agent's expectation of what other agents will be willing to pay for the bond in the next period.[6] The price of an $n-$period bond can be expressed as a function of an agent's period t expectation of his period $t+1$ discount factor, plus the agent's expectation in period t of the average agent's period $t+1$ expectation about the stochastic discount factor in period $t+2$, and so on.

The expression (4) may appear quite abstract and without a clear economic interpretation. However, one way to think about what it means is to note that stochastic discount factors generally have both a time-discount and risk-adjustment component. The bond price in period t thus depends on the current risk-free rate and risk premia as well as on expectations about other agents' future expectations about risk-free interest rates and risk premia further into the future. When other agents are expected to believe in period $t+1$ that interest rates or risk-premia will be high in period $t+2$, and so on, bond prices in period t will be low, since agents then expect that the next period resale value of the bond will be low.

It is well known that higher-order expectations are distinct from first order expectations when agents have heterogeneous information (e.g. Allen, Morris and Shin 2006 and Bacchetta and van Wincoop 2006). Heterogeneous information then introduces what Allen, Morris and Shin (2006) call a "Keynesian beauty contest" into asset markets, where heterogeneously informed agents attempt to predict not what the fun-

[6]An implicit assumption here is that bond prices and stochastic discount factors are jointly log-normal processes. We also assume that conditional variances are deterministic and common across agents. The latter assumption is not crucial for any of the results, but helps simplifying the notation. The intermediate steps necessary to derive Eq (4) from Eq (3) are described in Appendix A.

damental value of a bond is, but how much other agents will be willing to pay for it at the next trading opportunity. To quantify the importance of this type of speculative behavior, we need to write down an explicit model for both bond prices and agents' stochastic discount factors.

C. The risk-free short rate

To operationalize the insights from above, we need to specify explicit functional forms for agents' stochastic discount factors. These stochastic discount factors will determine how much compensation agents require to hold a long term bond relative to the risk-free short rate. As in affine full-information models, the risk-free short rate r_t is an affine function of an exogenous state vector x_t

$$r_t = \delta_0 + \delta_x' x_t. \tag{5}$$

where the d-dimensional vector x_t follows a first order vector auto regression

$$x_{t+1} = \mu^{\mathbb{P}} + F^{\mathbb{P}} x_t + C\varepsilon_{t+1} : \varepsilon_{t+1} \sim N(0, I). \tag{6}$$

The short rate r_t makes up a common component in every agent's stochastic discount factor.

D. Agents' information sets

We introduce heterogeneous expectations by relaxing the assumption that agents can observe the factors x_t directly. Instead, agents observe the signal vector x_t^j, which is the sum of the true vector x_t and an idiosyncratic noise component

$$x_t^j = x_t + Q\eta_t^j : \eta_t^j \sim N(0, I) \tag{7}$$

where the noise shocks η_t^j are independent across agents. The vector x_t^j is the source of agent-specific information about the unobservable exogenous state x_t. The precision of the signals x_t^j is common across agents and determined by the matrix Q. The agents use these signals to form rational, or model consistent, expectations about future risk-free rates and risk premia. This set-up is a simple way to capture the fact that, in practice, it is too costly for agents to pay attention to all available information that could potentially help predict bond prices. With slightly different vantage points and historical experiences, agents instead tend to observe different subsets of all available information. Since the signals contain information about a common vector of latent factors, information sets will be highly correlated across agents, but not perfectly so. Formally, our set-up is similar to the information structure in Diamond and Verrecchia (1981), Admati (1985), Singleton (1987), Allen, Morris and Shin (2006) and Bacchetta and van Wincoop (2006).

In addition to the vector of private signals x_t^j agents also observe the short rate r_t and a vector of bond prices y_t of maturity up to $\overline{n}$

$$y_t \equiv \begin{bmatrix} -\frac{1}{2}p_t^2 & -\frac{1}{3}p_t^3 & \cdots & -\overline{n}^{-1}p_t^{\overline{n}} \end{bmatrix}'.$$

Agent j's information set in period t is thus defined by the filtration $\Omega_t^j = \left\{x_t^j, r_t, y_t, \Omega_{t-1}^j\right\}$.

E. The stochastic discount factor of agent j

Following the full-information affine term structure models as closely as possible, we specify (the logarithm of) agent j's stochastic discount factor as

$$m_{t+1}^j = -r_t - \frac{1}{2}\Lambda_t^{j\prime}\Sigma_a\Lambda_t^j - \Lambda_t^{j\prime}a_{t+1}^j : a_{t+1}^j \sim N(0, \Sigma_a). \tag{8}$$

where a_{t+1}^j is a vector of one-period-ahead bond price forecast errors conditional on agent j's information set. It is defined as

$$a_{t+1}^j \equiv \begin{bmatrix} p_{t+1}^1 - E\left[p_{t+1}^1 \mid \Omega_t^j\right] \\ \vdots \\ p_{t+1}^{\overline{n}-1} - E\left[p_{t+1}^{\overline{n}-1} \mid \Omega_t^j\right] \end{bmatrix} \tag{9}$$

where $\overline{n}$ is maximum maturity of a bond in the market. The vector a_{t+1}^j thus spans the risk associated with agent j's bond holdings. The term $\Lambda_t^{j\prime}a_{t+1}^j$ in (8) introduces covariance between conditional bond price risk and the stochastic discount factor and the vector Λ_t^j determines the extent to which this risk is priced.

There are two main differences between the stochastic discount factor (8) and its full information counterpart. First, the vector of risk prices Λ_t^j is agent specific. Second, the vector a_{t+1}^j of risks is conditional on an agent's information set. In this paper we relax the assumption that agents can observe the factors x_t directly. One source of bond price risk thus arises from the fact that agents cannot observe the current state perfectly. Below we will also introduce maturity specific shocks to bond prices that are akin to random supply shocks. These shocks also contribute to the risk associated with holding bonds. Unlike in full information models, the vector of risks a^j_{t+1} is then generally not spanned by the period $t+1$ innovations to the factors x_t.

F. Conjectured bond prices and state equation

The vector of risk-prices Λ^j_t varies over time and to complete the model, we need to also be explicit about its functional form. However, it is convenient to first conjecture (and later verify) functional forms for bond prices and the law of motion of the state. In the solved model, it is possible to verify that bond prices can be written in the form

$$p^n_t = A_n + B'_n X_t + e'_{n-1} v_t : v_t \sim N(0, I) \tag{10}$$

where v_t is a vector of maturity specific shocks of the same dimension as the number of different long-maturity bonds. The vector e_n has a one in the n^{th} position and zeros elsewhere. The vector e_{n-1} thus picks out the shock v^n_t that directly affect the price of an n period bond.

The maturity specific shocks are akin to supply shocks in equilibrium models. In models where agents solve an explicit portfolio problem, a positive supply shock decreases the price of a risky asset since a higher expected excess return is necessary to convince risk averse agents to absorb the additional supply into their portfolios (e.g. Admati 1985 and Singleton 1987). The higher expected excess return due to the increased supply is thus compensation that agents require for holding a riskier portfolio with a larger share of the risky asset. Greenwood and Vayanos (2014) provide empirical evidence suggesting that an increase in bond supply do in fact increase expected excess returns. The framework presented here is consistent with this interpretation of the maturity specific disturbances v^n_t, though we do not model the portfolio decisions of agents explicitly.[7] Below, we introduce the maturity specific shocks v^n_t through the specification of the stochastic discount factor. The maturity specific shocks v^n_t are thus not pricing errors and the bond prices described by (10) are arbitrage-free.

When bond prices do not reveal the state perfectly, agents may have heterogeneous information sets. We then need to include higher-order expectations about the exogenous factors x_t as state variables. The state

[7] In the Internet Appendix we demonstrate that the equilibrium model in Barillas and Nimark (2016) is a special case of the affine framework derived here in which supply shocks enter the equilibrium price in exactly the same way as the maturity specific shocks v^n_t in (10).

vector X_t in (10) thus contains a hierarchy of average higher-order expectations about the latent factors x_t

$$X_t \equiv \left[\begin{array}{cccc} x_t' & x_t^{(1)\prime} & \cdots & x_t^{(\overline{k})\prime} \end{array}\right]' \tag{11}$$

where the average k order expectation $x_t^{(k)}$ is defined as

$$x_t^{(k)} \equiv \int E\left[x_t^{(k-1)} \mid \Omega_t^j\right] dj. \tag{12}$$

The recursive definition (12) can be started from the convention that $x_t^{(0)} = x_t$. In the Appendix we demonstrate that X_t follows a first order vector autoregression

$$X_{t+1} = \mu_X + \mathcal{F}X_t + \mathcal{C}u_{t+1} : u_{t+1} \sim N(0, I) \tag{13}$$

where $u_{t+1} = [\varepsilon_{t+1}' \quad v_{t+1}']'$.[8]

G. *Signal precision and state dynamics*

The law of motion of the state X_t is endogenous and depends on the precision of agents' information. If the variances of the idiosyncratic noise shocks are zero, the signal vector x_t^j reveals the factors x_t perfectly. The higher-order expectations in the hierarchy X_t then coincide with the true factors x_t and each element of X_t follows the same law of motion (6) as the exogenous factors. However, in general, an innovation to x_t is partly attributed to idiosyncratic sources so that on average, agents under-react to innovations to the factors. The presence of the idiosyncratic shocks thus changes the responses of expectations and bond prices to innovations in x_t, even though they average to zero in the cross-section.

[8] It is perhaps worth pointing out here that even though the state vector is high dimensional, this by itself will not increase our degrees of freedom in terms of fitting bond yields. The fact that the endogenous state variables $x_t^{(k)}$ are rational expectations of the lower order expectations in $x_t^{(k-1)}$ disciplines the law of motion (13). As a consequence, the matrices $\mathcal{F}$ and $\mathcal{C}$ are completely pinned down by the parameters of the process governing the true exogenous factors x_t and the parameters that govern how precise agents' signals about x_t are. How to find the matrices $\mathcal{F}$ and $\mathcal{C}$ in practice is described in Appendix E.

H. Risk prices

The final piece needed to complete the model is to specify the functional form for the vector of risk prices Λ_t^j. It is an affine function of the agent-specific state X_t^j

$$\Lambda_t^j = \Lambda_0 + \Lambda_x X_t^j + \Lambda_v v_t^j \tag{14}$$

where the vector v_t^j is the agent's expectations about the maturity specific shocks, i.e. $v_t^j \equiv E\left[v_t \mid \Omega_t^j\right]$. The agent-specific state X_t^j consists of the agent j specific exogenous factors x_t^j as well as of agent j's expectations (up to order $\bar{k}$) about the latent common factors x_t

$$X_t^j \equiv \left[\begin{array}{cccc} x_t^{\prime j} & E\left[x_t' \mid \Omega_t^j\right] & \cdots & E\left[x_t^{\prime(\bar{k})} \mid \Omega_t^j\right] \end{array} \right]' \tag{15}$$

so that $X_t = \int X_t^j dj$. The vector X_t^j determines both agent j's required compensation for risk as well as his expectations about future bond prices. To solve the model, we use that the no-arbitrage condition (2) has to hold for each agent, including the average agent who is defined as the agent whose state coincides with X_t.[9] The vector of agent j's risk-prices thus only depends on variables that are measurable functions of his information set. The exact form of Λ_t^j is chosen so that bond prices in equilibrium can be expressed as in the conjectured form (10). In the empirical specification section below we describe how Λ_t^j can be parameterized parsimoniously.

I. The bond price recursions

Substituting the stochastic discount factor (8) and the conjectured bond price equation (10) into the no-arbitrage condition (3) implies that A_n and B_n in (10) must satisfy the recursions

$$A_{n+1} = -\delta_0 + A_n + B_n' \mu_X + \frac{1}{2} e_n' \Sigma_a e_n - e_n' \Sigma_a \Lambda_0 \tag{16}$$

and

$$B_{n+1}' = -\delta_X + B_n' \mathcal{F} H - e_n' \Sigma_a \widehat{\Lambda}_x \tag{17}$$

[9]When the cross-sectional distribution of expectations has positive density at its mean, there exists an "average agent" whose expectations about future bond prices coincide with the cross-sectional average expectation.

where Σ_a is the covariance matrix of the vector of bond price risks (9). The matrix H is the average expectations operator $H : \mathbb{R}^{d(\bar{k}+1)} \to \mathbb{R}^{d(\bar{k}+1)}$ defined so that

$$\begin{bmatrix} x_t^{(1)} \\ \vdots \\ x_t^{(\bar{k})} \\ \mathbf{0}_{d\times 1} \end{bmatrix} = H \begin{bmatrix} x_t \\ x_t^{(1)} \\ \vdots \\ x_t^{(\bar{k})} \end{bmatrix}. \tag{18}$$

The matrix H thus moves the hierarchy X_t one step up in orders of expectations and sets $x_t^{(k)} = \mathbf{0}$ for expectations of orders higher than $\bar{k}$. The matrix $\widehat{\Lambda}_x$ is a translation of Λ_x and defined as $\widehat{\Lambda}_x = \Lambda_x + B(I - H)$.[10] As in a full information set-up, the recursions (16) and (17) can be started from

$$A_1 = -\delta_0 \tag{19}$$

$$B_1 = -\delta_X' \tag{20}$$

where $p_t^1 = -r_t$. Appendix C describes in detail each step involved in deriving the recursions above.

Readers familiar with full information affine term structure models will recognize that the recursive expressions for A_n and B_n' above are completely analogous to the corresponding expressions in the standard full information model. Replacing Σ_a by CC', $\mathcal{F}H$ by $F^{\mathbb{P}}$ and e_{n-1} by B_{n-1}' delivers the standard expressions. The interpretation of the corresponding matrices are also the same. Both Σ_a and CC' are the covariance of the vector of risks that agents require compensation for. Similarly, both $\mathcal{F}H$ and $F^{\mathbb{P}}$ are matrices that translate the current state into an (average) expectation about the next period's state. Finally, both e_{n-1} and B_{n-1}' are vectors that translate innovations to the respective risk vector a_{t+1}^j and $C\varepsilon_{t+1}$ into innovations to bond prices. The main difference between our recursions and those of the standard model is that here, risks associated with holding bonds arise not only from innovations to the true factors x_t but also from current state uncertainty, future innovations to higher-order expectations and maturity specific shocks.

[10]This specification ensures that the model nests the standard full information model by Joslin, Singleton and Zhu (2011) as a special case where agents can observe x_t perfectly and the maturity specific shocks have zero variance.

J. *Higher-order expectations in the affine model*

The average expectations operator H together with the state transition matrix $\mathcal{F}$ can be used to compute higher-order expectations about future short interest rates and risk premia. That is, the matrix product $\mathcal{F}H$ moves expectations one step "up" in orders of expectations and one step forward in time. The period k order expectation in period t of the state in period $t+k$ can thus be computed as $(\mathcal{F}H)^k X_t$.By recursive forward substitution of (17), the row vector B'_{n+1} can be decomposed into a term that captures higher-order expectations about the short rate and a term that captures higher-order expectations about future risk premia

$$B^{'}_{n+1} = -\delta_X \sum_{s=0}^{n} (\mathcal{F}H)^s - e_{n-s}\Sigma_a \Lambda_x \sum_{s=0}^{n-1} (\mathcal{F}H)^s \,. \tag{21}$$

The first term on the right hand side of (21) thus captures the effect of higher-order expectations about future risk-free interest rates on the price of a bond. The second term captures higher-order expectations about future risk premia. The difference between higher-order expectations and agents' first order expectations depends on the precision of agents' information.

K. *Solving the model*

The equilibrium dynamics of bond prices are completely described by the system

$$X_{t+1} = \mu_X + \mathcal{F}X_t + \mathcal{C}u_{t+1} \tag{22}$$

$$p_t^n = A_n + B'_n X_t + e'_{n-1} v_t \tag{23}$$

Solving the model thus implies finding $\mathcal{F}, \mathcal{C}, A_n$ and B'_n. The law of motion of the state depends on the bond price equation through the filtering problem of the agents. The bond price equation in turn depends on the law of motion of the state. It is therefore necessary to solve for (22) and (23) simultaneously. Appendix E describes how to do so using the method proposed in Nimark (2017).

II. Speculation in the Affine Model

In Section I above we showed that the price of a bond can be expressed as a function of agents' higher-order expectations about future risk-free interest rates and future risk premia. Here we first define the speculative component in a bond's price as the difference between the actual price, which depends on higher-order expectations, and the counterfactual price a bond would have if these higher-order expectations coincided with the average first order expectation. We then prove formally that the speculative term has to be orthogonal to publicly available information such as bond prices. This characteristic feature of the speculative term makes it statistically distinct from the classical components of the yield curve due to (first-order) expectations about future risk-free interest rates and risk premia.

A. The counterfactual consensus price

We will define the speculative component in the price of a bond as the difference between the actual price p_t^n and the counterfactual "consensus" price $\overline{p}_t^n$. As in Allen, Morris and Shin (2006), the consensus price is the hypothetical price a bond would have if by chance, all agents' higher-order expectations about future discount rates coincided with the current first order expectations of the average agent (while holding conditional variances fixed).

To see why this is a natural way to define the speculative component, note that in the affine model presented above, the forecasting problem of predicting other agents' future expectations about bond prices can be reduced to forming expectations about the current state and other agents' expectations about the current state X_t. The state summarizes all information that is possible to know about future states, so perceived agreement about the current state implies perceived agreement about expected future states. This means that if, by chance, an individual agent's first and higher-order expectations about the state x_t coincided, the agent must also believe that other agents share his predictions about future bond prices. There will then be no speculative motive for trade, since the agent then believes that other agents will only be willing to pay as much for the bond in the future as he expects himself to be willing to pay, were he to hold on to the bond until maturity.

We can thus specify the counterfactual consensus price $\overline{p}_t^n$ as the price of an n period bond that would prevail if average first and higher-order expectations about the latent state x_t coincided. It can be computed as

$$\overline{p}_t^n = A_n + B_n^{'} \overline{H} X_t + v_t^n \tag{24}$$

where the matrix $\overline{H}$ is the *consensus operator* $\overline{H} : \mathbb{R}^{d(\overline{k}+1)} \rightarrow \mathbb{R}^{d(\overline{k}+1)}$ defined so that

$$\begin{bmatrix} x_t \\ x_t^{(1)} \\ \vdots \\ x_t^{(1)} \end{bmatrix} = \overline{H} \begin{bmatrix} x_t \\ x_t^{(1)} \\ \vdots \\ x_t^{(\overline{k})} \end{bmatrix}. \tag{25}$$

That is, $\overline{H}$ is a matrix that takes a hierarchy of expectations about x_t and equates higher-order expectations with the first order expectation.

B. *The speculative component in bond prices*

We can use $\overline{H}$ to decompose the current n period bond price into a component that depends only on the average first order expectation and a speculative component that is the difference between the actual price and the counterfactual consensus price $\overline{p}_t^n$. By adding and subtracting the consensus price (24) on the right hand side of the bond price equation (10) we get the expression

$$p_t^n = A_n + B_n^{'} \overline{H} X_t + \underbrace{B_n^{'} \left(I - \overline{H}\right) X_t}_{speculative\ term} + v_t^n \tag{26}$$

since

$$p_t^n - \overline{p}_t^n = B_n^{\prime} \left(I - \overline{H}\right) X_t \tag{27}$$

The price of an n-period bond can thus be written as a sum of common components and a simple expression capturing the difference between the actual price and the answer you would get if you asked the average agent what he thinks the price would be if all agents, by chance, had the same state estimate as he did (while holding conditional uncertainty constant).[11]

C. *The speculative component and public information*

The speculative component in bond yields has a special characteristic compared to classical risk premia and terms due to expectations about future risk free interest rates. Since it depends on individual agents

[11] Bacchetta and van Wincoop (2006) refers to the equivalent object in their model as the "higher order wedge".

predicting that other agents will either over- or underestimate future discount rates, it must be orthogonal to public information in real time. That is, it is not possible for individual agents to predict other agents' forecast errors using public information. This result is stated more formally in Proposition 1.

PROPOSITION 1: *The speculative term $p_t^n - \overline{p}_t^n$ is orthogonal to public information in real time, i.e.*

$$E\left([p_t^n - \overline{p}_t^n]\,\omega_t\right) = 0 : \forall\ \omega_t \in \Omega_t \tag{28}$$

where Ω_t is the public information set at time t defined as the intersection of agents' period t information sets

$$\Omega_t \equiv \bigcap_{j\in(0,1)} \Omega_t^j. \tag{29}$$

Proof. In Appendix B. □

The proof follows directly from taking expectations of (27) conditional on the public information set Ω_t and using that $\Omega_t \subseteq \Omega_{t+s}^j$ for every j and $s > 0$. The intuition is straightforward: By construction, $\left(I - \overline{H}\right) X_t$ in (27) is a vector of higher-order predictions errors, i.e. a vector of differences between first and higher-order expectations about the latent state x_t. Since it is not possible to predict other agents' errors using publicly available information, any linear function of $\left(I - \overline{H}\right) X_t$ must be orthogonal to public information in real time.

If signals are noisy, differences between agents' first and higher-order expectations about the current state translate into differences between first and higher-order expectations about future states. That is, if an individual agent believes that other agents have a different estimate of the current state than he does, then it is rational to believe that other agents will also have different expectations from himself in the future, unless future signals will reveal the state perfectly. It is also rational for an individual agent to expect other rational agents' expectations to be revised towards his own expectations in the future when more signals have been observed. That is, unlike first order expectations, second (and higher) order expectations are not martingales. This is so because an individual agent expects the beliefs of other rational agents to converge towards his expectations about the true state, as other agents revise their beliefs in response to new information.

It is the fact that differences between first and higher-order expectations are expected to be persistent and to be revised in a predictable direction that induces speculative behavior in the model. If, for instance, it was common knowledge that everybody would observe a perfect signal about the state in the next period, there would be no motive to speculate since it would also be common knowledge that all agents would share the same valuation of the asset in the next period.

D. A three-way decomposition of the yield curve

Below, we will quantify the importance for bond yield dynamics of the speculative component derived above. While much of the focus in this paper is on the speculative component itself, it is also of interest to investigate how allowing for heterogeneous information may change our estimates of the classical components of the yield curve, i.e. short rate expectations and risk premia. To compare the implied estimates of (first order) short rate expectations and risk premia from our model to those produced by a standard affine common information model, we need to decompose the non-speculative component in (26) further. What we want is a decomposition of the form

$$p_t^n = \underbrace{A_n^{rp} + B_n^{rp\prime} X_t + \ v_t^n}_{\text{classic risk premia}} \quad + \quad \underbrace{A_n^r + B_n^{r\prime} X_t}_{\text{short rate expectations}} \quad + \quad \underbrace{B_n' \left(I - \overline{H}\right) X_t}_{\text{speculative term}} . \tag{30}$$

The classic risk premia terms can be found by subtracting the average first order expectations about future short rate expectations from the non-speculative component in (26). This implies that the A_n^{rp} and B_n^{rp} in (30) are given by

$$A_n^{rp} = A_n - A_n^r, \;\; B_n^{rp\prime} = B_n' \overline{H} - B_n^{r\prime}.$$

Using the law of motion (13) and the short rate equation (5), A_n^r and B_n^r can be computed as

$$A_n^r = -n \left(\delta_0 + \delta_X \mu_X\right) - \delta_X \sum_{s=0}^{n-1} \mathcal{F}^s \mu_X, \quad B_n^{\prime r} = -\delta_X \sum_{s=0}^{n-1} \mathcal{F}^s H.$$

The first two terms in (30) thus corresponds to the classic terms of the yield curve decomposition in Cochrane and Piazzesi (2008) and Joslin, Singleton and Zhu (2011) and are independent of any discrepancy between first and higher-order expectations. In the limit with perfectly precise signals, the speculative term tends to zero since both first and higher-order expectations about x_t then coincide with the true factors. The two classical terms, together with the maturity specific shocks, would then determine bond yields completely.

III. Empirical specification

In order to make the model presented in Section II operational we will need to be specific about some of the details that up until this point have been presented at a more general level. Here, we describe how the factor processes are normalized and how the prices of risk can be parameterized when higher order expectations enter as state variables. In this section we also describe how the cross-sectional dispersion of the individual responses in the Survey of Professional Forecasters can be exploited in likelihood based estimation of the model's parameters.

There are two principles that guide how we parameterize the model. First, we want a parsimonious specification that economizes on the number of free parameters. Second, we want a specification that nests a standard model as a special case so that we can isolate the effect of relaxing the full information assumption without changing other aspects of the model.

A. *Exogenous factor dynamics and the risk-free interest rate*

The first choice to be made is to decide how many factors to include in the exogenous vector x_t. In the estimated specification, x_t is a three dimensional vector so that in the special case with perfectly informed agents and no maturity specific shocks, the model collapses to a standard three factor affine Gaussian no-arbitrage model. Since the factors are latent we need to normalize their law of motion. We follow Joslin, Singleton and Zhu (2011) and let the risk neutral dynamics of the factors follow a first order vector autoregressive process

$$x_{t+1} = \mu^{\mathbb{Q}} + F^{\mathbb{Q}} x_t + C\varepsilon_{t+1} \tag{31}$$

with the restrictions that $\mu^{\mathbb{Q}} = 0$ and that the matrix $F^{\mathbb{Q}}$ is diagonal with the factors ordered in descending degree of persistence under the risk neutral dynamics. Furthermore, C is restricted to be lower triangular. Finally, δ_x in the short rate equation (5) is a vector of ones. These restrictions ensure that all parameters are identified in the special case of perfectly informed agents and no maturity specific shocks.

B. *Parameterizing the prices of risk Λ_t^j*

The state vector X_t^j and the vector of bond markets risks a_{t+1}^j are both high dimensional. As a consequence, leaving Λ_0 and Λ_X in the specification (14) of risk prices Λ_t^j completely unrestricted would result in a very large number of free parameters. To avoid an over-parameterized model we therefore restrict Λ_0 and Λ_x as

$$\Lambda_0 = \Phi \begin{bmatrix} \lambda_0 \\ \mathbf{0} \end{bmatrix}, \quad \Lambda_x = \Phi \begin{bmatrix} \lambda_x & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} - B\left(I - H\right) \tag{32}$$

where λ_0 is a 3×1 vector and λ_x is a 3×3 matrix that contain the freely estimated parameters. The matrices Φ and B are defined in the Appendix and do not contain any free parameters. The matrix Φ is a rotation that ensures that the estimated λ_0 and λ_x captures the compensation for risk associated with innovations to the true factors x_t and implies that the model nests the standard specification if agents' signals are perfectly precise and the variances of the maturity specific shocks v_t^n are zero. The number of estimated parameters in Λ_t^j is thus the same as in the price of risk specification in a standard Gaussian full information three-factor model, e.g. Duffee (2002) and Joslin, Singleton and Zhu (2011).

C. Implied physical dynamics

The physical dynamics of the factors (6) are implicitly defined by the combination of the risk neutral dynamics (31) and the prices of risk vector (32). The vector $\mu^{\mathbb{P}}$ and the coefficient matrix $F^{\mathbb{P}}$ in (6) are given by

$$\mu^{\mathbb{P}} = \mu^{\mathbb{Q}} + CC'\lambda_0, \quad F^{\mathbb{P}} = F^{\mathbb{Q}} + CC'\lambda_x \tag{33}$$

In the limit with perfectly precise signals, the risk neutral and physical dynamics of the affine model have the usual interpretation: While the latent factors follow the physical dynamics, bonds can be priced as if agents were risk neutral and the factors followed the risk neutral dynamics. The physical dynamics then also completely determine the law of motion of the extended state X_t.

D. Free parameters to be estimated

The parameters that we estimate are the elements of the matrices $F^{\mathbb{Q}}$ and C which govern the processes of the latent factors x_t under the risk neutral measure, the diagonal matrix Q which specifies the standard deviation of the idiosyncratic noise in the agent-specific signals about x_t, the constant δ_0 in the risk-free short rate equation (5), σ_v the standard deviation of the maturity specific disturbances v_t^n (specified so that $\sqrt{var\left(v_t^n\right)} = n\sigma_v$, i.e. so that the standard deviation of the impact on *yields* is constant across maturities) and the vector λ_0 and matrix λ_x which govern risk premia. The model has 26 parameters in total and relative to a canonical full information three-factor affine model, the only additional parameters are the three diagonal elements of Q that govern the precision of the agent-specific signals.

E. Agents' information sets

Agent j observes the factors x_t^j as defined in (7) which is the source of agent j's heterogeneous information about the common factors x_t. Each agent also observes the risk-free short rate r_t. In addition to these exogenous signals, all agents can observe all bond yields up to maturity $\overline{n}$, where $\overline{n}$ is the largest maturity used in the estimation of the model. Here, the longest maturity yield that we will use in estimation is a 10 year bond implying that $\overline{n} = 40$ with quarterly data.

F. Choosing the maximum order of expectation $\overline{k}$

Nimark (2017) demonstrates that it is possible to accurately represent the equilibrium dynamics of a model with heterogeneously informed agents by a finite dimensional state vector, despite of the infinite regress of agents "forecasting the forecasts of others" (e.g. Townsend 1983). We use the method proposed there to solve the model. As part of that procedure we need to choose the maximum order of expectation considered. We denote that maximum order $\overline{k}$ and here we set $\overline{k} = 40$. This implies a substantial redundancy as most of the bond price dynamics are captured by expectations of order lower than five.

G. Estimating the model using bond yields and survey data

The parameters of the model can be estimated by likelihood based methods. We use quarterly data on the short rate and bond yields with two, five, seven and ten years to maturity with the sample spanning the period 1971:Q4 to 2011:Q4. The zero-coupon yield data is taken from the Gurkaynak, Sack and Wright (2007) data set available from the Federal Reserve Board. In addition to bond yields we also use one quarter ahead forecasts of the T-Bill rate and the 1 quarter ahead forecasts of the 10 year bond rate from the Survey of Professional Forecasters (SPF). In the model, the cross-sectional distribution of agents' one-period-ahead forecasts of the risk-free short rate is Gaussian with mean and variance given by

$$E\left[r_{t+1} \mid \Omega_t^j\right] \sim N\left(-A_1 - B_1'\mu_X - B_1'\mathcal{F}HX_t,\ \ B_1'\mathcal{F}\Sigma_j\mathcal{F}'B_1\right) \tag{34}$$

where Σ_j is the cross-sectional covariance of expectations about the current state, i.e.

$$\Sigma_j \equiv E\left[H\left(X_t^j - X_t\right)\left(X_t^j - X_t\right)' H'\right] \tag{35}$$

As econometricians, we can thus treat the individual survey responses of T-Bill rate forecasts as measures of the average expectation of the short rate r_t where the variance of the individual responses' deviations from the average is determined by the model implied cross-sectional variance of short rate expectations.[12] The corresponding distribution for the one-period-ahead forecast of the 10 year yield is

$$E\left[y_{t+1}^{40} \mid \Omega_t^j\right] \sim N\left(-\frac{1}{40}A_{40} - \frac{1}{40}B_{40}'\mu_X - \frac{1}{40}B_{40}'\mathcal{F}HX_t,\ \frac{1}{40}B_{40}'\mathcal{F}\Sigma_j\mathcal{F}'B_{40}\frac{1}{40}\right) \tag{36}$$

The deviations of individual agents' forecasts from the average forecasts are caused by idiosyncratic shocks that are independent across agents. The individual survey responses are collected in the vectors $\mathbf{y}_{t+1|t}^1$ and $\mathbf{y}_{t+1|t}^{40}$. The cross-sectional covariance of $\mathbf{y}_{t+1|t}^1$ and $\mathbf{y}_{t+1|t}^{40}$ can thus be specified as the scalars $B_1'\mathcal{F}\Sigma_j\mathcal{F}'B_1$ and $\frac{1}{40}B_{40}'\mathcal{F}\Sigma_j\mathcal{F}'B_{40}\frac{1}{40}$ multiplied by an identity matrix.

H. The estimated state space system

Given the model and the data, we use the Kalman filter to evaluate the log likelihood function for the state space system

$$X_t = \mu_X + \mathcal{F}X_{t-1} + \mathcal{C}\overline{u}_t : \overline{u}_t \sim N(0, I) \tag{37}$$

$$\overline{z}_t = \overline{\mu}_t^z + \overline{D}_t X_t + \overline{R}_t \overline{u}_t \tag{38}$$

where $\overline{z}_t$ is the vector of observables

$$\overline{z}_t = \begin{bmatrix} r_t & y_t^8 & y_t^{20} & y_t^{28} & y_t^{40} & \mathbf{y}_{t+1|t}^{1\prime} & \mathbf{y}_{t+1|t}^{40\prime} \end{bmatrix}' \tag{39}$$

The vector $\overline{\mu}_t^z$ and the matrices $\bar{D}_t$ and $\bar{R}_t$ in the measurement equation (38) are defined in Appendix F.

The number of survey responses varies over time and surveys are not available at all for the period before 1981:Q3. Therefore, the dimensions of $\overline{\mu}_t^z$, $\bar{D}_t$ and $\bar{R}_t$ are also time-varying. This fact may influence

[12] Appendix G contains details of how to compute the cross-sectional variance Σ_j in practice.

the precision of our estimates of the state, i.e. we will have more precise estimates of the latent state X_t when there is a large number of survey responses available. Using individual survey responses and likelihood based methods also naturally incorporates the feature that we have more precise information about the cross-sectional average expectations of agents when there are 50 responses (the sample maximum) compared to when there are only 9 responses (the sample minimum). This information is lost when using measures of central tendency of the survey responses, such as the mean or median forecast.[13]

I. Estimation procedure

The next section will present empirical results based on the heterogeneous information model described above as well as the nested full information model without maturity specific shocks of Joslin *et al* (2011). To obtain parameter estimates for the full information model we first estimate a model without using survey data following the procedure in Joslin *et al* (2011) which reliably finds the maximum of the likelihood in a model estimated using only yields. Subsequently, we proceed to estimate the same full information model with both survey data and bond yields, taking the estimates from the yields-only estimates as starting values. To incorporate the surveys into the full information model we treat the individual responses as noisy measures of the model implied common expectation about future yields as in Kim and Orphanides (2005) and Chernov and Muller (2012). We thus also need to estimate the variance of the measurement errors in the survey data. The maximum likelihood estimates of the surveys plus yields model are found by first using a numerical optimizer and then the Metropolis-Hastings algorithm.[14] Importantly, in the full information model, there is no interaction between the cross-sectional dispersion of the survey data and the estimated dynamics of the factors. This is in contrast to the heterogeneous information model. There, the dynamics of the factors must not result in too informative bond prices since the model implied dispersion would then be inconsistent with the observed dispersion of survey forecasts.

The parameters of the heterogeneous information model are estimated using Bayesian methods with (improper) uniform priors so that the posterior mode coincides with the maximum likelihood estimates

[13]The SPF data is publicly available, with only a short delay between collection and release, and ideally, we would like to treat it as being observable also to the agents in the model. Unfortunately, this is not feasible using existing model solution techniques and the main difficulty arises because allowing agents to observe the individual survey data would make their information sets time varying. In the Online Appendix, we discuss the implications of this assumption and present empirical results based on an alternative estimation method using only information in the average dispersion of the survey responses. This strategy does not presume that we as econometricians have access to information unavailable to the agents in the model, but yields very similar results.

[14]We experimented with a number of alternative starting values and optimization routines. Of these alternatives, the procedure reported here resulted in the highest posterior likelihood.

and the posterior density is proportional to the likelihood function. We take 1,000,000 draws using the Metropolis-Hastings algorithm (e.g. Geweke 2005). The modes and the probability intervals presented in the next section are based on the last 500,000 draws.

IV. Empirical Results

This section contains the main empirical results of the paper. Here, we first present the parameter estimates and discuss how these are influenced by the fact that individual survey responses are used in estimation. This is followed by a decomposition of historical bond yields into risk premia, first order expectations about the risk-free short interest rate and a speculative component driven by differences between first and higher order expectations. In this section we also compare estimates of historical risk premia and short rate expectations from our heterogeneous information model with estimates of the same quantities from the nested full information model of Joslin, Singleton and Zhu (2011). The section ends by quantifying the total effect of information imperfections and by an assessment of how useful the agent-specific signals are to the agents for predicting excess returns.

A. Parameter estimates and the dispersion of survey responses

Table 1 presents the posterior modes along with 95% probability intervals (in brackets). Since the factors are latent most of the estimated parameter values are of no particular interest when viewed in isolation. However, using the individual survey responses in estimation has interesting implications for those parameters of the model that govern how informed agents are about the latent factors.

Using the full cross-section of surveys to estimate the parameters of the model clearly influences the estimates of the parameters that directly determine the precision of the agent-specific signals. However, the relationship between the precision of the agent-specific signals and the model implied dispersion is non-monotonic. When the agent-specific signals are very precise, the cross-sectional dispersion is close to zero and the dynamics of bond yields will be close to those of the full information model. When the agent-specific signals are very imprecise, agents attach little weight to them, and again, the cross-sectional dispersion will be close to zero. To match the substantial dispersion observed in the survey responses, intermediate values for the parameters that govern the precision of the agent-specific signals are required. The estimates of $Q_{1,1}$, $Q_{2,2}$ and $Q_{3,3}$ that govern the precision of the private signals thus cannot be neither too large, nor too small relative to the estimates of $C_{1,1}$, $C_{2,2}$ and $C_{3,3}$ that govern the standard deviation of the innovations to the latent factors. At the posterior mode, the standard deviation of the idiosyncratic noise in the agent-specific signals are between 1 to 4 times as large as the innovations to the true factors.

Less obviously, the cross-sectional dispersion in surveys will also discipline the estimated dynamics of bond prices. In our model, depending on the parameters, bond prices may or may not reveal the state perfectly. If bond prices are too revealing about the latent exogenous factors x_t, the cross-sectional dispersion will be too low relative to the dispersion in the survey data, regardless of the precision of the agent-specific signals. How informative bond yields are depends on how different the persistence of each factor is under the risk-neutral dynamics. The intuition is straightforward: The risk neutral dynamics determine how each factor affect the yield across different maturities. Factors with low risk-neutral persistence only affect the short end of the yield curve, factors with a high risk-neutral persistence affect the entire yield curve. If there is only one factor that is persistent enough under the risk neutral dynamics to move the long end of the yield cure, a change in long maturity yields can then only be caused by a change in the high persistence factor. Similarly, if there is only one factor with very low persistence under the risk neutral measure, a change that is exclusive to the short end of the yield curve can only be caused by a change to the low persistence factor. More generally, observing bond yields will be very informative about the latent factors if each factor has a very different implication for the shape of the yield curve.

If simply observing the yield curve would be enough to get very precise estimates of the latent factors, agents would put little or no weight on their private signals. The cross-sectional dispersion of expectations would then be too concentrated relative to the survey data. Estimating the model using the full cross-section of survey responses thus imposes strong restrictions on the risk-neutral dynamics and these restrictions are binding in practice: The posterior estimates of the second and third eigenvalue of the factor process under the risk neutral dynamics, i.e. $F^{\mathbb{Q}}_{2,2}$ and $F^{\mathbb{Q}}_{3,3}$, are very similar at *0.927* and *0.924,* respectively. This means that it is virtually impossible for the agents in the model to disentangle the individual effects of the second and third factor on the yield curve, even when the variances of the maturity specific shocks are small. While the second and third factor have very similar effects on the current yield curve, they have very different implications for future bond yields as is evidenced by the distinct eigenvalues of the matrix $F^{\mathbb{P}}$ that governs the factors' physical dynamics. Disentangling the effect from the second and third factors is thus important for agents in order to make accurate predictions about future bond returns.

Table I

Posterior Parameter Estimates

Physical Factor Dynamics

$$F^{\mathbb{P}} = \begin{bmatrix} \underset{(0.950-0.971)}{0.967} & \underset{(-0.003--0.0002)}{-0.001} & \underset{(-0.015--0.001)}{-0.011} \\ \underset{(0.001-0.003)}{0.002} & \underset{(0.870-0.896)}{0.873} & \underset{(-0.020--0.015)}{-0.018} \\ \underset{(-0.002--0.0002)}{-0.001} & \underset{(-0.11-0.00)}{-0.008} & \underset{(0.658-0.673)}{0.670} \end{bmatrix}$$

$$C = \begin{bmatrix} \underset{(0.006-0.014)}{0.012} & 0 & 0 \\ \underset{(0.0015-0.0021)}{0.0018} & \underset{(0.0007-0.0010)}{0.0009} & 0 \\ \underset{(-0.0011-0.0015)}{-0.0006} & \underset{(-0.0002-0.0074)}{0.0056} & \underset{(0.015-0.022)}{0.021} \end{bmatrix}$$

$$\mu_X^{\mathbb{P}} = \begin{bmatrix} \underset{(0.00-0.0003)}{0.0002} \\ \underset{(-0.0069--0.0043)}{-0.0055} \\ \underset{(-0.0034--0.0013)}{-0.0027} \end{bmatrix}$$

Risk Neutral Factor Dynamics

$$F^{\mathbb{Q}} = \begin{bmatrix} \underset{(0.995-1.004)}{0.997} & 0 & 0 \\ 0 & \underset{(0.924-0.944)}{0.927} & 0 \\ 0 & 0 & \underset{(0.921-0.944)}{0.924} \end{bmatrix}$$

Short Rate Constant *Maturity Specific disturbances*

$$\delta_0 = \underset{(0.057-0.11)}{0.09} \qquad \sigma = \underset{(0.43-0.53)}{0.0051}$$

Private signal noise s.d.

$$Q = \begin{bmatrix} \underset{(0.027-0.029)}{0.028} & 0 & 0 \\ 0 & \underset{(0.003-0.005)}{0.004} & 0 \\ 0 & 0 & \underset{(0.014-0.023)}{0.022} \end{bmatrix}$$

Numbers in round brackets indicate 95% probability intervals.

The restrictions on the risk neutral dynamics that ensures that bond prices are not too revealing are similar to those that imposes that some factors are unspanned, e.g. Joslin, Priebsch and Singleton (2014), Duffee (2011) and Barillas (2013). An unspanned factor is by definition not priced, i.e. does not affect current bond yields. The flip side of this definition is that an unspanned factor cannot be extracted from the yield curve by inverting the bond price function. Yet, Joslin *et al* (2014) and Duffee (2011) demonstrate that unspanned factors can help forecast future interest rates even after conditioning on the current yield curve. The restriction imposed on our model by fitting the cross-section of survey forecasts is thus similar to imposing an unspanned factor structure: Only parameterizations that ensures that the state is not an invertible function of bond yields will leave room for the agent-specific signals to play a role. Parameterizations that imply that the state is almost perfectly revealed by bond yields will be rejected by the data.

An alternative strategy to use the survey data is to treat individual responses as noisy measures of a common expectation held by all agents as in Kim and Orphanides (2005) and Chernov and Mueller (2012). This is also the strategy that we follow when estimating the full information model with survey data. Others have used a measure of central tendency from the surveys, like a mean or median, to represent a noisy measure of the expectations of a representative agent (see Piazzesi and Schneider (2011)). Because these papers by construction feature a degenerate distribution of expectations, they cannot exploit the information in the cross-sectional dispersion of survey forecast when estimating their models. In contrast, because agents in our model do have heterogeneous expectations, we can use the information contained in the cross-sectional forecast dispersion to discipline the parameters of our model.

B. Model fit

The model does a good job of fitting the cross-sectional dispersion in the survey data. The model implied dispersion of the one-quarter-ahead forecasts of the short interest rate has a cross-sectional standard deviation of 26 basis points, compared to the 40 basis points sample average in the survey data. The model implied dispersion of the one-quarter-ahead forecasts of the 10 year yield is 29 basis points versus 40 basis points in the data. That the posterior dispersion is somewhat lower than in the survey data suggests there is a trade-off between fitting bond yields and the cross-sectional dispersion. One possible explanation is that traders in reality are better and more uniformly informed than survey respondents and this may be inferable from bond yield dynamics.

The model also provides a good unconditional fit of bond yields. The model implied unconditional yields are within a few basis points of the sample averages. The standard deviation of the maturity specific shocks is 51 basis points. In the heterogeneous information model, these shocks are priced factors and thus not pricing errors. The maturity specific shocks are different from traditional pricing errors not only because they are priced factors. To see how, note that agents' observe bond yields and use that information to form expectations about the state. An innovation to v_t^n therefore does not affect only y_t^n but also indirectly affects bond yields of maturities other than n. Due to the persistence in agents' estimates of the state, the effect on bond yields of a single maturity specific shock will also last for several periods after impact. This means that our model is not subject to the critique in Hamilton and Wu (2011) who argue that the independent white noise assumption of classical pricing errors is testable and rejected by the data in standard affine term structure models. Thinking about the maturity specific shocks as pricing or measurement errors is thus not correct and their standard deviation should not be used to judge the fit of the model. However, to understand their relative importance for explaining the observed variation in bond yields, we can note that their standard deviation is comparable to the 38 basis points standard deviation of the measurement errors in the full information model estimated using the same yields and survey data.

C. Historical decompositions

We can use the estimated model to measure how large the speculative term has been historically. From Proposition 1 we know that the speculative term must be orthogonal to public information available to all agents in real time, such as bond prices. However, as econometricians we have access to the full sample of data and can use information from period $t+1, t+2, ..., T$, together with the survey data to form an estimate of the speculative term in period t. The Kalman smoother (see for instance Durbin and Koopman 2002) can be used to back out an estimate of the state X_t conditional on the entire history of observables. Since the speculative components of the term structure are linear functions of the state, the smoothed state history $E\left[X^T \mid \overline{z}^T\right]$ can be used to perform the decomposition of historical bond yields as described by (30).[15]

[15]The smoothed estimates of the state X^T are plotted in the Appendix.

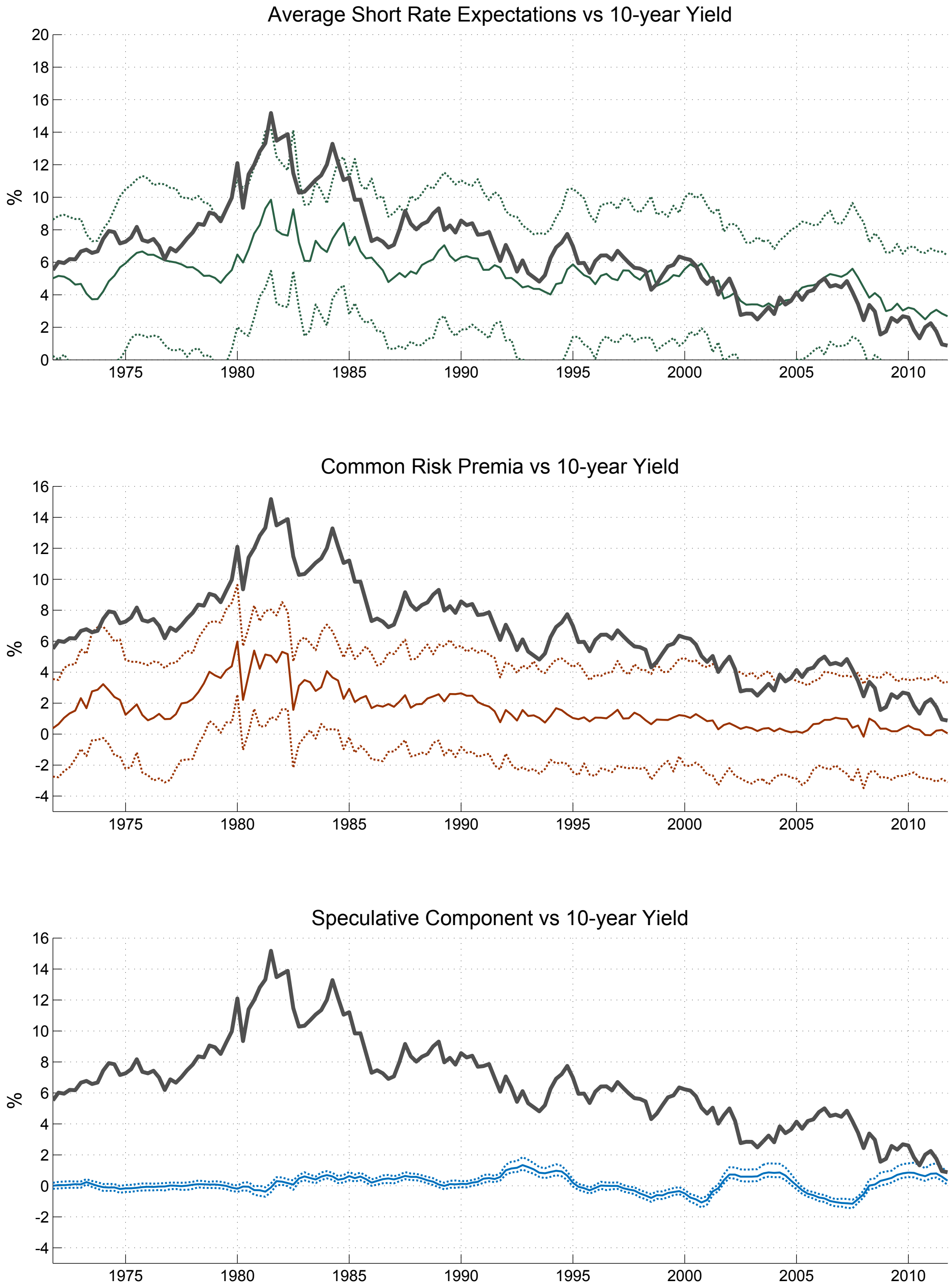

Figure 1 Decomposition of the 10-year yield. This figure plots the 10-year yield along with estimates of average short rate expectations, common risk premia and the speculative component at the posterior mode. The sample is 1971Q2 to 2011Q4. Dotted lines indicate 95% posterior probability intervals.

Figure 1 plots the history of the 10 year yield together with a decomposition, splitting the yield into the terms based on average expectations about future short rates, classical risk premia and the speculative term. Most of the variation in yields is driven by variation in average first order expectations about the short rate. The standard deviation of classical risk premia is 121 basis points, which is about 2 1/2 times the speculative term's standard deviation of 48 basis points.

In absolute terms, the speculative component is largest in the early 1990s when it accounts for about 125 basis points (out of 6 per cent) of the 10 year yield. The speculative component's contribution as a fraction of the total yield has been most important in the low yield environment of the last decade, reducing the 10 year yield by about 100 basis points at a time when the 10 year yield was on around 3 percent. This period coincides with the period when former Federal Reserve Chairman Alan Greenspan described the low long yields as a "conundrum" (see Federal Reserve Board 2005) . We can see in Figure 2 that just before Greenspan's remark in February 2005, the speculative term in the 10 year yield decreased sharply from a positive 75 to a negative 100 basis points. The speculative term thus contributed substantially to the observation that inspired Greenspan's remark.

Speculative dynamics are present at all maturities $n > 2$, but are quantitatively more important in medium- to long-maturity bonds. This is illustrated in Figure 2, where the estimated speculative components in the 1- 5- and 10-year yields are plotted. The speculative components are almost perfectly correlated across maturities and thus appear to have a one-factor structure. The speculative term is most volatile in the 10-year bond yield, but only marginally more so than for the 5-year bond. However, since long bond yields are less volatile than shorter bond yields, the speculative term contributes to a larger fraction of the variance of 10 year yields than to the variance of 5 year yields. The standard deviation of the speculative term in the 1-year bond is substantially lower than for medium to long maturity yields and it never accounts for more than 25 basis points of the 1 year yield in the sample.

The probability intervals of the decomposition in Figure 1 are much narrower around the speculative term than around the terms capturing expectations about future risk-free rates and risk premia. This is mostly likely the case because the speculative term is statistically distinct from the two classic terms by virtue of being orthogonal to public information in real time. Expectations about future risk free rates and risk premia can both be predicted based on public information and are in a statistical sense substitutes. It is thus more difficult to determine wether yields are high because expectations of future risk-free rates are high or because risk premia are high.

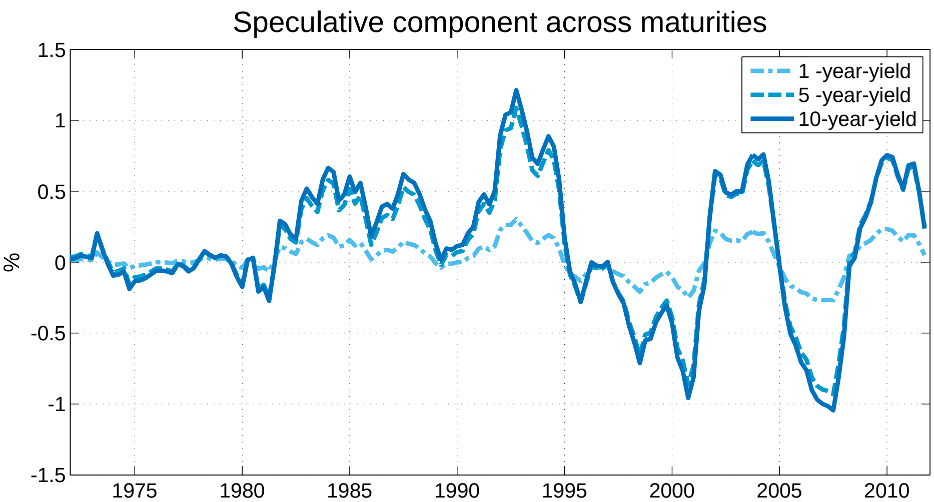

Figure 2 Speculative component across maturities. This figure plots estimates of the speculative component in the 1- 5- and 10-year yield at the posterior mode. The sample is 1971Q2 to 2011Q4.

D. Speculation and the expectations hypothesis

The result that the speculative component is quantitatively important for explaining bond yields relates in two different ways to the expectations hypothesis of the term structure of interest rates. One way to think of the expectations hypothesis is in terms of a 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 contributes to explaining the failure of the expectations hypothesis since it provides a second wedge, in addition to classical risk-premia, between bond yields and (first order) expectations about future short rates.

A second way to think about the expectations hypothesis is in terms of the predictability of excess returns. As proved in Proposition 1, the speculative term must be orthogonal to public information available to all agents in real time and can thus not help explaining the well-documented empirical regularity that future excess returns are predictable based on the current yield curve (as well as many other variables).

E. Comparison to a full information model

Gaussian affine term structure models have been used, for instance by Cochrane and Piazzesi (2005) and Joslin, Priebsch and Singleton (2014), to decompose the term structure into risk premia and expected future short rates. Allowing for heterogeneous information may potentially change our estimates of historical risk premia and short rate expectations. In the bottom panel of Figure 3 we have plotted the posterior estimate of the risk premia in the 10-year bond yield extracted using our model with heterogeneously informed agents together with the risk premia extracted using the full information model of Joslin *et al* (2011), which our model nests as a special case. The full information model is estimated using the same bond yields (i.e. 3 months, 2, 5, 7 and 10 year) and the same individual survey responses about the one quarter ahead forecasts for the T-Bill 3 month yield and 10 year bond yield as observables. Since the full information model implies that all agents share the same expectations about future bond yields, we treat the survey data as noisy measures of a common expectation held by all agents as in Kim and Orphanides (2005) and Chernov and Mueller (2012) when we estimate the full information model.

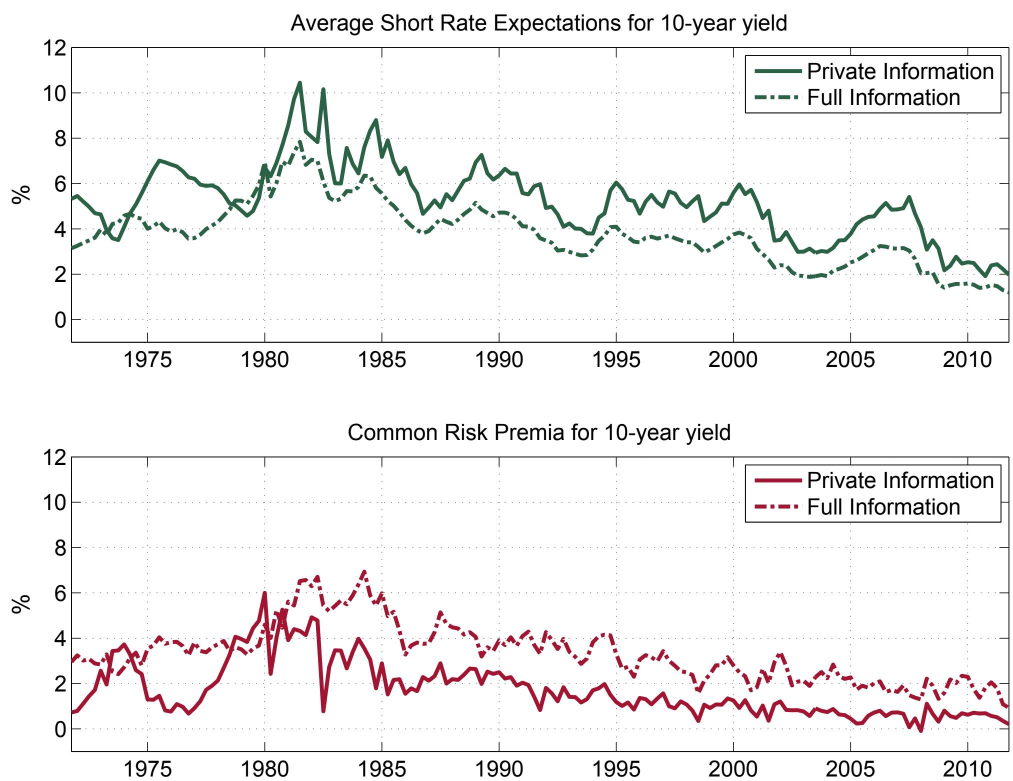

Figure 3 Average short rate expectations and common risk premia for the 10-year yield in the private and full information model This figure plots estimates of average short rate expectations and risk premia for the 10-year yield for the full information model and the private information model. Both estimates are obtained at the posterior mode (MLE estimates). The sample is 1971Q2 to 2011Q4.

The standard deviations of the risk premia are similar at 126 basis points in the full information model and 121 basis points in the heterogeneous information model. However, allowing for heterogeneous information changes the cyclical properties of risk premia. The correlation between the common risk premia term in the heterogeneous information model and the risk premia extracted using the full information model is 0.72 and the two time series diverge substantially at several occasions. For instance, the full information model and the heterogeneous information model provide very different interpretations of the Volcker disinflation in the early 1980s. Inspecting the two panels of Figure 3 reveals that the relative to the full information model, the heterogeneous information model attributes much less of the high and volatile long maturity yields in the late 1970s and early 1980s to high risk premia and much more to expectations about future risk-free interest rates. The structural origins of risk premia in our model must therefore be quite different from those that could explain risk premia in the full information model.

F. What drives speculative dynamics?

In order to address the question of what drives speculative dynamics we can decompose the variance of the speculative terms into four orthogonal sources: The three innovations to the exogenous factors in x_t and the maturity specific disturbances v_t. Table II displays variance decompositions of the 1-, 5- and 10 year yields, the speculative components in the 1-, 5- and 10 year yields and the first three principal components.

Shocks to the first factor explain between 80 and 86 per cent of the variance of yields and explain 85 percent of the variance of the first and third (i.e. the level and the curvature) principal components. Shocks to the first factor also explain about half the variance of the speculative component in 1 and 5 year bond yields and about a quarter of the variance of 10-year yields.

Shocks to the second factor explains less than 3 per cent of the variance of any bond yield or principal component, but explains between 10 and 21 per cent of the variance of the speculative terms. Shocks to the third factor explains 75 per cent of the variance of the second (slope) principal component and a substantial fraction of the variance of the speculative terms. Interestingly, the maturity specific shocks, which explain less than 2 per cent of the variance of any bond yield, explain between 23 and 32 per cent of the variance of the speculative term.

It would be interesting to rotate the model into an equivalent representation where the states are the principal components and analyze how each of the principal components affect the speculative term at different maturities. However, except in the full information limit, no such equivalent representation exists. Since the principal components are directly observable from the cross-section of yields, such a representation could not capture the speculative dynamics since by Proposition 1 these must be orthogonal to current bond

Table II **Variance decomposition of yields, speculative components and principal component of yields**

This table reports results of a variance decomposition of yields, speculative components and principal components of yields. These were computed at the posterior mode (MLE estimates) of our model. The sample period is from 1971:Q2 to 2011Q4. In the first three columns we show the percentage of the variance of the 1-, 5- and 10-year yield that is explained by the shocks of the model. The next three columns show related quantities for the 1-, 5- and 10-year speculative component of yields. The last three columns report the results for the first three principal components of yields.

	Yields			Speculative Component			Principal Components		
	$y_t^{(1)}$	$y_t^{(5)}$	$y_t^{(10)}$	$spec_t^{(1)}$	$spec_t^{(5)}$	$spec_t^{(10)}$	pc_t^1	pc_t^2	pc_t^3
ε_t^1	80.7	84.9	86.2	54.5	48.9	28.7	84.8	13.8	85.0
ε_t^2	2.7	2.3	2.1	10.6	10.9	21.1	2.4	2.7	2.1
ε_t^3	16.0	11.8	10.2	11.7	15.3	17.9	12.7	74.6	3.8
v_t	0.6	1.0	1.4	23.2	24.9	32.3	0.1	8.9	9.1

yields.

G. *How useful are the agent-specific signals for predicting excess returns?*

We can use the estimated model to quantify how useful the agent-specific signals are in terms of helping agents to forecast future bond prices and excess returns. For each maturity n we first compute the model implied unconditional variance of quarterly excess returns given by

$$var(rx_t^n) = E\left[(p_{t+1}^{n-1} - p_t^n - r_t) - (A_{n-1} - A_n - \delta_0) - (B_{n-1} - B_n - \delta_X)\,\mu_X\right]^2. \tag{40}$$

We then compute the variance of expected excess returns by conditioning on the information set of agent j. Dividing the latter by the former gives the model implied R^2 of excess returns for an n-period bond from the perspective of an agent in the model.

The R^2 of excess returns from the agents' perspective is U-shaped in maturities. It is largest for very short and very long maturities at around 0.18 and 0.16 respectively. The R^2 is smallest for 1 year bonds at about 0.14. These estimates of the model implied R^2 are somewhat smaller than suggested by simple predictability regressions on yields only, but larger than the R^2 of the Sharpe-ratio constrained affine models in Duffee (2010).

Comparing the predictability of excess returns from the agents' perspective with the R^2 of excess returns conditional only on bond yields suggests that the agent-specific signals increase the R^2 of excess returns by approximately 3 per cent, more or less uniformly across maturities.

These results are related to the findings in Ludvigson and Ng (2009) and Joslin, Priebsch and Singleton (2014). These papers document that macro factors are useful for predicting excess returns after conditioning on (principal components of) bond yields. One important difference is that our result does not imply that we as econometricians will necessarily do better in terms of predicting bond yields by using a model that allow for heterogeneously informed agents: The result above only states that the agents *inside the model* do better than an outside econometrician conditioning only on bond yields would.

H. Total effect of information imperfections

The speculative term quantified above is a function of perceived disagreement between individuals' first and higher order expectations. The speculative term would thus be identically zero if all agents could observe the state perfectly. Yet, the speculative term does not capture the total effect of information imperfections in the model. Independently of the speculative term, agents may have incorrect first order expectations about the latent factors.

We can use the estimated heterogeneous information model to quantify the historical importance for bond yields of inaccurate first order expectations. To this end, define the counterfactual *full information price* p_t^{n*} as the price that would prevail if agents' first and higher order expectations coincided with the true factors, that is, if $x_t^{(k)} = x_t$ for every k. The total effect of information imperfections is then captured by the difference between the actual price and the counterfactual full information price. In Figure 4 we have plotted the difference between the actual 10 year bond yield and the counterfactual full information bond yield, i.e. $n^{-1}(p_t^n - p_t^{n*})$ alongside the speculative component.

The estimated total effect of information imperfections on historical bond yields is quite large, at times accounting for up to 3 1/2 percentage points of 10 year bond yields. It is almost perfectly correlated with the speculative component (which makes up part of the total effect). The difference between the total effect and the speculative component is due to agents' incorrect first order expectations about the latent factors. With a historical standard deviation of 139 basis points, the total effect of imperfect information accounts for more of the historical variation in bond yields than classical risk premia.

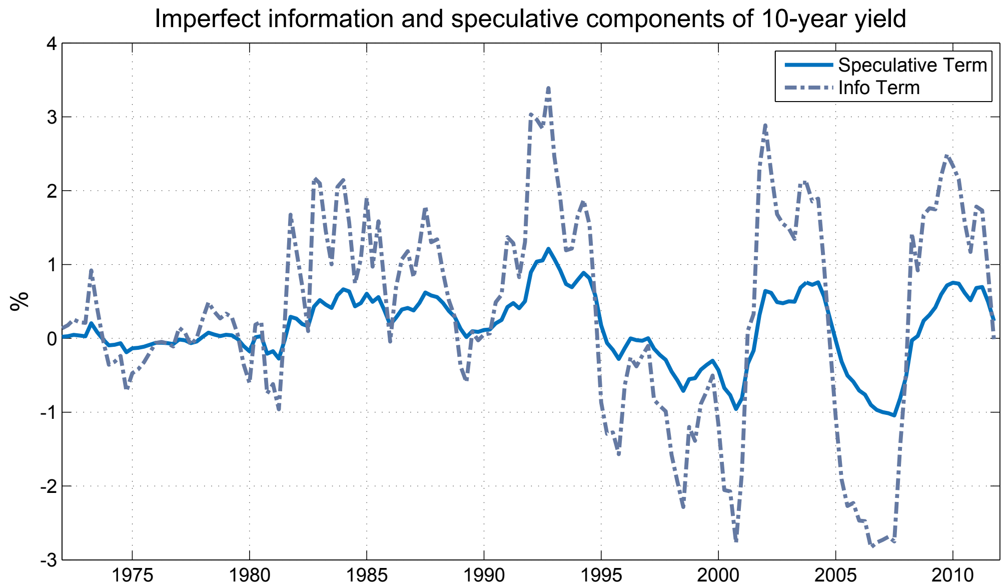

Figure 4 Total effect of information imperfections and the speculative component for the 10-year yield This figure plots the speculative component for the 10-year yield along with the counterfactual yield (Info Term) that would prevail if all order of expectations coincided with the true factors. Estimates are obtained at the posterior mode (MLE estimates). The sample is 1971Q2 to 2011Q4.

V. Conclusions

In this paper we have developed a flexible no-arbitrage framework for empirically quantifying the importance of information heterogeneity in bond markets. We have showed that allowing for heterogeneous information introduces a speculative component in bond prices that has been overlooked by the standard empirical term structure literature. The speculative component explain a substantial fraction of the variation in historical bond yields and we proved formally that the it in must be orthogonal to public information in real time. This makes it statistically distinct from the traditional yield components due to risk-premia and expectations about future risk free rates. In addition, the model is also consistent with the observed dispersion in survey forecasts.

Allowing for heterogeneous information changes our view on the relative historical importance of risk premia and expectations about risk-free short rates. Compared to a standard model, the heterogeneous information model attributes less of the high bond yields during the early 1980s to high risk premia and it attributes more to high expectations about future risk-free short rates. Affine no-arbitrage models like the one presented here impose a minimal structure on the data and thus are not well-suited for interpreting these findings in terms of economic decisions. However, the results presented here should guide our search for more structural models that could help us gain a more complete understanding of the documented phenomena. We should thus perhaps look for models that can explain why investor's interest rate expectations were unusually high in this period rather than look for reasons why investors required more compensation for risk than usual. One step in this direction is provided by Struby (2017) who builds on our framework by estimating a model where agents have heterogeneous information about aggregate macro economic variables such as inflation and output.

Finally, the stochastic discount factor framework presented here is general and can also be used to price other asset classes. In the present paper, we found that speculative dynamics were quantitatively important in treasury markets even though the value at maturity of a zero-coupon default-free bond is known with certainty, and the only source of uncertainty is future discount rates. The prices of other classes of assets such as stocks and corporate bonds also depend on expectations about future discount rates, but are subject to additional sources of uncertainty due to stochastic cash-flows and the probability of default. It seems plausible that speculative dynamics could be even more important in those asset classes where prices depend on a richer set of variables.

References

Admati, Anit R., (1985), "A noisy rational expectations equilibrium for multi-asset securities markets", *Econometrica*, Vol. 53, pp. 629-657.

Allen, F. S. Morris and H.S. Shin, (2006), "Beauty Contests and Iterated Expectations in Asset Markets", *Review of Financial Studies*, 19, pp719 – 752.

Bacchetta, Phillippe and Eric van Wincoop, (2006) "Can Information Heterogeneity Explain the Exchange Rate Determination Puzzle?", *American Economic Review* vol 96, pp552-576.

Barillas, F (2013) "Can we Exploit Predictability in Bond Markets?", Working Paper, Emory University.

Barillas, F and K. Nimark, (2016), "Speculation and the Term Structure of Interest Rates", forthcoming, *Review of Financial Studies.*

Chernov, M and Mueller, P (2012), "The Term Structure of Inflation Expectations", forthcoming in *Journal of Financial Economics*.

Christiano, L. J., Eichenbaum, M. and Evans, C. L. (2005), "Nominal rigidities and the dynamic effects of a shock to monetary policy", *Journal of political Economy*, 113(1), 1-45.

Chun, A.L. (2011), Expectations, Bond Yields, and Monetary Policy, *Review of Financial Studies*, vol 24(1), pp208-47.

Cochrane, J. and M. Piazzesi, (2005), "Bond Risk Premia", *American Economic Review*.

Cochrane, J. and M. Piazzesi, (2008), "Decomposing the Yield Curve", mimeo University of Chicago.

Constantinides, G.M. and Duffie, D., 1996, "Asset pricing with heterogeneous consumers", *Journal of Political Economy*, 104(2), pp.219-240.

Dai, Q and K. Singleton, 2000, "Specification Analysis of Affine Term Structure Models," *Journal of Finance*, vol. 55(5), pages 1943-78.

d'Amico, S., Kim, D.H. and Wei, M., 2014, "Tips from TIPS: the informational content of Treasury Inflation-Protected Security prices", Finance and Economics Discussion Series 2014-024, Federal Reserve Board of Governors.

Diamond, D and R. Verrecchia, (1981), "Information Aggregation in a Noisy Rational Expectations Economy", *Journal of Financial Economics* vol. 9, pages 221-235.

Duffee, Gregory (2002), "Term Premia and Interest Rate Forecasts in Affine Models," *Journal of Finance*, 57, pp.405-443.

Duffee, Gregory, (2010), "Sharpe ratios in term structure models", mimeo, Johns Hopkins University.

Duffee, Gregory, (2011), "Information in (and not in) the term structure of interest rates", forthcoming, *Review of Financial Studies.*

Duffee, Gregory, (2012), "Bond pricing and the macroeconomy, prepared for the Handbook of the Economics of Finance.

Duffie, Darrell and Rui Kan (1996), "Yield Factor Models of Interest Rates," *Mathematical Finance*, 64, pp.379-406.

Durbin, J. and S.J. Koopman, (2002), "A simple and efficient simulation smoother for state space time series analysis", *Biometrica*, 89, pp. 603-615.

Fama, E.F. and Bliss, R.R., (1987), "The Information in Long-Maturity Forward Rates", *American Economic Review*, 77(4), pp. 680-92.

Federal Reserve Board, (2005), *Testimony of Chairman Alan Greenspan Federal Reserve Board's semiannual Monetary Policy Report to the Congress Before the Committee on Banking, Housing, and Urban Affairs, U.S. Senate February 16, 2005.*

Geweke, John, (2005), *Contemporary Bayesian Econometrics and Statistics*, Wiley-Interscience.

Greenwood, R. and D. Vayanos, (2014), "Bond Supply and Excess Bond Returns", *Review of Financial Studies*, vol. 27, pp. 663-713.

Grossman, Sanford, 1976, "On the Efficiency of Competitive Stock Markets Where Trades Have Diverse Information", *Journal of Finance*, Vol. 31, No. 2, Papers and Proceedings. pp. 573-585.

Grossman, S. and J. Stiglitz, (1980), "On the Impossibility of Informationally Efficient Markets", *American Economic Review* vol. 70(3), pages 393-408.

Gurkaynak, R. S., Sack, B., and Wright, J. H., (2007), "The U.S. treasury yield curve: 1961 to the present", *Journal of Monetary Economics* vol. 54, pages 2291-2304.

Harrison, J.M. and D. Kreps, (1978), “Speculative Behaviour in a Stock Market with Heterogeneous Expectations”, *Quarterly Journal of Economics*, pp323-336.

Hamilton, J and Wu, C (2011), “Testable Implications of Affine-Term-Structure Models”, forthcoming, *Journal of Econometrics*.

Hellvig, Martin (1980), “On the aggregation of information in competitive markets,” *Journal of Economic Theory*, 22, issue 3 pp.477-498.

Joslin, S., M. Priebsch and K. Singleton, (2014), “Risk Premiums in Dynamic Term Structure Models with Unspanned Macro Risks”, *Journal of Finance*, No.3, pp.1197-1233.

Joslin, S., K. Singleton and H. Zhu (2011), ”A New Perspective on Gaussian DTSMs”, *Review of Financial Studies*, March 2011, v. 24, iss. 3, pp. 926-70.

Kim, D. H., Orphanides, A., (2005). ”Term structure estimation with survey data on interest rate forecasts” working paper, Board of Governors of the Federal Reserve System.

Leeper, E.M., Sims, C.A. and T Zha, (1996), “What does monetary policy do?”, *Brookings papers on economic activity*, pp1-78.

Ludvigson, S. and S. Ng, (2009), “Macro Factors in Bond Risk Premia”, *Review of Financial Studies* 22(12), 5027-5067.

Mankiw, N.G., 1986, “The equity premium and the concentration of aggregate shocks”, *Journal of Financial Economics*, 17(1), pp.211-219.

Nimark, K., (2017), “Dynamic Higher Order Expectations”, Center for Economic Policy Research DP 11863.

Piazzesi, M. and M. Schneider, (2011), “Trend and Cycle in Bond Premia”, unpublished manuscript.

Sims, C.A., (1980) “Macroeconomics and reality”, *Econometrica*, pp1-48.

Singleton, Kenneth J., (1987), “Asset prices in a time series model with disparately informed, competitive traders”, in *New Approaches to Monetary Economics,* Eds. W.A. Burnett and K.J. Singleton, Cambridge University Press.

Struby, Ethan, 2017, “Macroeconomic Disagreement in Treasury Yields ”, Working Paper, Boston College.

Townsend, Robert M., (1983), Forecasting the Forecasts of Others, *Journal of Political Economy*, vol 91, pp546-588.

Xiong, W. and H. Yan, (2010),“Heterogeneous Expectations and Bond Markets”, *Review of Financial Studies*, vol. 23(4), pp. 1433-1466.

Appendix A. Derivation of Eq (4)

Here, we show the intermediate steps required to go from equation (3) to equation (4) in the main text. Start with the no-arbitrage condition for agent j in logs

$$p_t^n = E\left[m_{t+1}^j \mid \Omega_t^j\right] + E\left[p_{t+1}^{n-1} \mid \Omega_t^j\right] + \frac{1}{2}Var\left(m_{t+1}^j + p_{t+1}^{n-1} \mid \Omega_t^j\right) \tag{A.1}$$

The assumption that individual agents are price takers means that when we evaluate the no-arbitrage condition (A.1) for agent j, we replace the expectation of the next period price p_{t+1}^{n-1} by agent j's expectation of what other agents will be willing to pay for the bond in the next period. Leading the no-arbitrage condition (A.1) and using it to substitute out p_{t+1}^{n-1} from agent j's expectation above gives

$$\begin{aligned} p_t^n \quad = \quad & E\left[m_{t+1}^j \mid \Omega_t^j\right] \\ & + E\left[E\left(m_{t+2}^i \mid \Omega_{t+1}^i\right) \mid \Omega_t^j\right] + E\left[E\left(p_{t+2}^{n-2} \mid \Omega_{t+1}^i\right) \mid \Omega_t^j\right] \\ & + \frac{1}{2}Var\left(m_{t+1}^j + p_{t+1}^{n-1} \mid \Omega_t^j\right) + \frac{1}{2}Var\left(m_{t+2}^i + p_{t+2}^{n-2} \mid \Omega_{t+1}^i\right) \end{aligned} \tag{A.2}$$

where the superscript i is used to indicate any agent i such that $i \neq j$. In this paper, agent j does not have any information relevant for predicting the expectations and stochastic discount factors of any other particular agent. Agent j's expectations of any other agent's expectation then coincide with agent j's expectation about the average expectation. That is, agent j's expectation about agent i's future expectation about bond prices and discount factors coincide with agent j's expectations about the future cross-sectional average expectation about the same quantities. We can use this fact to substitute

$$E\left(E\left[p_{t+2}^{n-2} \mid \Omega_{t+1}^i\right] \mid \Omega_t^j\right) = E\left(\int E\left[p_{t+2}^{n-2} \mid \Omega_{t+1}^i\right] di \mid \Omega_t^j\right) \tag{A.3}$$

and

$$E\left(E\left[m_{t+2}^i \mid \Omega_{t+1}^i\right] \mid \Omega_t^j\right) = E\left(\int E\left[m_{t+2}^i \mid \Omega_{t+1}^i\right] di \mid \Omega_t^j\right) \tag{A.4}$$

into (A.2). The log price of an n period bond can be written as

$$
\begin{aligned}
p_t^n \quad = \quad & E\left[m_{t+1}^j \mid \Omega_t^j\right] \\
& +E\left[\left(\int E\left[m_{t+2}^i \mid \Omega_{t+1}^i\right] di+\int E\left[p_{t+2}^{n-2} \mid \Omega_{t+1}^i\right] di\right) \mid \Omega_t^j\right] \\
& +\frac{1}{2} Var\left(m_{t+1}^j+p_{t+1}^{n-1} \mid \Omega_t^j\right)+\frac{1}{2} Var\left(m_{t+2}^i+p_{t+2}^{n-2} \mid \Omega_{t+1}^i\right)
\end{aligned} \tag{A.5}
$$

Continued recursive substitution of expectations about future prices from equation (A.1) gives the log price of a bond as the sum of higher order expectations about future stochastic discount factors

$$
\begin{aligned}
p_t^n \quad = \quad & E\left[m_{t+1}^j \mid \Omega_t^j\right] \\
& +E\left[\int E\left[m_{t+2}^i \mid \Omega_{t+1}^i\right] di \mid \Omega_t^j\right]+\ldots \\
& \ldots+E\left[\int E\left[\ldots \int E\left[m_{t+n}^{i'} \mid \Omega_{t+n-1}^{i'}\right] di' \ldots\right] di \mid \Omega_t^j\right] \\
& +\frac{1}{2} \sum_{s=0}^{n-1} Var\left(m_{t+1+s}^j+p_{t+1+s}^{n-1-s} \mid \Omega_{t+s}^j\right)
\end{aligned} \tag{A.6}
$$

which is equation (4) in the main text.

Appendix B. Proof of Proposition 1

We want to prove that the speculative term $p_t^n-\overline{p}_t^n$ is orthogonal to public information in real time, i.e.

$$
E\left(\left[p_t^n-\overline{p}_t^n\right] \omega_t\right)=0: \forall \;\; \omega_t \in \Omega_t \tag{B.1}
$$

where Ω_t is the public information set.

Start by taking expectations of $p_t^n-\overline{p}_t^n$ conditional on the public information set Ω_t to get

$$
E\left[p_t^n-\overline{p}_t^n \mid \Omega_t\right]=B_n E\left[\left(I-\overline{H}\right) X_t \mid \Omega_t\right]. \tag{B.2}
$$

and then use that $\left(I - \overline{H}\right) X_t$ is a vector of higher order prediction errors of the form

$$\left(I - \overline{H}\right) X_t = \begin{bmatrix} x_t - x_t \\ x_t^{(1)} - x_t^{(1)} \\ x_t^{(2)} - x_t^{(1)} \\ \vdots \\ x_t^{(k)} - x_t^{(1)} \end{bmatrix}.$$

The definition of the public information set (29) implies that $\Omega_t \subseteq \Omega_t^j$ for all j. Applying the law of iterated expectations then implies that $E\left[x_t^{(k)} \mid \Omega_t\right] = E\left[x_t \mid \Omega_t\right]$ for all k so that $E\left[\left(I - \overline{H}\right) X_t \mid \Omega_t\right] = \mathbf{0}$ which completes the proof.

Appendix C. Deriving the bond price equation

To find A_n and B_n in the conjectured bond price equation (10), start by substituting in the expressions (8) for the SDF into the no-arbitrage condition (2) to get

$$p_t^n = \log E\left[\exp\left(-r_t - \frac{1}{2}\Lambda_t^{j\prime}\Sigma_a\Lambda_t^j - \Lambda_t^{j\prime}a_{t+1}^j + p_{t+1}^{n-1}\right)|\Omega_t^j\right] \tag{C.1}$$

We will substitute out the price p_{t+1}^{n-1} from (C.1) via three intermediate steps. First, use the definition (9) to write p_{t+1}^{n-1} as the sum of agent j's expectations about the price and his forecast error

$$p_{t+1}^{n-1} = E\left[p_{t+1}^{n-1}|\Omega_t^j\right] + e_{n-1}'a_{t+1}^j \tag{C.2}$$

where e_n is a vector with a one in the nth element and zeros elsewhere. Second, note that the conjectured price equation (10) and the law of motion of the state (13) together with rational expectations imply that agent j's expectation of the next period price can be expressed as a function of his expectations about the current state, i.e.

$$E\left[p_{t+1}^{n-1}|\Omega_t^j\right] = A_{n-1} + B_{n-1}'\mu_X + B_{n-1}'\mathcal{F}E\left[X_t|\Omega_t^j\right]. \tag{C.3}$$

Third, agent j's expectation of the current state can by definition (15) be written as

$$E\left[X_t \mid \Omega_t^j\right] = HX_t^j \tag{C.4}$$

where H is the average expectations operator $H : \mathbb{R}^{d(\bar{k}+1)} \to \mathbb{R}^{d(\bar{k}+1)}$

$$H \equiv \begin{bmatrix} \mathbf{0} & I_{d\bar{k}} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \tag{C.5}$$

The matrix H increases each order of expectation in a hierarchy by annihilating the zero order expectation and replacing it with the first order expectation and by replacing the first order expectation with the second order expectation, and so on.

The expressions (C.2) - (C.4) can then be used in reverse order to substitute out p_{t+1}^{n-1} from (C.1). After simplifying the resulting expression is:

$$\begin{aligned} p_t^n &= -r_t + A_{n-1} + B_{n-1}'\mu_X + B_{n-1}'\mathcal{F}HX_t^j \\ &\quad + \frac{1}{2}e_{n-1}'\Sigma_a e_{n-1} - e_{n-1}'\Sigma_a\Lambda_t^j. \end{aligned} \tag{C.6}$$

That is, the price of an n-period bond is a function of the risk-free interest rate r_t, a number of constants and terms specific to agent j. The no-arbitrage condition (2) has to hold for all agents at all times. This implies that we could choose any agent j's state X_t^j as being the state variable that bond prices are a function of. However, the most convenient choice from a modeling perspective is to let bonds be priced by the SDF of the fictional agent whose state X_t is defined to coincide with the cross-sectional average state so that

$$X_t \equiv \int X_t^j dj \tag{C.7}$$

The identity of the average agent will change over time as idiosyncratic shocks change an individual agent's relative position in the cross-sectional distribution. However, the identity of the average agent is of no consequence and the advantage of letting the average agent's SDF price bonds is that it allows us to write log bond prices in the conjectured form (10), i.e. as a function of the average state X_t. We can thus substitute in X_t for X_t^j in (14) and (C.6).

The last step required to find the conjectured form of the bond price equation is to substitute in the definition (14) of Λ_t^j and the process for the short rate (5) to replace r_t in (C.6). After simplifying, we get

$$\begin{aligned} p_t^n &= -\delta_0 + A_{n-1} + B_{n-1}'\mu_X + \frac{1}{2}e_{n-1}'\Sigma_a e_{n-1} - e_{n-1}'\Sigma_a\Lambda_0 \\ & -\delta_X X_t + B_{n-1}'\mathcal{F}HX_t - e_{n-1}'\Sigma_a\Lambda_x X_t \\ & -e_{n-1}'\Sigma_a\Lambda_v \int E\left[v_t \mid \Omega_t^j\right] dj \end{aligned} \tag{C.8}$$

where

$$\delta_X' = \begin{bmatrix} \delta_x' & \mathbf{0} \end{bmatrix} \tag{C.9}$$

Restricting Λ_t^j

The final step needed to get the price p_t^n in the conjectured form (10) involves imposing two restrictions on Λ_t^j. The first is to set Λ_v equal to the negative of the inverse of the covariance of a_{t+1}^j, i.e.

$$\Lambda_v = -\Sigma_a^{-1} \tag{C.10}$$

Specifying Λ_v as in (C.10) has the additional advantage that letting Λ_t^j depend on $E\left[v_t \mid \Omega_t^j\right]$ does not introduce additional free parameters relative to the standard model without maturity specific shocks.

Second, we impose the normalization

$$\Lambda_x = \widehat{\Lambda}_x - B\left(I - H\right) \tag{C.11}$$

where

$$B = \begin{bmatrix} B_1' \\ \vdots \\ B_{\overline{n}} \end{bmatrix} \tag{C.12}$$

This restriction allows us to write the price in the conjectured form, i.e. as a function of the true maturity specific shocks, rather than the average trader's expectation of v_t. But clearly, it does not by itself restrict the elements of Λ_x.To see why this helps, first note that since agents have model consistent expectations, the observation of the price p_t^n must be consistent with each agent's expectations of X_t and v_t so that for any agent j

$$p_t^n = E\left[A_n + B_n X_t + v_t^n \mid \Omega_t^j\right] \tag{C.13}$$

or equivalently

$$p_t^n = A_n + B_n H X_t^j + e_{n-1}' E\left[v_t \mid \Omega_t^j\right] \tag{C.14}$$

must hold.

Using that $\int X_t^j dj \equiv X_t$ to equate the right hand side of (C.14) for the average agent with the right hand side of the conjectured bond price equation (10) gives

$$A_n + B_n H X_t + e_{n-1}' \int E\left[v_t \mid \Omega_t^j\right] dj = A_n + B_n X_t + v_t^n \tag{C.15}$$

which can be rearranged to

$$e_{n-1}' \int E\left[v_t \mid \Omega_t^j\right] dj = B_n \left(I - H\right) X_t + v_t^n \tag{C.16}$$

We can thus substitute the expression (C.16) into the bond price equation above. Using the restriction (C.10) and (C.11) and simplifying delivers the conjectured form

$$\begin{aligned} p_t^n \quad = \quad & -\delta_0 + A_{n-1} + B_{n-1}'\mu_X + \frac{1}{2}e_{n-1}'\Sigma_a e_{n-1} - e_{n-1}'\Sigma_a \Lambda_0 \\ & -\delta_X X_t + B_{n-1}'\mathcal{F}HX_t - e_{n-1}'\Sigma_a \widehat{\Lambda}_x X_t \\ & +v_t^n \end{aligned} \tag{C.17}$$

The bond price recursions

The bond price recursions for A_{n+1} and B_{n+1} in the bond price equation (10) are thus given by

$$A_{n+1} = -\delta_0 + A_n + B_n'\mu_X + \frac{1}{2}e_n'\Sigma_a e_n - e_n'\Sigma_a\Lambda_0 \tag{C.18}$$

and

$$B_{n+1}' = -\delta_X + B_n'\mathcal{F}H - e_n'\Sigma_a\widehat{\Lambda}_x \tag{C.19}$$

As in a full information set-up, the recursions (16) and (17) can be started from

$$A_1 = -\delta_0 \tag{C.20}$$

$$B_1 = -\delta_X' \tag{C.21}$$

where $p_t^1 = -r_t$.

Appendix D. The estimated parameters in Λ_t^j

In order to avoid over-parameterizing the model, we need to limit the number of free parameters in Λ_t^j. We will impose such restrictions so that the model nests the standard full information model of Joslin, Singleton and Zhu (2011) in the special case where $x_t^j = x_t$ and with no maturity specific shocks.

First, note that the vector of bond price innovations a_{t+1}^j can be expressed as a linear function of innovations to the state X_{t+1} and the vector of period $t+1$ maturity specific shocks v_{t+1}

$$a_{t+1}^j = \Psi \begin{bmatrix} X_{t+1} - E(X_{t+1} \mid \Omega_t^j) \\ v_{t+1} \end{bmatrix} \tag{D.1}$$

where the matrix Ψ is defined as

$$\Psi = \begin{bmatrix} B_1' & \\ \vdots & I_{\overline{n}-1} \\ B_{\overline{n}-1}' & \end{bmatrix}. \tag{D.2}$$

We thus have the equality

$$\Lambda_t^{j\prime} a_{t+1}^j = \Lambda_t^{j\prime}\Psi \begin{bmatrix} X_{t+1} - E(X_{t+1} \mid \Omega_t^j) \\ v_{t+1} \end{bmatrix} \tag{D.3}$$

In the model of Joslin, Singleton and Zhu (2011), the priced risks are the innovations to the true factors x_t and the price of risk varies over time as a function of the true factors x_t. Our framework will nest that model if the term $\Lambda_t^{j\prime} a_{t+1}^j$ in the SDF (8) satisfies the equality

$$\Lambda_t^{j\prime} a_{t+1}^j = \left(\begin{bmatrix} \lambda_0 \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} \lambda_x & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} X_t^j \right)' \begin{bmatrix} X_{t+1} - E(X_{t+1} \mid \Omega_t^j) \\ v_{t+1} \end{bmatrix}. \tag{D.4}$$

For (D.4) to hold, we need first of all that the maturity specific shocks are equal to zero. Second. we need the equalities

$$\begin{bmatrix} \lambda_0 \\ \mathbf{0} \end{bmatrix}' \begin{bmatrix} X_{t+1} - E(X_{t+1} \mid \Omega_t^j) \\ v_{t+1} \end{bmatrix} = \Lambda_0' a_{t+1}^j \tag{D.5}$$

and

$$\left(\begin{bmatrix} \lambda_x & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} X_t^j \right)' \begin{bmatrix} X_{t+1} - E(X_{t+1} \mid \Omega_t^j) \\ v_{t+1} \end{bmatrix} = \widehat{\Lambda}_x' a_{t+1}^j \tag{D.6}$$

to hold, which can be achieved by setting

$$\Lambda_0 = \Psi \left(\Psi'\Psi\right)^{-1} \begin{bmatrix} \lambda_0 \\ \mathbf{0} \end{bmatrix}, \quad \widehat{\Lambda}_x = \Psi \left(\Psi'\Psi\right)^{-1} \begin{bmatrix} \lambda_x & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \tag{D.7}$$

The matrix Φ in Equation (32) of Section IV.B is thus given by

$$\Phi \equiv \Psi \left(\Psi'\Psi\right)^{-1} \tag{D.8}$$

The free parameters in Λ_t^j are thus the 3×1 vector λ_0 and the 3×3 matrix λ_x, i.e. the same number as in the model of Joslin, Singleton and Zhu (2011).

Appendix E. Solving the model

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

$$X_{t+1} = \mu_X + \mathcal{F} X_t + \mathcal{C} u_{t+1} \tag{E.1}$$

where

$$X_t \equiv \begin{bmatrix} x_t^{(0)} \\ x_t^{(1)} \\ \vdots \\ x_t^{(\bar{k})} \end{bmatrix}$$

That is, to solve the model, we need to find the matrices $\mathcal{F}$ and $\mathcal{C}$ as functions of the parameters governing the short rate process, the maturity specific disturbances 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 (2011) 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, i.e. model consistent, expectations of x_t thus imply a law of motion for average expectations $x_t^{(1)}$ which can then be treated as a new stochastic process. Knowledge that other agents are rational means that second order expectations $x_t^{(2)}$ are determined by the average across agents of the rational expectations of the stochastic process $x_t^{(1)}$. The average third order expectation $x_t^{(3)}$ is then the average of the rational expectations of the process $x_t^{(2)}$, and so on. Imposing this structure on all orders of expectations allows us to find the matrices $\mathcal{F}$ and $\mathcal{C}$. Section A below describes how this is implemented in practice.

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

(i) The variances of higher order expectations of the factors 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

$$cov\left(x_t^{(k+1)}\right) \leq cov\left(x_t^{(k)}\right) \tag{E.2}$$

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

$$x_t^{(k)} \equiv x_t^{(k+1)} + \zeta_t^{(k+1)} \tag{E.3}$$

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

$$cov\left(x_t^{(k)}\right) = cov\left(x_t^{(k+1)}\right) + cov\left(\zeta_t^{(k+1)}\right). \tag{E.4}$$

Now, since covariances are positive semi-definite we have that

$$cov\left(\zeta_t^{(k+1)}\right) \geq \mathbf{0} \tag{E.5}$$

and the inequality (E.2) follows immediately. (This is an abbreviated description of a more formal proof available in Nimark 2011.)

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.

The law of motion of higher order expectations of the factors

To find the law of motion for the hierarchy of expectations X_t we use the following strategy. For given $\mathcal{F},\mathcal{C}$ in (E.1) and B_n' in (17) we will derive the law of motion for agent j's expectations of X_t , denoted $X^j_{t|t} \equiv E\left[X_t \mid \Omega^j_t\right]$. First, write the vector of signals z^j_t as a function of the state, the aggregate shocks and the idiosyncratic shocks

$$z_t^j \equiv \begin{bmatrix} x_t^j \\ r_t \\ y_t \end{bmatrix} = \mu_z + DX_t + R \begin{bmatrix} u_t \\ \eta_t^j \end{bmatrix} \tag{E.6}$$

where μ_z and D are given by

$$\mu_z = \begin{bmatrix} 0 \\ \delta_0 \\ -\frac{1}{2}A_2 \\ \vdots \\ -\frac{1}{\overline{n}}A_{\overline{n}} \end{bmatrix}, \quad D = \begin{bmatrix} I_d \quad \mathbf{0} \\ B_1' \\ \vdots \\ \overline{n}^{-1}B_{\overline{n}}' \end{bmatrix} \tag{E.7}$$

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

$$R = \begin{bmatrix} R_u & R_\eta \end{bmatrix}. \tag{E.8}$$

The matrix R_u picks out the appropriate maturity specific shocks v_t^n from the vector of aggregate shocks u_t so that

$$R_u = \begin{bmatrix} \mathbf{0} \\ V \end{bmatrix} \tag{E.9}$$

and R_η adds the idiosyncratic shocks $Q\eta_t^j$ to the exogenous state x_t to form the agent j specific signal vector x_t^j, i.e.

$$R_\eta = \begin{bmatrix} Q \\ \mathbf{0} \end{bmatrix} \tag{E.10}$$

Agent j's updating equation of the state $X_{t|t}^j \equiv E\left[X_{t+1} \mid \Omega_t^j\right]$ estimate will then follow

$$X_{t|t}^j = \mu_X + \mathcal{F}X_{t-1|t-1}^j + K\left(z_t^j - D(\mu_X + \mathcal{F}X_{t-1|t-1}^j)\right) \tag{E.11}$$

Rewriting the observables vector z_t^j as a function of the lagged state and current period innovations and taking averages across agents using that $\int \zeta_t(j)dj = 0$ yields

$$\begin{aligned} X_{t|t} &= \mu_X + \mathcal{F}X_{t-1|t-1} && \text{(E.12)}\\ &\quad + K\left(D\left(\mu_X + \mathcal{F}X_{t-1}\right) + \left(D\mathcal{C} + R_u\right)u_t - D(\mu_X + \mathcal{F}X_{t-1|t-1}^j)\right) && \text{(E.13)}\\ &= \mu_X + \left(\mathcal{F} - KD\mathcal{F}\right)X_{t-1|t-1} + KD\mathcal{F}X_{t-1} + K\left(D\mathcal{C} + R_u\right)u_t && \text{(E.14)} \end{aligned}$$

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

$$\begin{bmatrix} x_t \\ X_{t|t} \end{bmatrix} = \mu_X + \mathcal{F}\begin{bmatrix} x_{t-1} \\ X_{t-1|t-1} \end{bmatrix} + \mathcal{C}u_t \tag{E.15}$$

where $\mathcal{F}$ and $\mathcal{C}$ are given by

$$\begin{aligned} \mathcal{F} &= \begin{bmatrix} F^{\mathbb{P}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{0}_{d\times d} & \mathbf{0} \\ \mathbf{0} & \left[\mathcal{F} - KD\mathcal{F}\right]_- \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \left[KD\mathcal{F}\right]_- \end{bmatrix} && \text{(E.16)}\\ \mathcal{C} &= \begin{bmatrix} C & 0 \\ \mathbf{0} & \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \left[K\left(D\mathcal{C} + R_u\right)\right]_- \end{bmatrix} && \text{(E.17)} \end{aligned}$$

where $[\cdot]_{\_}$ indicates that the a last row or column has been truncated to make a the matrix $[\cdot]$ conformable, i.e. implementing that $x_t^{(k)} = 0 : k > \overline{k}$. The Kalman gain K in (E.11) is given by

$$\begin{aligned} K &= \left(\Sigma_{t+1|t} D' + \mathcal{C} R_u\right) \left(D \Sigma_{t+1|t} D' + R R'\right)^{-1} && \text{(E.18)} \\ \Sigma_{t+1|t} &= \mathcal{F} \left(\Sigma_{t+1|t} - \left(\Sigma_{t+1|t} D' + \mathcal{C} R_u\right) \left(D \Sigma_{t+1|t} D' + R R'\right)^{-1} \left(\Sigma_{t+1|t} D' + \mathcal{C} R_u\right)'\right) \mathcal{F}' && \text{(E.19)} \\ &\quad + \mathcal{C}\mathcal{C}' \end{aligned}$$

The model is solved by finding a fixed point that satisfies (17), (E.16), (E.17), (E.18) and (E.19).

Appendix F. The matrices in the estimated state space system

The objects in the state space system left undefined in Section IV.G of the main text are given by

$$\mu_t^{\overline{z}} = \begin{bmatrix} -\frac{1}{4} A_4 \\ -\frac{1}{8} A_{20} \\ -\frac{1}{40} A_{40} \\ -\mathbf{1}_{(m\times 1)} \times (A_1 + B_1' \mu_X) \\ -\frac{1}{40} \mathbf{1}_{(m\times 1)} \times (A_{40} + B_{40}' \mu_X) \end{bmatrix}, \quad \bar{D}_t = \begin{bmatrix} -\frac{1}{4} B_4' \\ -\frac{1}{8} B_{20}' \\ -\frac{1}{40} B_{40}' \\ -\mathbf{1}_{(m\times 1)} \times B_1' \mathcal{F} H \\ -\frac{1}{40} \mathbf{1}_{(m\times 1)} \times B_{40}' \mathcal{F} H \end{bmatrix} \tag{F.1}$$

$$\overline{R}_t = \begin{bmatrix} \sigma_v \times \begin{bmatrix} e_{d+4-1} & e_{d+20-1} & e_{d+40-1} \end{bmatrix}' & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & I_m \times B_1' \mathcal{F} \Sigma_j^{1/2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & I_m \times \frac{1}{40} B_{40}' \mathcal{F} \Sigma_j^{1/2} \end{bmatrix} \tag{F.2}$$

where m is the number of survey responses available in period t and e_i is a vector with a one in the i^{th} position and zeros elsewhere.

Appendix G. Computing the cross-sectional variance Σ_j

The idiosyncratic noise shocks η_t^j are white noise processes that are orthogonal across agents and to the aggregate shocks v_t and ε_t. This implies that the cross-sectional variance of expectations is equal to the part of the unconditional variance of agent j's expectations that is due to idiosyncratic shocks. This quantity can be computed by finding the variance of the estimates in agent j's updating equation (E.11), but with the aggregate shocks v_t and ε_t "switched off". The covariance Σ_j of agent j's state estimate due to idiosyncratic shocks is defined as

$$\Sigma_j \equiv E\left(E\left[X_t \mid \Omega_t^j\right] - \int E\left[X_t \mid \Omega_t^i\right] dj\right)\left(E\left[X_t \mid \Omega_t^j\right] - \int E\left[X_t \mid \Omega_t^i\right] dj\right)' \tag{G.1}$$

and given by the solution to the Lyapunov equation

$$\Sigma_j = (I - KD)\,\mathcal{F}\Sigma_j\mathcal{F}'\,(I - KD)' + KR_\eta R_\eta K'. \tag{G.2}$$

which can be found by simply iterating on (G.2).

Appendix H. Higher order expectations about the factors

In this appendix, we discuss in more detail how the higher order expectations about the current factors x_t translate into the speculative component. The speculative component is non-zero only when agents on average believe that other agents either over- or underestimate future bond yields. Since the state X_t summarizes all information that is possible to know about future states, perceived disagreement about future bond yields can be reduced to perceived disagreement about the current state X_t. It is thus only when there is perceived disagreement about the current state, i.e. when higher order expectations about x_t differ from agents' first order expectations, that the speculative term will be non-zero.

The role played by perceived disagreement for determining the speculative component is illustrated in Figure 5. There, we have plotted smoothed estimates of the exogenous factors x_t (solid black line), together with the average first order expectation (dashed black line) and the average higher order expectations (grey lines) about the individual components in x_t. It is clear from the figure that in general, higher order expectations do not coincide with first order expectations. Comparing Figure 5 with the time series of the speculative component in Figures 1 and 2 shows that the speculative component is large (in absolute terms) when the discrepancy between the first and the higher order expectations is large. In those periods in which there is no perceived disagreement, i.e. in those periods when the higher order expectations about x_t coincide with the first order expectations, the term $\left(I - \overline{H}\right)X_t$ in (30) equals zero. As a consequence, the speculative component must then also be zero.

For some perspective on these results, note that in a full information model in which all agents can

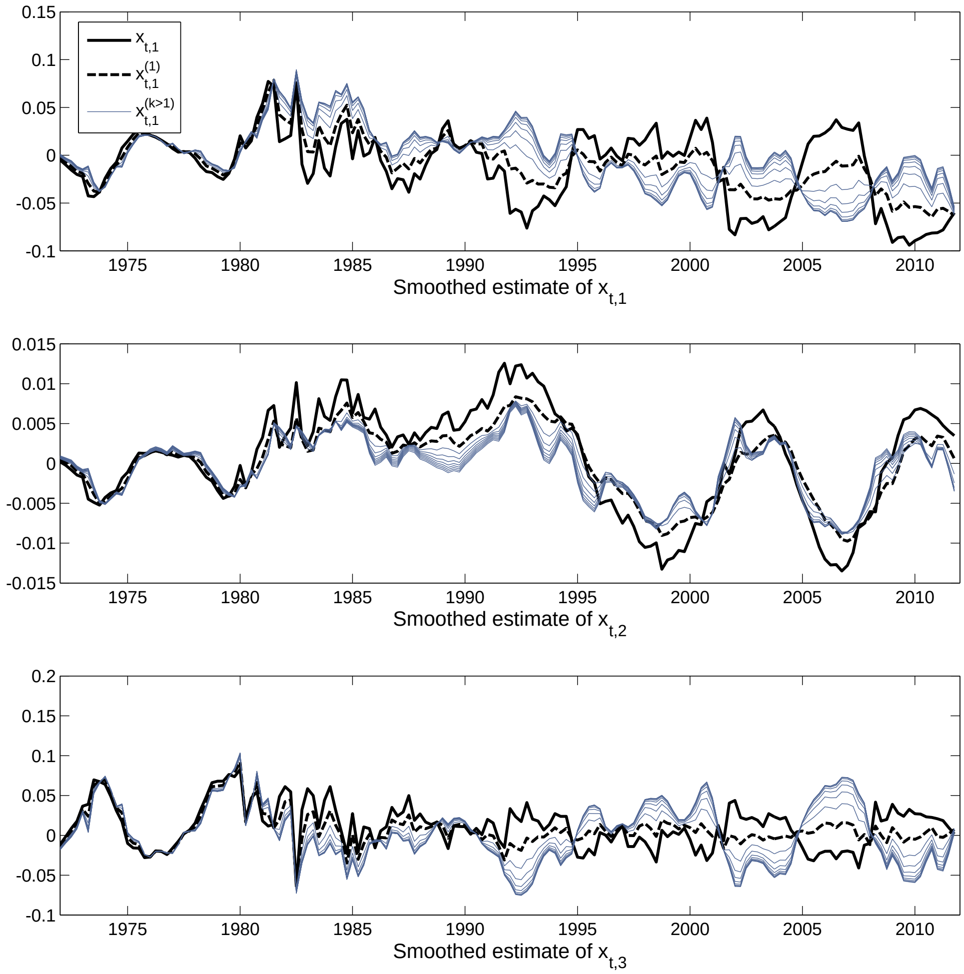

Figure 5 Smoothed estimates of state X_t. This figure plots the elements of the smoothed estimates of X_t. That is, the true factors x_t as well as the first and higher order expectations of x_t. The smoothed estimate is computed using the posterior mode of the model parameters. The sample is 1971Q2 to 2011Q4.

observe x_t perfectly and this fact is common knowledge, higher order expectations always coincide with the true factors. A single line would then represent all orders of expectations. Of course, in a full information model there would be no role for information driven speculation. Similarly, in periods when first and higher order expectations in our model by chance coincide with the true factors, there is perceived agreement about future bond yields and again no speculative motive. From the agents' perspective, there is substantial uncertainty about the true exogenous factors. The discrepancy between the solid black line and the dashed black line captures the average agent's expectation error about x_t. Agents are clearly not able to filter out the exogenous factors perfectly. If they were, the average first order expectation would coincide with the actual factors and the dashed line would lie on top of the solid black line.