Beauty Contests and the Term Structure*

Martin Ellison[†] Andreas Tischbirek[‡]

April 2019

Abstract

A novel decomposition highlights the scope for information to influence the term structure of interest rates. Based on the law of total covariance, we show that real term premia in macroeconomic models contain a component that depends on covariances of *realised* stochastic discount factors and a component that depends on covariances of *expectations* of those stochastic discount factors. The impact of different informational assumptions can then be identified by looking at their effect on the second, expectational, component. If agents have full information about technology in a simple macro-finance model then the conditional covariance of expectations is low, which contributes to the real term premia implied by the model being at least an order of magnitude too small, a result that is unchanged if some components of technology are unobservable or observed with noise. To generate realistic term premia, we draw on the beauty contest literature by differentiating between private and public information and introducing the possibility of strategic complementarities in the formation of expectations. A quantitative version of the model is found to explain a significant proportion of observed term premia when estimated using data on expectations of productivity growth from the Survey of Professional Forecasters.

Keywords: Yield Curve, Term Premia, Information Friction, Beauty Contest, Asset Pricing
JEL classification: E40, E43, E70, G12

*We would like to thank Philippe Bacchetta, Paul Beaudry, Kenza Benhima, John Cochrane, Stefano Eusepi, Piero Gottardi, Chris Gust, Refet Gürkaynak, Julian Kozlowski, Elmar Mertens, Emanuel Moench, Jean-Paul Renne, Florian Scheuer, Cédric Tille, Eric van Wincoop, Mirko Wiederholt, Fabian Winkler, and seminar participants at the Bank of Canada, Federal Reserve Board, the Graduate Institute Geneva, the University of Oxford, Simon Fraser University, the 2018 Growth and Business Cycle Research Group Conference in Manchester, the 2nd EABCN Conference on Asset Prices and the Macroeconomy in Mannheim, the European Summer Meeting of the Econometric Society in Cologne and the Oxford-Indiana Macroeconomic Policy Conference for helpful discussions at different stages of this project, as well as Simon Tièche for research assistance.

[†]University of Oxford, NuCamp and CEPR, e-mail: martin.ellison@economics.ox.ac.uk

[‡]HEC Lausanne, University of Lausanne, e-mail: andreas.tischbirek@unil.ch

Since investors lack any clear sense of objective evidence regarding prices of speculative assets, the process by which their opinions are derived may be especially social.

Robert J. Shiller, 1984

1 Introduction

The way that developments in the real economy affect the pricing of financial assets is central to much recent research in monetary economics and macro-finance. Despite this, quantitative DSGE models of the type developed by Christiano et al. (2005) and Smets and Wouters (2003, 2007) continue to have difficulty generating risk premia of magnitude anything like those seen in financial markets. For example, Rudebusch and Swanson (2012) find that the average term premium on a default-free nominal 10-year zero-coupon bond is less than one basis point in a medium-scale DSGE model with nominal rigidities, expected utility preferences and a reasonable coefficient of relative risk aversion. Updated estimates from Adrian et al. (2013) suggest that the term premium on 10-year US Treasury Bills 1999-2017 was, at 114.7 bps, at least two orders of magnitude higher.

This paper presents a new decomposition that stresses the importance of informational assumptions for the emergence of sizeable term premia in a broad class of macroeconomic models. We focus our attention on the relationship between information and the real term premium. Many models struggle to explain positive nominal term premia because they imply *negative* real term premia. For example, in models with long-run risk or rare disasters the real yield curve is typically downward-sloping. Such models then have to rely on unrealistically large inflation risk premia to match the upward-sloping average nominal yield curve found in US data.[1] By understanding better what is needed to generate positive real term premia, we make significant progress with nominal term premia too.

Using the standard no-arbitrage pricing condition, we divide the real term premium into a component that is directly affected by information and a component that is not. By applying the law of total covariance, the mean of the real term premium at any maturity can be shown to depend on covariances of successive *realised* stochastic discount factors and covariances of successive *expectations* of stochastic discount factors, with the latter being directly affected by the informational assumptions imposed on the model. The new decomposition prompts us to attribute the quantitative failure of medium-scale DSGE models to informational assumptions that are unable to deliver sufficient covariance in expectations. Having identified the channels by which information impacts on term premia, the new decomposition provides a useful framework in which to explore alternative assumptions about information.

[1] See Albuquerque et al. (2016) for an excellent discussion of this point and evidence from Treasury Inflation Protected Securities (TIPS) showing that the average US real yield curve has been steeply upward sloping. Correspondingly, Hördahl and Tristani (2012) estimate that in the euro area "most of the average slope in the nominal term structure is due to compensation for real risks, rather than inflation risk".

We begin our analysis by focusing on the real term premium on two-period bonds in a simplified yet tractable model where persistent and transitory technology shocks are the only exogenous disturbances. In keeping with quantitative results from medium-scale DSGE models, we find that the term premium is small if agents are assumed to have full information about the current level of technology. This is because the unconditional covariance of both realised and expected stochastic discount factors is low for any realistic paramaterisation of the technology process and any reasonable level of risk aversion. The unconditional covariance of expectations is especially small, a result that still holds if we assume that the representative agent has only partial or noisy information and so has to infer the current level of technology and break it down into persistent and transitory components.

In thinking of alternative informational assumptions that may increase the covariance of expected stochastic discount factors, we are guided by the quote of Shiller (1984) above the start of this introduction. We interpret his hypothesis of a social process driving opinions as motivation for exploring informational assumptions that introduce strategic complementarities into the formation of expectations. Drawing on the literature on beauty contests, we consider a framework in which agents have access to a noisy signal with a public and a private component and price the real term premium according to forecasts of technology that reward the agent for being similar not only to fundamentals but also to the average forecast across all agents.[2] In this scenario, the greater the strategic complementarity in forecasting the more that agents are willing to incorporate aggregate and idiosyncratic noise into their forecasts of future technology. If the aggregate noise is persistent then it will induce additional unconditional autocovariance in expectations of the stochastic discount factor, which increases term premia as agents demand compensation for the extra risk they face.

The potential for complementarities in forecasting to occasion significant term premia is first demonstrated in our simple analytical framework, where as informational assumptions they gain traction by increasing the unconditional covariance of expectations of the stochastic discount factor. We find the same result in a more general model in which the degree of strategic complementarities is tightly disciplined by data on individual forecasts of productivity from the Survey of Professional Forecasters. When estimated, the general model generates a real premium in the price of 5-year zero-coupon bonds of 26.7 bps if strategic complementarities are present, a distinct improvement on the 3.7-10.7 bps that is estimated when they are absent, and close to 50% of the 57.2 bps term premium on 5-year US Treasury Bills 1997-2017 implied by the updated estimates of Adrian et al. (2013). Our decomposition attributes all the extra term premium to a larger covariance in expectations of the stochastic discount factor.

The paper is organised as follows. Section 2 discusses related literature before Section 3 applies the law of total covariance to derive our new decomposition. Section 4 presents the implications of different informational assumptions for the premium on a two-period bond in

[2] The term "beauty contest" should be understood here as in Angeletos et al. (2018) as "a coordination game with linear best responses and incomplete information".

our simplified analytical model. Section 5 presents the more general business cycle model and discusses quantitative results from its estimation. Section 6 concludes.

2 Related literature

The finding that standard general equilibrium asset pricing models are inconsistent with data on the equity premium and the term structure of interest rates dates back to Mehra and Prescott (1985), Backus et al. (1989), and Hansen and Jagannathan (1991). A large literature evolved in response to this apparent "puzzling" disconnect between models and data, including explanations based on recursive preferences (Epstein and Zin, 1989), long-run risk (Bansal and Yaron, 2004; Croce, 2014), rare disasters (Rietz, 1988; Barro, 2006) and habit formation (Constantinides, 1990; Jermann, 1998; Abel, 1999; Campbell and Cochrane, 1999). However, models which include one or more of the above mechanisms typically explain data on risk premia only at the expense of implausibility still remaining at some other margin.

Models that explain the nominal term structure by combining recursive preferences with long run risk or rare disasters typically generate *negative* real term premia. Agents in these models are worried that their future real consumption growth may decline and, since they have a preference for the early resolution of uncertainty, are very willing to hold real long-term bonds as a hedge against future bad states of the world. The price of real long-term bonds is driven up and the real term premium is negative (Piazzesi and Schneider, 2007).[3] To match the positive nominal term premium in the data, these models need inflation risk premia to be large enough to dominate the negative real term premium that results. This is usually achieved by generating a negative correlation between consumption growth and inflation.[4] Albuquerque et al. (2016) find no empirical evidence of such a significantly negative relationship between consumption growth and inflation over a longer time horizon, and furthermore argue that evidence points to the real term premium being positive rather than negative.[5]

Another problem is the long-standing tension between the degree of risk aversion needed to match financial as opposed to macroeconomic variables. Tallarini (2000) shows that asset pricing models with recursive preferences can simultaneously match data on the equity risk premium and the risk-free rate with a coefficient of relative risk aversion of 50, considerably less than what is needed to justify the risk premium with standard preferences but still well above the levels usually estimated in micro data. Rudebusch and Swanson (2012) ask whether a medium-scale DSGE model with long-run risk and recursive preferences can match moments of macroeconomic aggregates and the average term premium on a nominal 10-year bond. The

[3] See, for example, Kung (2015) for the case of long-run risk and Nakamura et al. (2013) for rare disasters.

[4] Shocks that persistently lower consumption growth and increase inflation imply that the real pay-off from holding nominal long-term bonds is positively correlated with consumption growth. Their price falls as consumption-smoothing investors demand compensation for holding them.

[5] Albuquerque et al. (2016) propose an asset pricing model in which shocks to the rate of time preference give rise to an upward-sloping real yield curve. This complements our approach, which instead concentrates on the role played by information in a real business cycle model with a full set of cross-equation restrictions.

answer is yes with a plausible degree of long-run inflation risk, but their preferred specification still relies on a coefficient of relative risk aversion of 110.

An attempt to rationalise the ostensibly high coefficients of relative risk aversion has been made by Barillas et al. (2009) in the literature on ambiguity aversion. They show that the market price of risk includes components that compensate agents not only for known risks but also for ambiguity surrounding the true data generating process for returns, in which case the high coefficients of relative risk aversion in recursive preference formulations can be re-interpreted as measuring reasonable levels of aversion to both risk and ambiguity. Results from van Binsbergen et al. (2012) caution against applying this reasoning to the results of Rudebusch and Swanson (2012) though, finding that inflation would still need to be unrealistically volatile for a standard DSGE model with nominal frictions to square with the US Treasury yield curve, even if the high coefficient of relative risk aversion is rationalised by ambiguity aversion.

A number of authors have explored how amending the standard informational assumptions in macroeconomic models can help explain the behaviour of selected financial assets. Cogley and Sargent (2008) create significant equity premia by requiring agents to re-learn the law of motion for consumption growth after a bout of pessimism brought on by the Great Depression, although the premia eventually dissipate as the influence of the initial pessimism declines. Collard et al. (2018) report similar findings in a model where ambiguity-averse agents fear that their model of the joint distribution of future consumption and dividends may be misspecified. Another promising line of research is by Luo (2010) and Luo and Young (2016), who examine simple portfolio choice models in which agents solve a rational inattention problem of the type introduced by Sims (2003). Models in which higher-order expectations play a role in the dynamics of equity prices have been proposed in work by Allen et al. (2006), Bacchetta and Wincoop (2008), Kasa et al. (2014) and Barillas and Nimark (2018) that is broadly related to the beauty contest literature of Morris and Shin (2002) and Angeletos and Pavan (2007). The idea that strategic motives can amplify volatility in aggregate expectations is studied by Angeletos and La'O (2013), who demonstrate how random matching of agents uncertain about their idiosyncratic productivity makes expectations volatile even when aggregate productivity is common knowledge. Benhabib et al. (2015) make a related point in a model with multiple equilibria. We regard these approaches as complementary to ours. Comparable to Andrade et al. (2016), we rely on survey data to evaluate the significance of the informational imperfections in our model.

3 Decomposing the term premium

This section explains bond pricing in dynamic stochastic general equilibrium and applies the law of total covariance to derive our novel decomposition of risk premia. In what follows we fix ideas by assuming that financial assets are priced by a representative agent, leaving until Section 4.5 the decomposition of risk premia in beauty contests with heterogeneous information and strategic complementarities in expectations formation.

3.1 The household's optimisation problem

The household's expected utility is

$$\mathrm{E}_t \sum_{s=t}^{\infty} \beta^{s-t} u(c_s, l_s) \tag{1}$$

in all periods $t \geq 0$, where $\beta \in (0,1)$ is the discount factor, E_t is the conditional expectation operator with respect to the information set $\mathcal{I}_t$, $(c_t, l_t) \in \Omega \subseteq (\mathbb{R}^+)^2$ is a consumption and labour supply choice and $u : \Omega \to \mathbb{R}$ is the period utility function. The flow budget constraint is

$$c_t + \sum_{n=1}^{N} p_t^{(n)} b_t^{(n)} = w_t l_t + d_t + \sum_{n=1}^{N} p_t^{(n-1)} b_{t-1}^{(n)} \tag{2}$$

where w_t is the real wage and d_t is lump-sum income. The household can invest in non-contingent zero-coupon real bonds that have a redemption value of one unit of consumption at maturity. Their period t holdings of bonds due to mature in n periods are denoted by $b_t^{(n)}$ for $n = 1, 2, \ldots, N$, with corresponding prices $p_t^{(n)}$. Note that bond holdings $b_{t-1}^{(n)}$ inherited from period $t-1$ are priced at $p_t^{(n-1)}$ in period t. The price of a bond at maturity is its redemption value, hence $p_t^{(0)} = 1$.

The representative household maximises utility (1) subject to the sequence of flow budget constraints (2) and conditions ruling out Ponzi schemes. An interior solution satisfies N consumption Euler equations of the form

$$p_t^{(n)} = \mathrm{E}_t \left(m_{t+1} p_{t+1}^{(n-1)} \right) \tag{3}$$

for $n \in \{1, 2, \ldots, N\}$, where m_{t+1} is the stochastic discount factor defined by

$$m_{t+1} \equiv \beta \frac{u_c(c_{t+1}, l_{t+1})}{u_c(c_t, l_t)} \tag{4}$$

Equation (3) is the standard asset pricing condition equating the price of an asset to its expected discounted price in the following period. If the law of iterated expectations holds and the stochastic discount factor $m_t \in \mathcal{I}_t$ is in the information set at time t, it implies that the price of a bond $p_t^{(n)}$ is given by the expected product of successive stochastic discount factors over the maturity of the bond.[6]

$$p_t^{(n)} = \mathrm{E}_t \prod_{j=1}^{n} m_{t+j} \tag{5}$$

Bond prices can be translated into yields by defining the continuously-compounded yield of an n-period zero-coupon bond as the value of $i_t^{(n)}$ that satisfies $p_t^{(n)} \exp(n i_t^{(n)}) = 1$, in which case the yield in per-period terms is

$$i_t^{(n)} = -\frac{1}{n} \ln p_t^{(n)} \tag{6}$$

[6] See Appendix A.1 for details.

We denote the yield on a one period bond as $i_t \equiv i_t^{(1)}$ without a superscript to simplify notation.

3.2 Comovement and the term premium

The stochastic nature of bond prices implies that bonds of maturity $n > 1$ are a source of risk for the household, even if there is no possibility of default.[7] It is well-known from the capital asset pricing model (CAPM) that a risk-averse household demands compensation for holding bonds if there is undesirable comovement between bond prices and the household's marginal utilities of consumption. Since marginal utilities depend on realised stochastic discount factors and bond prices are a function of expected stochastic discount factors, the term premium will depend on the joint autocovariance structure of realisations and expectations.

The real term premium at maturity n is defined relative to the hypothetical price of an n-period bond under risk neutrality. Following the literature in Andreasen (2012), Gürkaynak and Wright (2012) and Rudebusch and Swanson (2008, 2012), the risk-neutral price is assumed to be

$$\tilde{p}_t^{(n)} = \mathrm{e}^{-i_t} \mathrm{E}_t \tilde{p}_{t+1}^{(n-1)} = \mathrm{E}_t \prod_{j=0}^{n-1} \mathrm{e}^{-i_{t+j}} \tag{7}$$

which is the redemption value of an n-period bond discounted by expected future one-period bond yields rather than expected household stochastic discount factors. Translating prices into yields as we did for equation (6), the risk-neutral yield at maturity n is

$$\tilde{i}_t^{(n)} \simeq \frac{1}{n} \mathrm{E}_t \sum_{j=0}^{n-1} i_{t+j} \tag{8}$$

to an approximation that ignores a Jensen's inequality term. It coincides with the bond yield predicted by the expectations theory of the term structure. The per-period real term premium in bond prices is defined by

$$\psi_t^{(n)} \equiv \frac{1}{n} \left(\tilde{p}_t^{(n)} - p_t^{(n)} \right) \tag{9}$$

with $\psi_t^{(1)} = 0$ because $\tilde{p}_t^{(1)} = p_t^{(1)}$. With regard to yields the term premium is $i_t^{(n)} - \tilde{i}_t^{(n)}$.

3.3 A useful decomposition

The real term premium on bonds for $n = 2$ is given by equations (3), (7) and (9) as

$$\psi_t^{(2)} = \frac{1}{2} \left[\mathrm{e}^{-i_t} \mathrm{E}_t \tilde{p}_{t+1}^{(1)} - \mathrm{E}_t \left(m_{t+1} p_{t+1}^{(1)} \right) \right] \tag{10}$$

where the second term on the right hand side of (10) can be expanded to

$$\mathrm{E}_t \left(m_{t+1} p_{t+1}^{(1)} \right) = \mathrm{e}^{-i_t} \mathrm{E}_t \mathrm{e}^{-i_{t+1}} + \mathrm{Cov}_t \left(m_{t+1}, p_{t+1}^{(1)} \right) \tag{11}$$

[7] Holding bonds of maturity $n = 1$ is not risky as they always deliver one unit of consumption next period.

by the definition of conditional covariance and by substituting in for $\mathrm{E}_t m_{t+1} = \mathrm{e}^{-i_t}$ and $\mathrm{E}_t p_{t+1}^{(1)} = \mathrm{E}_t \mathrm{e}^{-i_{t+1}}$ from the Euler equation (3) and the bond yield equation (6).[8] The first term on the right hand side of (10) similarly expands to

$$\mathrm{e}^{-i_t}\mathrm{E}_t \tilde{p}_{t+1}^{(1)} = \mathrm{e}^{-i_t}\mathrm{E}_t \mathrm{e}^{-i_{t+1}} \tag{12}$$

and the real term premium in the price of a two-period bond

$$\psi_t^{(2)} = -\frac{1}{2}\mathrm{Cov}_t\left(m_{t+1}, p_{t+1}^{(1)}\right) \tag{13}$$

is positive if the conditional covariance between the stochastic discount factor in $t+1$ and the price of a one-period bond in $t+1$ is negative.

To proceed to our decomposition, we begin by writing the term premium solely in terms of stochastic discount factors. Since $p_{t+1}^{(1)} = m_{t+2} + (\mathrm{E}_{t+1}m_{t+2} - m_{t+2})$, we have

$$\begin{aligned}\psi_t^{(2)} &= -\frac{1}{2}\mathrm{Cov}_t\left(m_{t+1}, p_{t+1}^{(1)}\right)\\ &= -\frac{1}{2}\left[\mathrm{Cov}_t\left(m_{t+1}, m_{t+2}\right) + \mathrm{Cov}_t\left(m_{t+1}, \mathrm{E}_{t+1}m_{t+2} - m_{t+2}\right)\right]\end{aligned} \tag{14}$$

where the second covariance term in equation (14) is zero in a large class of models considered in the literature because the one-step ahead forecast error $\mathrm{E}_{t+1}m_{t+2} - m_{t+2}$ is orthogonal in equilibrium to the stochastic discount factor m_{t+1}. This is the case for example under rational expectations, or more generally when making the standard assumptions that the law of iterated expectations holds and bonds are priced in period $t+1$ drawing on knowledge of m_{t+1}.[9] The unconditional mean of the real term premium under such conditions follows as

$$\begin{aligned}\mathrm{E}\psi_t^{(2)} &= -\frac{1}{2}\mathrm{E}\left[\mathrm{Cov}_t\left(m_{t+1}, m_{t+2}\right)\right]\\ &= \frac{1}{2}\left[-\mathrm{Cov}(m_{t+1}, m_{t+2}) + \mathrm{Cov}(\mathrm{E}_t m_{t+1}, \mathrm{E}_{t+1}m_{t+2})\right]\end{aligned} \tag{15}$$

by applying the law of total covariance.[10]

The usefulness of decomposition (15) lies in its ability to separate out the unconditional mean real term premium into the covariance of successive *realisations* of the stochastic discount factor and the covariance of successive *expectations* of the stochastic discount factor. The nominal term premium can be decomposed in an analogous way based on the nominal pricing kernel. The macroeconomic literature has focused on the first component in the decomposition. Holding the information structure and expectations formation process fixed, authors have introduced new types of preferences and shocks that cause large autocovariance in the realised stochastic discount factor. An example is explanations based on the recursive preference specification of Epstein

[8] The notation Cov_t signifies that covariance is conditional on period t information.

[9] The second covariance term in equation (14) also vanishes if bonds are held to maturity rather than being tradable every period.

[10] See Appendix A.2 for details.

and Zin (1989) and Weil (1989), which rely on high risk aversion and a negative relationship between consumption growth and inflation to induce a large negative autocovariance in the nominal stochastic discount factor that leads to significant nominal term premia.[11] Our focus is on the second component in the decomposition. We are interested in the role that expectations play in the determination of the real term premium, since the inability of standard models to generate significantly upward-sloping real yield curves is instrumental in their failing to match the nominal yield curve data.

It is instructive to decompose term premia at longer maturities in this way too. For example, for $n = 3$ the conditional term premium is

$$\psi_t^{(3)} = \frac{1}{3}\{-\mathrm{Cov}_t\left(m_{t+1}, m_{t+2}m_{t+3}\right) \\ +\mathrm{e}^{-i_t}\left[-\mathrm{Cov}_t\left(m_{t+2}, m_{t+3}\right) + \mathrm{Cov}_t\left(\mathrm{E}_{t+1}m_{t+2}, \mathrm{E}_{t+2}m_{t+3}\right)\right]\} \tag{16}$$

and the corresponding unconditional term premium is

$$\mathrm{E}\psi_t^{(3)} = \frac{1}{3}\{-\mathrm{Cov}\left(m_{t+1}, m_{t+2}m_{t+3}\right) + \mathrm{Cov}\left(\mathrm{E}_t m_{t+1}, \mathrm{E}_{t+1}m_{t+2}m_{t+3}\right) \\ + \mathrm{e}^{-i_t}\left[-\mathrm{Cov}\left(m_{t+2}, m_{t+3}\right) + \mathrm{Cov}\left(\mathrm{E}_{t+1}m_{t+2}, \mathrm{E}_{t+2}m_{t+3}\right)\right] \\ + \mathrm{Cov}\left[\mathrm{e}^{-i_t}, -\mathrm{Cov}_t\left(m_{t+2}, m_{t+3}\right) + \mathrm{Cov}_t\left(\mathrm{E}_{t+1}m_{t+2}, \mathrm{E}_{t+2}m_{t+3}\right)\right]\} \tag{17}$$

Since a three-period bond issued in t becomes a two-period bond in $t+1$, the second line of equation (17) contains the discounted mean term premium at two-period maturity (15) discussed above. The first line, which includes covariance terms of realisations and expectations of the one-period stochastic discount factor with the successive two-period stochastic discount factor $m_{t+2}m_{t+3}$, reflects the compensation demanded for the additional risk incurred between t and $t+1$. The final line contains a small corrective term that accounts for potential covariance of the risk-free rate with the future term-premium components it discounts.

Proposition 1 and Corollary 1 in Appendix A.3 present the general case. The resulting expressions are independent of assumptions about the stochastic processes of external disturbances and rely on much weaker informational assumptions than those typically imposed in macroeconomic models, which gives the decomposition power to inform across many different environments. The only requirements are that $\{m_s\}_{s=0}^{t} \in \mathcal{I}_t \;\forall\; t$ and that the law of iterated expectations holds. We are particularly interested in how the expectations component reacts to different informational assumptions.

[11] Fuerst (2015) gives a detailed discussion of this fact, pointing out that "To match the term premium, we need a peculiar time series behavior of the nominal pricing kernel."

4 A simple analytical model

This section uses the new decomposition to connect informational assumptions and term premia in simple analytical models, in preparation for the more general investigation in Section 5.

4.1 Households, firms and technology

The economy consists of a representative household and a representative firm. The representative household has inelastic labour supply $l_t = \bar{l}$ and solves the household's optimisation problem in Section 3.1. The representative firm produces output y_t according to

$$y_t = A_t \bar{l}^{1-\alpha} \tag{18}$$

where the fixed capital stock is normalised to one, $\alpha \in (0, 1)$ is a parameter and the logarithm of technology $a_t \equiv \ln A_t$ follows the exogenous process

$$a_t = x_t + \eta_t \tag{19}$$

$$x_t = \rho x_{t-1} + \varepsilon_t \tag{20}$$

with $\rho \in (-1, 1)$, $\eta_t \sim N(0, \sigma_\eta^2)$, $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$ and $\text{Cov}(\eta_t, \varepsilon_t) = 0$. The logarithm of technology hence has an AR(1) persistent component x_t and an i.i.d. transitory component η_t. The household's optimisation problem defines the general form (4) of the stochastic discount factor. We assume that utility is logarithmic in consumption so the stochastic discount factor is

$$m_{t+1} = \beta \left(\frac{c_{t+1}}{c_t} \right)^{-1} \tag{21}$$

and the coefficient of relative risk aversion is fixed at one. In equilibrium $c_t = A_t \bar{l}^{1-\alpha}$ for all t, since in the household's budget constraint (2) labour is paid its marginal product, firms make zero profits and bonds are in zero net supply. To a log-linear approximation the stochastic discount factor implied by the model becomes

$$m_{t+1} = \beta \left(\frac{A_{t+1} \bar{l}^{1-\alpha}}{A_t \bar{l}^{1-\alpha}} \right)^{-1} \approx \beta \left(1 + a_t - a_{t+1}\right) \tag{22}$$

It is instructive to begin with the shortest maturity $n = 2$ at which bonds have a non-zero real term premium, for which the unconditional mean real term premium is given by equation (15). The first term on the right hand side of (15) is the unconditional autocovariance of successive realised stochastic discount factors. When the stochastic discount factor is approximated by equation (22), the technology process (19) and (20) implies

$$-\text{Cov}(m_{t+1}, m_{t+2}) = \beta^2 \left(\frac{1-\rho}{1+\rho} \sigma_\varepsilon^2 + \sigma_\eta^2 \right) \tag{23}$$

The autocovariance of realised stochastic discount factors $\text{Cov}(m_{t+1}, m_{t+2})$ in equation (23) is independent of the household information set $\mathcal{I}_t$, which means that informational assumptions are fully reflected in the expectations component $\text{Cov}(\text{E}_t m_{t+1}, \text{E}_{t+1} m_{t+2})$ in the simple model.

4.2 Full information

The benchmark informational assumption is that households have full information about current and past values of the persistent and transitory components of technology. The assumption is that $m^t, x^t, \eta^t, a^t \subset \mathcal{I}_t^*$, where $\mathcal{I}_t^*$ is the information set of the representative household and the superscript notation $z^t \equiv \{z_s\}_{s=0}^t$ indicates the entire history of a variable up to and including period t. Taking expectations of the stochastic discount factor implied by the model (22) with respect to $\mathcal{I}_t^*$ gives its conditional expectation as

$$\text{E}(m_{t+1}|\mathcal{I}_t^*) = \beta(1 + a_t - \rho x_t) \tag{24}$$

and the unconditional autocovariance of successive expected stochastic discount factors is

$$\text{Cov}\left[\text{E}(m_{t+1}|\mathcal{I}_t^*), \text{E}(m_{t+2}|\mathcal{I}_{t+1}^*)\right] = \beta^2 \frac{\rho(1-\rho)}{1+\rho} \sigma_\varepsilon^2 \tag{25}$$

With the unconditional autocovariance of successive realised stochastic discount factors (23) being independent of informational assumption, the unconditional mean of the real term premium can be calculated analytically. Substituting the covariance terms (23) and (25) into (15) it is

$$\text{E}\psi_{FI}^{(2)} = \frac{1}{2}\beta^2 \left[(1-\rho)\sigma_\varepsilon^2 + \sigma_\eta^2\right] \tag{26}$$

Figure 1 depicts the mean real term premium $\text{E}\psi_{FI}^{(2)}$ with full information in our model. The units on the vertical axis are such that 1×10^{-5} represents a term premium of 0.1 basis points. In both panels we fix the discount factor $\beta = 0.99$ and the overall volatility in technology a_t at $\text{Var}(a_t) = 0.01^2$, which leaves freedom to explore term premia in two dimensions. The left panel plots the real term premium against the degree of persistence ρ in the persistent component x_t of the technology process, holding constant its relative contribution to overall volatility in technology by setting σ_ε^2 and σ_η^2 such that $\text{Var}(x_t)/\text{Var}(a_t) = 0.9$. In the right panel we have the real term premium plotted against the relative contribution of the persistent component, adjusting σ_ε^2 and σ_η^2 while fixing the persistence of the persistent component at $\rho = 0.8$. The term premium in each panel is decomposed into covariance terms $-\text{Cov}(m_{t+1}, m_{t+2})$ and $\text{Cov}\left[\text{E}(m_{t+1}|\mathcal{I}_t^*), \text{E}(m_{t+2}|\mathcal{I}_{t+1}^*)\right]$ that respectively depend on realisations and expectations of stochastic discount factors.

The real term premium for the model with full information is small and decreasing in the persistence parameter ρ and the relative contribution of the persistent component of technology. Most of the real term premium comes from the component that depends on realisations, which decreases in ρ in the left panel of Figure 1 because a_t and a_{t+1} enter equation (22) for the stochastic discount factor with opposite signs. Higher autocovariance in technology therefore

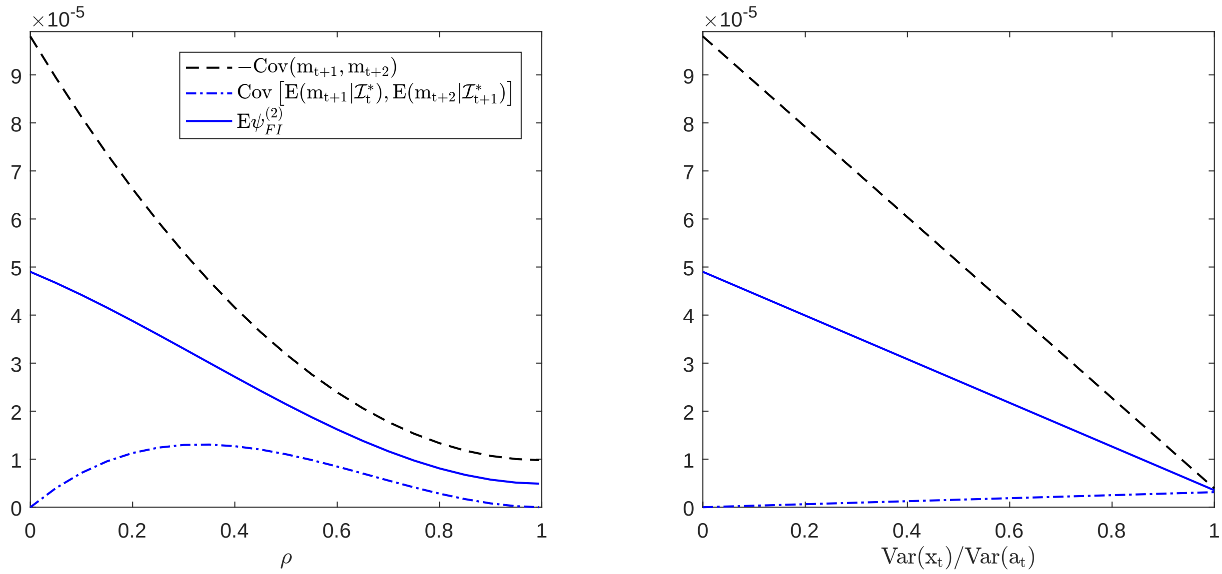

Figure 1: Components of the mean real term premium ($n = 2$) with full information (FI). $\beta = 0.99$ and $\text{Var}(a_t) = 0.01^2$. Left panel: $\text{Var}(x_t)/\text{Var}(a_t) = 0.9$. Right panel: $\rho = 0.8$.

translates into lower autocovariance in realised stochastic discount factors. The component of the real term premium that depends on expectations first rises with ρ as the autocovariance of the persistence component of technology increases. However, it eventually falls as a higher ρ reduces the extent to which that autocovariance is loaded into expectations of the stochastic discount factor by equation (24). The largest contribution from the expectations component comes at the value of ρ that makes technology sufficiently persistent yet still tracked by expectations. At the extremes expectations make no contribution to the real term premium, at $\rho = 0$ because there is no persistence in the model and at $\rho = 1$ as a random walk in technology removes all persistence from the stochastic discount factor. The expectations component increases with the relative contribution of the persistent component to technology in the right panel of Figure 1, although its contribution remains small and dominated by the component that depends on realisations.

4.3 Partial information

The first relaxation of the full information benchmark supposes that the household knows current and past values of technology but does not observe its decomposition into transitory and persistent components. Formally, the representative household's information set $\mathcal{I}'_t$ contains m^t and a^t but there is no period $s \in \{0, 1, \ldots, t\}$ such that $x_s \in \mathcal{I}'_t$ or $\eta_s \in \mathcal{I}'_t$.

The household forms expectations over future stochastic discount factors according to

$$\text{E}(m_{t+1}|\mathcal{I}'_t) = \beta\left(1 + a_t - \rho \text{E}(x_t|\mathcal{I}'_t)\right) \tag{27}$$

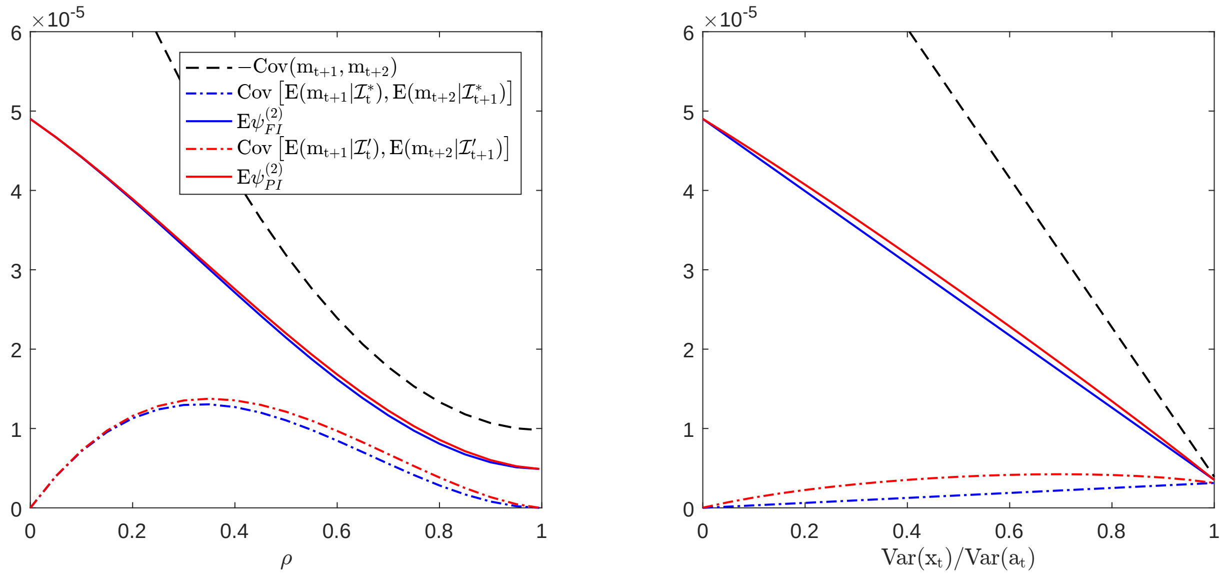

Figure 2: Components of the mean real term premium ($n = 2$) with full information (FI) and partial information (PI). $\beta = 0.99$ and $\mathrm{Var}(a_t) = 0.01^2$. Left panel: $\mathrm{Var}(x_t)/\mathrm{Var}(a_t) = 0.9$. Right panel: $\rho = 0.8$.

which requires them to infer the fraction of current technology that comes from its persistent component. This is a standard signal extraction problem, requiring an estimate of the state x^t from a sequence of noisy signals a^t. The solution satisfies the Kalman filter recursion

$$\rho \mathrm{E}(x_t|\mathcal{I}'_t) = \rho \mathrm{E}(x_t|\mathcal{I}'_{t-1}) + K_t \left(a_t - \mathrm{E}(x_t|\mathcal{I}'_{t-1})\right) \tag{28}$$

$$\Sigma_{t+1} = \rho(\rho - K_t)\Sigma_t + \sigma_\varepsilon^2 \tag{29}$$

$$K_t = \frac{\rho \Sigma_t}{\Sigma_t + \sigma_\eta^2} \tag{30}$$

where $\Sigma_t \equiv \mathrm{Var}(x_t|\mathcal{I}'_{t-1})$ and K_t is the Kalman gain. We further assume that $x_0 \sim N(0, \Sigma)$, in which case Σ_t and K_t are constant for all t and the resulting autocovariance of expected stochastic discount factors can be calculated numerically using Monte Carlo methods. The autocovariance of realised stochastic discount factors is independent of the household's information set, so identical to that under full information. What these calculations imply for the real term premium and its decomposition is shown in Figure 2.

If the majority of the volatility in technology comes from its persistent component then a_t is a precise signal about x_t and expectations formed using the Kalman filter are close to those with full information. As can be seen from the left panel of Figure 2 drawn for $\mathrm{Var}(x_t)/\mathrm{Var}(a_t) = 0.9$, the expectations component of the real term premium under $\mathcal{I}'_t$ is practically identical to that under the full information set $\mathcal{I}^*_t$ at all levels of ρ. The right panel of Figure 2 shows a marginally larger expectations component at intermediate levels of $\mathrm{Var}(x_t)/\mathrm{Var}(a_t)$ when ρ is fixed at 0.8, but requiring the household to decompose technology into transitory and persistent components only has a very weak impact on the real term premium.

4.4 Noisy information

Adding noise to the household's signal extraction problem does not generally result in larger term premia in our model. To see this, we change the informational assumptions so that the representation agent only observes noisy signals of current and past values of technology when making decisions that determine the real term premium. A household with complete knowledge of the structure of the economy could in principle invert the equilibrium relationships to infer the current state of technology even if it is not directly observable. To prevent this we assume that the household is unable to do so, or at least only able to do so up to an error term reflected in the noise contained in their signal.[12] As a result, the household takes positions in financial markets until the real term premium on all maturities of bonds is arbitrage-free with respect to the noisy signals in its information set.

The signal s_t has the form $s_t = a_t + \xi_t$ with noise $\xi_t \sim N(0, \sigma_\xi^2)$ that may correlate with the transitory component of technology according to $\sigma_{\eta\xi} \equiv \text{Cov}(\eta_t, \xi_t)$. $\mathcal{I}_t''$ is the information set of the representative household, defined such that $m^t, s^t \subset \mathcal{I}_t''$ but with no period $s \in \{0, 1, \dots, t\}$ such that $a_s \in \mathcal{I}_t''$, $x_s \in \mathcal{I}_t''$ or $\eta_s \in \mathcal{I}_t''$. The conditional expectation of m_{t+1} is

$$\text{E}\left(m_{t+1}|\mathcal{I}_t''\right) = \beta\left(1 + (1-\rho)\text{E}\left(x_t|\mathcal{I}_t''\right)\right) \tag{31}$$

and the representative household has to infer the persistent component of technology from a noisy signal, as before when it was unobservable. The solution satisfies the same Kalman filter recursion defined by equations (28) and (29) in Section 4.3, only now the signal $s_t = x_t + \eta_t + \xi_t$ has noise $\eta_t + \xi_t \sim N(0, \sigma_\eta^2 + \sigma_\xi^2 + 2\sigma_{\eta\xi})$ so instead of (30) the Kalman gain is

$$K_t^s = \frac{\rho\Sigma_t^s}{\Sigma_t^s + \sigma_\eta^2 + \sigma_\xi^2 + 2\sigma_{\eta\xi}} \tag{32}$$

with $\Sigma_t^s \equiv \text{Var}(x_t|\mathcal{I}_{t-1}'')$. What this implies for the real term premium is illustrated in Figure 3, drawn for $\text{Var}(\xi_t) = \text{Var}(a_t)/2$ and $\sigma_{\eta\xi} = 0$.

Assuming that technology is observed through a noisy signal reduces the expectations component of the real term premium in both panels of Figure 3, most notably at intermediate values of the persistence parameter ρ. Noise has the potential to inject additional volatility and persistence into estimates of the persistent component of technology, but any effect on the autocovariance of expected stochastic discount factors is offset by expectations no longer directly reacting to realisations of technology. Expectations of the stochastic discount factor react to $a_t - \rho x_t$ with full information by equation (24) or $a_t - \rho\text{E}(x_t|\mathcal{I}_t')$ with partial information by equation (27). But when there is noisy information they track $(1-\rho)\text{E}\left(x_t|\mathcal{I}_t''\right)$ by equation (31).

[12] This would be an optimal outcome if the agent were faced with an intrinsic capacity constraint in the spirit of Sims (2003).

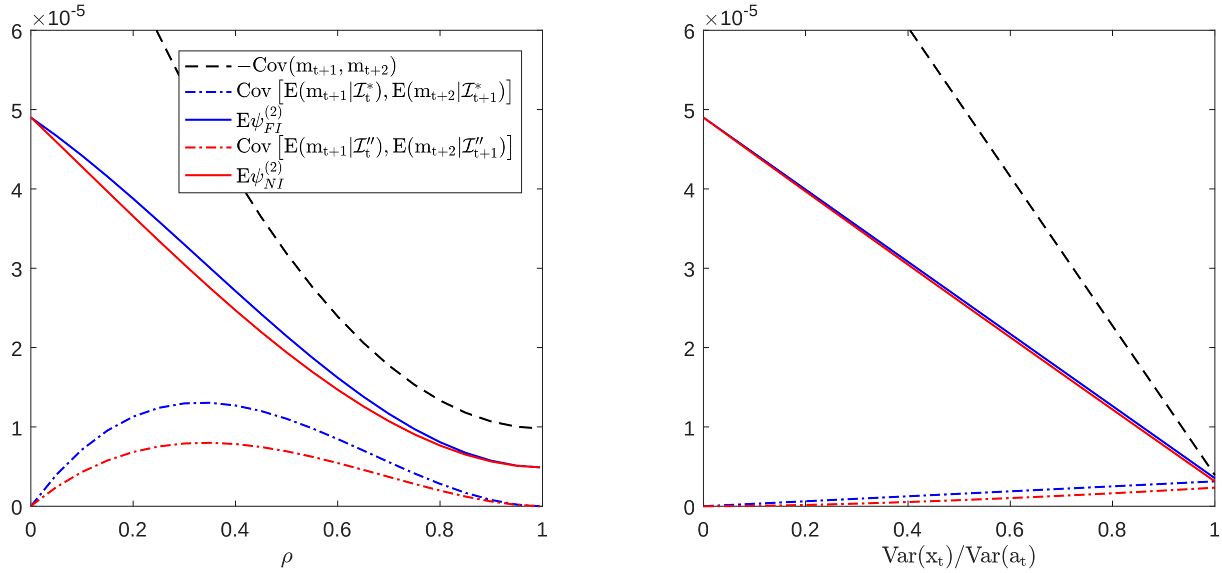

Figure 3: Components of the mean real term premium ($n = 2$) with full information (FI) and noisy information (NI). $\beta = 0.99$, $\mathrm{Var}(a_t) = 0.01^2$, $\sigma_\xi^2 = \mathrm{Var}(a_t)/2$ and $\sigma_{\eta\xi}^2 = 0$. Left panel: $\mathrm{Var}(x_t)/\mathrm{Var}(a_t) = 0.9$. Right panel: $\rho = 0.8$.

The absence of a direct effect of a_t reduces the autocovariance of expectations and explains the fall in the expectations component of the term premium. If the aim is to generate term premia that match estimates from the data then adding noise in this way is counterproductive.

The finding that noise attenuates real term premia extends to the general case in which the representative household chooses the informational content of the noisy signals they observe, subject to a limited information processing capacity. Luo and Young (2014) show that rational inattention and signal extraction problems are isomorphic when the task is to extract an estimate of the persistent component of an exogenous process in a linear-quadratic Gaussian framework. Their results apply to our model, so our findings extend to the general case under rational inattention and we conclude that the expectations component of the real term premium is dampened by noise, however that noise is introduced.

4.5 Beauty contests

The inability of either partial or noisy information to engender the autocovariance in expectations required to support meaningful real term premia suggests a need to adopt more radical informational assumptions. In this section we do just that by introducing private information and assuming that households demand term premia on the basis of heterogeneous forecasts of technology that are conditioned to be similar to the forecasts of other households. Our setup mirrors the classic Keynesian beauty contest, where to predict the winner it is necessary to identify not only the prettiest contestant but also who other people think is the prettiest. This coordinates households in our model as they form expectations of the expectations of others.

4.5.1 Information

There is a continuum of *ex ante* identical households indexed by $i \in [0,1]$. As in the previous section, households do not observe the state of technology directly and are assumed to be unable to infer it from equilibrium relationships. Each household receives a signal $s_{i,t}$ about current technology that has public and private noise components. Common (public) noise n_t and idiosyncratic (private) noise $n_{i,t}$ follow mean-zero AR(1) processes with common persistence parameter ρ and respective innovations $\xi_t \sim N(0, \sigma_\xi^2)$ and $\zeta_{i,t} \sim N(0, \sigma_\zeta^2)$. The signal is

$$s_{i,t} = \ a_t + n_t + n_{i,t} = \ x_t + n_t + n_{i,t} + \eta_t \tag{33}$$

where the sum of all persistent components $x_{i,t}^n \equiv x_t + n_t + n_{i,t}$ evolves as

$$x_{i,t}^n = \rho x_{i,t-1}^n + \varepsilon_{i,t}^n \tag{34}$$

with $\varepsilon_{i,t}^n \equiv \varepsilon_t + \xi_t + \zeta_{i,t} \sim N(0, \sigma_\varepsilon^2 + \sigma_\xi^2 + \sigma_\zeta^2)$.[13]

The information set $\mathcal{I}_{i,t}$ of household i is such that $m_i^t, s_i^t \subset \mathcal{I}_{i,t}$, where $m_{i,t}$ is household i's stochastic discount factor, but with no period $s \in \{0, 1, \ldots, t\}$ in which $a_s \in \mathcal{I}_{i,t}$, $x_s \in \mathcal{I}_{i,t}$, $n_s \in \mathcal{I}_{i,t}$ or $n_{i,s} \in \mathcal{I}_{i,t}$. Since it is common knowledge that dispersion in beliefs about the current state of technology must be entirely a result of noise, heterogeneity in beliefs does not translate into heterogeneity in consumption so that $m_{i,t} = m_t \ \forall \ i$ in equilibrium. Section 4.5.5 expands on this argument. We assume that the transitory component of technology η_t is observable, following the negligible effect of partial information on the term structure in Section 4.3. This makes the sum of persistent components $x_{i,t}^n$ observable too and $\eta^t, x_i^{n,t} \subset \mathcal{I}_{i,t}$. The setting is otherwise identical to before.

4.5.2 Strategic complementarity

The expectations component of the term premium demanded by household i depends as before on the autocovariance of their expectations of stochastic discount factors. The equilibrium stochastic discount factor for household i is analogous to that in equation (22), so the expectation of its stochastic discount factor $\hat{\mathrm{E}}\left(m_{t+1}|\mathcal{I}_{i,t}\right)$ is constructed as

$$\hat{\mathrm{E}}\left(m_{t+1}|\mathcal{I}_{i,t}\right) = \beta\left(1 + (1-\rho)\hat{\mathrm{E}}\left(x_t|\mathcal{I}_{i,t}\right)\right) \tag{35}$$

in period t, which highlights the role played by forecasts $\hat{\mathrm{E}}\left(x_t|\mathcal{I}_{i,t}\right)$ of the permanent component of technology. We target our strategic complementarity at this forecasting problem.[14] Strate-

[13] It is possible to relax the assumption of a common ρ across persistent components and the assumption that innovations to the persistent component are uncorrelated. We refrain from doing so as it delivers few additional insights at a considerable burden to notation and presentation.

[14] In our linear model it could equally be assumed that there are complementarities in forecasting the stochastic discount factor. We prefer to work with forecasts of technology since they have a clear empirical counterpart that

gic complementarities are introduced by assuming that household i sets their forecast of the permanent component of technology to minimise

$$(1-\omega)\mathrm{E}\left[(\hat{\mathrm{E}}\left(x_t|\mathcal{I}_{i,t}\right)-x_t)^2|\mathcal{I}_{i,t}\right]-\omega\mathrm{E}\left(\int_0^1\hat{\mathrm{E}}\left(x_t|\mathcal{I}_{j,t}\right)dj\middle|\mathcal{I}_{i,t}\right)\hat{\mathrm{E}}\left(x_t|\mathcal{I}_{i,t}\right) \tag{36}$$

where $0 \leq \omega \leq 1$ is a parameter measuring the degree of strategic complementarity. We denote the strategic forecast by $\hat{\mathrm{E}}\left(x_t|\mathcal{I}_{i,t}\right)$ to distinguish it from the rational expectation $\mathrm{E}\left(x_t|\mathcal{I}_{i,t}\right)$.

The first term in (36) gives the household an incentive to minimise the mean squared deviation of its forecast from the true value of the persistent component of technology. This reflects the household wanting to produce a forecast that best estimates its stochastic discount factor. The second term captures a strategic complementarity by rewarding the household for a forecast that has the same sign as the expected average forecast of all households. If the average is expected to be positive then the household adjusts its forecast upwards, if negative the impetus is for the household to adjust downwards. This acts to coordinate forecasts and expectations. Constructs of this type have been used to introduce strategic interactions in a variety of settings.[15] The strategic complementarity in our model could be rationalised by fears that investors suffer a liquidity shock and so need to liquidate their bond holdings within the period, in which case expectations have to be formed about the price on liquidation. Similarly, investors who forecast a favourable state of the economy when forecasts are negative on average may appear less well informed and incur higher costs for example in the acquisition of funding. A simple foundation for strategic complementarities and objective (36) based on such funding concerns appears in Appendix A.4.

4.5.3 Equilibrium signal extraction

The first order condition of household i implies that it sets its forecast according to

$$\hat{\mathrm{E}}\left(x_t|\mathcal{I}_{i,t}\right)=\mathrm{E}\left(x_t|\mathcal{I}_{i,t}\right)+\frac{\omega}{2(1-\omega)}\mathrm{E}\left(\int_0^1\hat{\mathrm{E}}\left(x_t|\mathcal{I}_{j,t}\right)dj\middle|\mathcal{I}_{i,t}\right) \tag{37}$$

which adjusts the mean squared error minimising rational expectation of x_t to account for the strategic complementarity. If $\omega = 0$, forecasts coincide with rational expectations. In Section 5, we examine whether survey data favour a strictly positive value of ω. The rational expectation of the persistent component of technology is extracted from household i's noisy signal $x^n_{i,t}$ as

$$\mathrm{E}\left(x_t|\mathcal{I}_{i,t}\right)=\left(\frac{\sigma^2_\varepsilon}{\sigma^2_\varepsilon+\sigma^2_\xi+\sigma^2_\zeta}\right)x^n_{i,t} \tag{38}$$

proves useful in disciplining estimation of the general model in Section 5.

[15] In management science, Dessein and Santos (2006) and Dessein et al. (2016) posit a function similar to ours to capture strategic complementarities in organising tasks. Haltiwanger and Waldman (1989) do likewise for strategic complementarities in production, as does Vives (2014) in a finance environment. All these applications shy away from the microfoundations of strategic complementarities, as does the microeconomic theory in Bergemann et al. (2015) and the seminal work by Morris and Shin (2002).

That the expectation is linear in the signal suggests the existence of a symmetric equilibrium in which the forecasts of all households are a linear function of their respective signals

$$\hat{\mathrm{E}}\left(x_t|\mathcal{I}_{j,t}\right) = \theta x^n_{j,t} \quad \forall j \tag{39}$$

where θ is a parameter to be determined in equilibrium, in which case household i's expectation of household j's forecast solves the complementary signal extraction problem as

$$\mathrm{E}\left(\hat{\mathrm{E}}\left(x_t|\mathcal{I}_{j,t}\right)|\mathcal{I}_{i,t}\right) = \mathrm{E}\left(\theta x^n_{j,t}|\mathcal{I}_{i,t}\right) = \theta\left(\frac{\sigma^2_\varepsilon + \sigma^2_\xi}{\sigma^2_\varepsilon + \sigma^2_\xi + \sigma^2_\zeta}\right) x^n_{i,t} \tag{40}$$

Symmetric linear equilibrium is confirmed with

$$\theta = \frac{(1-\omega)\sigma^2_\varepsilon}{(1-\omega)(\sigma^2_\varepsilon + \sigma^2_\xi + \sigma^2_\zeta) - \frac{\omega}{2}(\sigma^2_\varepsilon + \sigma^2_\xi)} \tag{41}$$

The equilibrium exists with $0 \leq \theta < \infty$ provided $2(1-\omega)/\omega > (\sigma^2_\varepsilon + \sigma^2_\xi)/(\sigma^2_\varepsilon + \sigma^2_\xi + \sigma^2_\zeta)$, which is satisfied if there is a sufficiently large private noise component σ^2_ζ in signals or only a limited degree of strategic complementarity ω. The maximal strategic complementarity the model can support is therefore bounded from above by a limit that depends on the variance of noise relative to the variance of the persistent component of technology. We restrict ourselves throughout to configurations of the model for which the condition holds.

The value of θ determines how much the household's forecasts $\hat{\mathrm{E}}\left(x_t|\mathcal{I}_{i,t}\right)$ and $\hat{\mathrm{E}}\left(m_{t+1}|\mathcal{I}_{i,t}\right)$ react to the noisy signals they observe. It is increasing and convex in ω in equilibrium, meaning that forecasts react progressively more to signals when there are more strategic complementarities. When $\omega = 0$ the household forecasts in isolation, being indifferent to the forecasts of others and having no reason to consider whether the noise in their signal comes from a common or idiosyncratic component. When $\omega > 0$ the household can no longer do this. They know that the forecasts of others react to signals that have a common noise component, so to keep their forecast in line with others they would ideally react to the common noise too. Since the common noise is unobservable and cannot be identified, the best the household can do is react more to their own noisy signal.

4.5.4 Term premium

The real term premium demanded by household i is again an average of covariances of successive realisations and expectations of the household's stochastic discount factor. Our decomposition (15) is still valid, albeit with a convexity adjustment term that arises because the law of iterated expectations does not hold fully in the beauty contest environment.[16] The strategic complementarities in the beauty contest have no effect on realisations, which continue to have

[16] The convexity adjustment term is negligible for relevant parameterisations of our model, see Appendix A.5 for full details. That the law of iterated expectations does not hold fully in beauty contests is known from the work of Allen et al. (2006).

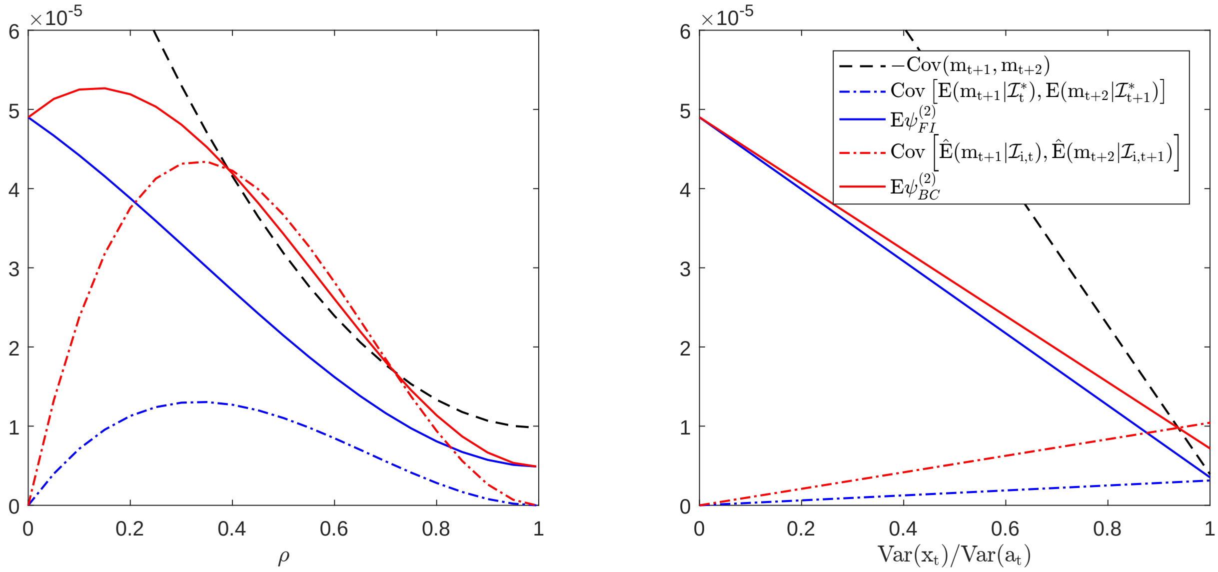

Figure 4: Components of the mean real term premium ($n = 2$) with full information (FI) and with strategic complementarities from a beauty contest (BC). $\beta = 0.99$, $\text{Var}(a_t) = 0.01^2$, $\sigma_\xi^2 = \sigma_\zeta^2 = 0.05\sigma_\varepsilon^2$ and $\omega = 0.5$. Left panel: $\text{Var}(x_t)/\text{Var}(a_t) = 0.9$. Right panel: $\rho = 0.8$.

the autocovariance calculated in Section 4.1. The autocovariance of expectations mirrors that in forecasts of the stochastic discount factor, which inherit the properties of signals by their equilibrium loading θ into the forecasts of the persistent component of technology. Combing equations (15), (23), (35), (40) and (41) delivers the unconditional autocovariance of expected stochastic discount factors

$$\text{Cov}\left[\hat{\text{E}}\left(m_{t+1}|\mathcal{I}_{i,t}\right), \hat{\text{E}}\left(m_{t+2}|\mathcal{I}_{i,t+1}\right)\right] = \beta^2 \frac{\rho(1-\rho)}{1+\rho}\theta^2(\sigma_\varepsilon^2 + \sigma_\xi^2 + \sigma_\zeta^2) \tag{42}$$

and the unconditional mean real term premium

$$\text{E}\psi_{BC}^{(2)} = \frac{1}{2}\beta^2 \left\{(1-\rho)\sigma_\varepsilon^2 + \sigma_\eta^2 + \frac{\rho(1-\rho)}{1+\rho}\left[\theta^2(\sigma_\xi^2 + \sigma_\zeta^2) - (1-\theta^2)\sigma_\varepsilon^2\right]\right\} \tag{43}$$

Figure 4 compares the mean real term premium with strategic interactions to the full information benchmark. The assumption is that $\omega = 0.5$ and that there are equal quantities of common and idiosyncratic noise that together account for 10% of the variance of signals. The way that strategic complementarities persuade households to react more to their noisy persistent signals raises the autocovariance of expectations, boosting the expectations component of the term premium. The willingness of households to countenance fluctuating but coordinated forecasts increases their demand for risk compensation and so drives up term premia.

The conditional real term premium demanded by a household depends on how much forecasts react to signals in equilibrium, but strikingly is independent of the particular signal that the household receives. The reason is the standard capital asset pricing model (CAPM) intuition that the term premium is a compensation for the risk that the household gives up consumption in a period when the marginal utility is high and receives a claim to consumption in a period when the marginal utility is low. The risk depends on what is expected to happen to the household's stochastic discount factor, and is compensated for according to our decomposition of the term premium into components that depend on the autocovariances of realisations and expectations. The autocovariance of realisations is the same for all households because they have a common stochastic discount factor. The autocovariance of expectations is a function of the volatility in expectations and how much a household expects their forecasts to persist into the future. These are identical for all households, volatility by construction and persistence because each household projects their forecasts forward using the persistence parameter that governs the persistent component in their signal. Whether a household thinks technology is currently high or low, they all project forward in the same way, the autocovariance of expectations is common across households, and the no-arbitrage real term premium is independent of the signal a household receives.

4.5.5 More on information heterogeneity

The model with heterogeneous information remains close to a representative agent model, despite there being private noise components in the signals received by each household. Our model is concordant with the conditions of the no-trade theorem in Milgrom and Stokey (1982), since the initial allocation before signals are observed is Pareto efficient and it is common knowledge under rational expectations that trades must be mutually beneficial to both parties. This means that there is no trading of bonds in equilibrium and consumption is identical for all households, which implies that we can incorporate heterogeneity in bond market expectations without having to keep track of heterogeneity in the distribution of wealth. Intuitively, whilst households' forecasts of future technology differ, the households are otherwise identical and understand that any differences in their bond valuations must be due solely to noisy signals and cannot be based on fundamentals. The only reason to trade would be to take advantage of another household, but the absence of gains from trade means that such a trade must be disadvantageous to one of the households and so cannot take place under rational expectations. Households therefore hold zero bonds in equilibrium at all maturities, as in a standard representative agent framework with bonds in zero net supply. Since bonds are not traded in equilibrium, the model with heterogeneous information cannot predict trading volumes or identify a marginal household that determines the price of a bond in the market. However, there is no heterogeneity in perceptions of second and higher moments so all households demand the same term premia, irrespective of the noisy signal they receive. There is complete agreement about the risk compensation required at each horizon, the object of our analysis.

5 A more general model

The results with the simple analytical model suggest that real term premia may be sizeable if there is a beauty contest element to the formation of expectations. In this section we show that this still applies when some of the seemingly restrictive assumptions of the simple model are relaxed. The main innovation is to allow for endogeneity in labour supply. This is done by specifying a form for household preferences that incorporates the disutility of labour and recognises that the household's coefficient of relative risk aversion is not necessarily equal to one. We further generalise by working with real term premia at horizons of up to 20 quarters and by using the exact stochastic discount factor rather than its log-linear approximation. The other features of the simple analytical model are retained, in particular that a no-trade theorem prevents non-zero bond holdings. Sections 5.1-5.4 outline the general model and present its quantitative implications through a series of numerical examples, which sets the scene for taking the model to data in Section 5.5.

5.1 Model

The economy consists of heterogeneously-informed households and a representative firm. The firm produces according to

$$y_t = A_t L_t^{1-\alpha} \tag{44}$$

and maximises profits by demanding labour L_t until the marginal product of labour is equal to the wage rate. The household receives the same noisy signal of technology as they did in the simple analytical model and has period utility

$$u(c_t, l_t) = \frac{1}{1-\sigma}\left(c_t - \chi_0 \frac{l_t^{1+\chi}}{1+\chi}\right)^{1-\sigma} \tag{45}$$

following Greenwood et al. (1988). Commonly referred to as GHH preferences, this functional form implies that labour supply is independent of wealth in equilibrium, and so provides a tractable route to modelling realistic fluctuations in hours without recourse to nominal rigidities or labour market frictions (Gertler et al., 2012).[17] The equilibrium conditions are standard. The households' stochastic discount factor can be expressed in terms of current and future technology

$$m_{t+1} = \beta \left(\frac{\gamma_1 A_{t+1}^{\gamma_2} - \frac{\chi_0}{1+\chi}\left(\frac{(1-\alpha)}{\chi_0} A_{t+1}\right)^{\gamma_2}}{\gamma_1 A_t^{\gamma_2} - \frac{\chi_0}{1+\chi}\left(\frac{(1-\alpha)}{\chi_0} A_t\right)^{\gamma_2}} \right)^{-\sigma} \tag{46}$$

[17] Schmitt-Grohé and Uribe (2012) estimate a rich DSGE model with Jaimovich and Rebelo (2009) preferences that nest GHH preferences. Results with postwar US data provide strong support for the GHH specification, using both Bayesian and maximum likelihood methods. The specification also allows us to adjust risk aversion without introducing counterfactual volatility in hours worked, see (Gertler et al., 2012).

with γ_1 and γ_2 functions of parameters of the model. The full derivation of the equilibrium conditions and the stochastic discount factor is presented in Appendix A.6.

5.2 Bond prices and valuations

The price of an n-period bond with full-information is

$$p_t^{(n)}(x_t, \eta_t) = \mathrm{E}\left(m_{t+1} p_{t+1}^{(n-1)} \middle| \mathcal{I}_t^* \right) \tag{47}$$

The equilibrium price is a function of the current levels of the persistent and transitory components of technology x_t and η_t. Expectations are conditional on the full information set $\mathcal{I}_t^*$ and taken over the joint distribution of future technology and its components. The risk-neutral price under full information is defined by

$$\tilde{p}_t^{(n)}(x_t, \eta_t) = p_t^{(1)} \mathrm{E}\left(\tilde{p}_{t+1}^{(n-1)} \middle| \mathcal{I}_t^* \right) \tag{48}$$

The valuation $p_{i,t}^{(n)}$ of an n-period bond when household i receives a noisy signal that has common and idiosyncratic noise components is

$$p_{i,t}^{(n)}(x_{i,t}^n, \eta_t) = \hat{\mathrm{E}}\left(m_{t+1} p_{i,t+1}^{(n-1)} \middle| \mathcal{I}_{i,t} \right) \tag{49}$$

It is a function of the sum of the persistent technology and noise components in their signal $x_{i,t}^n$ and the idiosyncratic component of technology η_t, both of which are observed by assumption. The expectation is conditional on household i's information set $\mathcal{I}_{i,t}$ and is defined over the joint distribution of the household's forecasts of future technology, its components, and the sum of the persistent components in the signal. The corresponding risk-neutral bond valuation is

$$\tilde{p}_{i,t}^{(n)}(x_{i,t}^n, \eta_t) = p_{i,t}^{(1)} \hat{\mathrm{E}}\left(\tilde{p}_{i,t+1}^{(n-1)} \middle| \mathcal{I}_{i,t} \right). \tag{50}$$

5.3 Computation

The equations for the pricing and valuing of bonds in Section 5.2 have a recursive structure that enables us to compute term premia up to any desired horizon. Our algorithm has four steps.

1. The exogenous processes for technology and noise are discretised into Markov chains using standard methods.

2. The joint conditional distributions of current and future endogenous variables are discretised using a moment matching procedure.

3. Conditional real term premia are calculated up to any desired maturity by recursively integrating over the joint conditional distributions of exogenous and endogenous variables at successively increasing horizons.

4. The unconditional real term premium is approximated by Monte Carlo simulation.

Steps 1 to 4 deliver an accurate numerical characterisation of the real term premium, provided that enough nodes are used when discretising and that the Monte Carlo simulation has converged. We find that 15 nodes and simulating for 10^6 periods is sufficient. Full details of the algorithm are presented in Appendix A.7.

5.4 Term premia

The real term premia in the general model are illustrated in Figure 5. Each panel plots the average real term premium in bond prices against maturity, at different intensities of strategic interaction $\omega \in \{0, 0.4, 0.5, 0.6\}$ and for a given calibration of risk aversion σ. The coefficient of relative risk aversion varies from $\sigma = 1$ in panel a) through to $\sigma = 4$ in panel d). In red are the corresponding average real term premia under the full information benchmark.

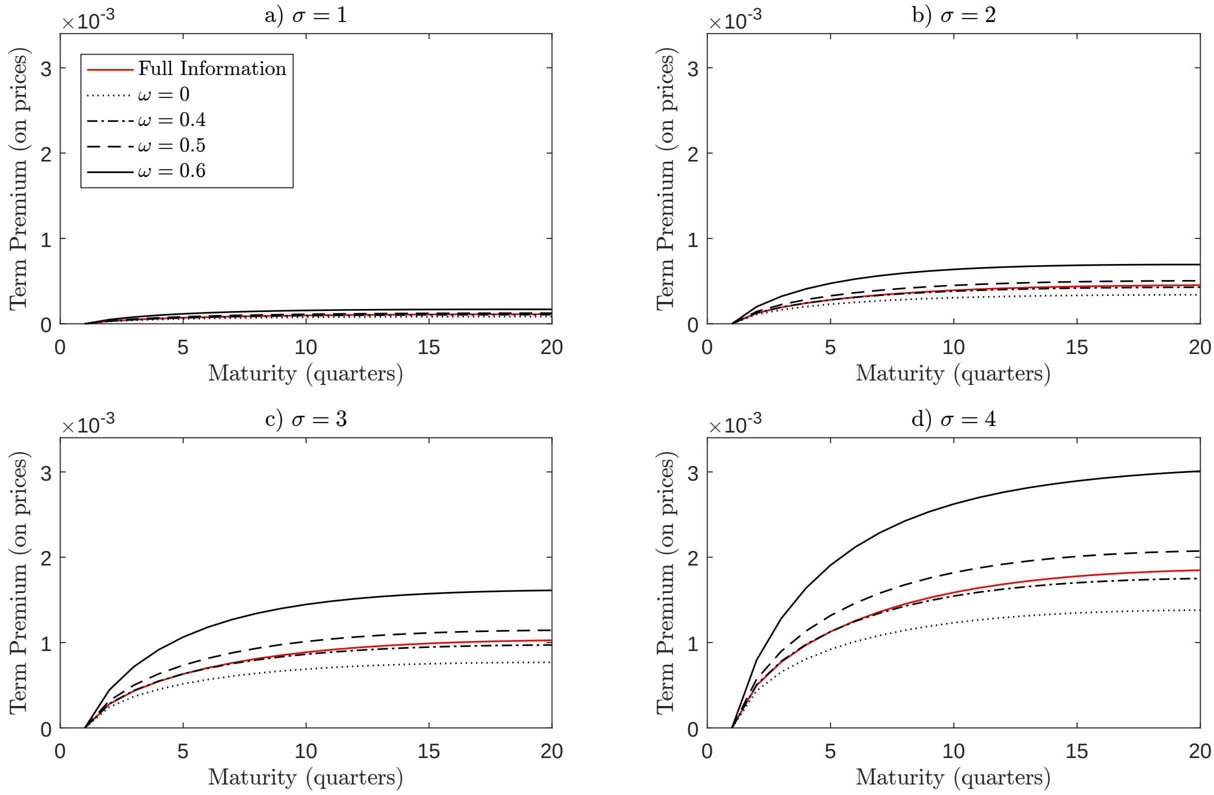

Figure 5: Real term premia with different strategic complementarities and risk aversion

The model is parameterised as in Section 4.5 where possible, so the discount factor β is 0.99, the overall volatility in technology $\text{Var}(a_t)$ is 0.01^2, the degree of persistence ρ in the persistent component x_t of the technology process is 0.8, the variances of the two components of technology σ_ε^2 and σ_η^2 are such that $\text{Var}(x_t)/\text{Var}(a_t) = 0.9$, and there are equal quantities of common and idiosyncratic noise that contribute 10% to the variance of signals. The parameters not shared with the simple analytical model are the labour share of income $1-\alpha$, the inverse Frisch elasticity of labour supply χ and the utility weight on labour χ_0. These are set at standard values 0.67, 0.5 and 1.67 so that steady-state hours worked are one third of the time endowment.

The general model has an increasing and concave term structure. A higher degree of strategic complementarity ω is associated with a steeper term structure, as expected from the discussion in Section 4.5.3. Real term premia are below their full information equivalents when ω is low because the reaction of households to their signal is muted and there is only limited volatility in their forecasts. As ω rises so does the reaction coefficient θ and enough noise enters households' forecasts that real term premia rise above those in the full information case. The coefficient of relative risk aversion σ raises term premia and amplifies the effects of ω.

5.5 Quantitative analysis

The model is now confronted with US data. Our interest is in how much the model can explain average term premia at different maturities. To find out, we estimate the model and derive what its parameter estimates entail for the average real term premium at different horizons. We only use data at the shortest maturity when estimating the model, preferring instead to identify ω in the model with strategic complementarities by requiring the cross-sectional and time-series distributions of forecasts implied by the model to be consistent with those of productivity growth forecasts in the Survey of Professional Forecasters.

5.5.1 Data

The sample period for estimation is 1999q1-2017q2. The Survey of Professional Forecasters asks participants at the beginning of each year to forecast the average annual growth in labour productivity over the next ten years. The survey responses are shown in Figure 6, where the solid line is the median of the cross-sectional distribution each period and the dashed lines are the lower and upper quartiles. Not all forecasters respond every time they are asked. Over the 26 years for which data is available, the number of respondents ranges from 21 to 46 with an average of 30.5. Moments of this series are used when estimating the parameters governing forecasts in the model with strategic complementarities. We calculate moments over the whole period 1992-2017 for which we have data.[18]

Consumption is measured as the quarterly sum of real personal consumption expenditure on non-durables and services recorded by the Bureau of Economic Analysis. It is transformed into per capita terms using data on the civilian non-institutional population from the Bureau of Labour Statistics. Labour supply is taken as the quarterly average weekly hours worked in the non-farm business sector, also reported by the Bureau of Labour Statistics. Consumption and labour supply data are used to estimate the exogenous processes for technology that drive the household's stochastic discount factor.

[18] Calculating moments from 1999 to match the sample period for estimation does not lead to significant changes.

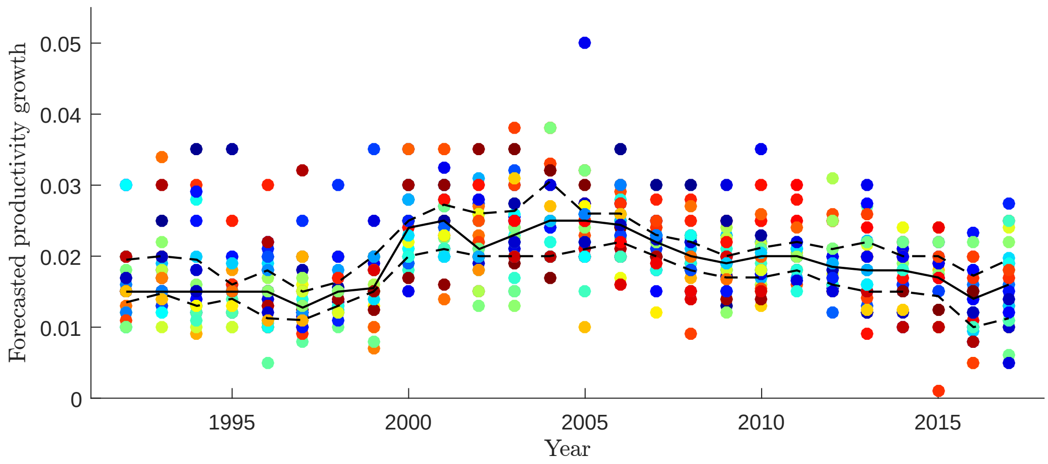

Figure 6: Forecasts of productivity growth from the Survey of Professional Forecasters. Solid line median, dashed lines lower and upper quartiles.

Data on the term premium provide a benchmark against which models can be judged. Estimates of the term premium on nominal zero-coupon Treasuries are available from Adrian et al. (2013) for maturities between one and ten years.[19] The premia estimated on yields are converted into premia on bond prices using $p_t^{(n)} = \exp(-ni_t^{(n)})$. The term premium is calculated as the annualised difference between the price of a bond and its risk-neutral counterpart. High frequency data is averaged where necessary.

5.5.2 Calibrated parameters

Table 1 lists a subset of parameters that are calibrated before estimating the model. The discount factor is set to match the steady-state annualised yield on a four-quarter bond in the model to its counterpart in the data, which we construct as the average yield on one-year zero-coupon nominal Treasuries minus the time-average of the median one-year inflation expectations from the Survey of Professional Forecasters. The mean nominal Treasury rate is 2.05% in the data, which includes an extended period when overnight rates were at the effective lower bound, and average inflation expectations were 1.91%, giving rise to a very small real discount rate that translates into a value for β of 0.9997. The labour share of income $1 - \alpha$ is fixed at 61.6%, the mean share of labour compensation in GDP in the data. The parameter χ is equal to the inverse Frisch elasticity of labour supply. It can be shown that $\text{Var}(\ln l_t)/\text{Var}(\ln c_t) = 1/(1+\chi)^2$ in equilibrium, so χ is chosen such that the relative standard deviation of hours and consumption in the model matches that in detrended data. The value that results lies in the range of common calibrations. The weight on labour disutility χ_0 is calibrated so that 1/3 of the time endowment

[19] Adrian et al. (2013) use a regression-based approach to estimate an affine term structure model with five pricing factors. Their frequently-updated estimates are available at https://www.newyorkfed.org/research/data_indicators/term_premia.html.

is spent working in steady state.[20]

Parameter	Value	Description	Target (Data)
β	0.9997	Discount factor	$i^{(4)} = 0.0205 - 0.0191$ (Treasury yields, Adrian et al. (2013), 4/1/99 - 30/6/17; Inflation expectations, SPF, 1999q1-2017q2)
α	0.384	1 - Labour share	$1 - \alpha = 0.6160$ (Share of labour compensation in GDP, Penn World Table, 1999-2014)
χ	0.708	Inverse Frisch elasticity	$\text{Var}(\ln l_t)/\text{Var}(\ln c_t) = 0.3428$ (Consumption of nondurables and services, BEA; Population and hours, BLS, 1999q1-2017q2)
χ_0	2.04	Labour utility weight	$l = 1/3$

Table 1: Calibrated parameters

5.5.3 Estimation procedure

The model parameters to be estimated relate to the exogenous processes for technology and noise, the degree of strategic complementarities and the coefficient of relative risk aversion. We estimate them using a method-of-moments procedure that fits the model to moments of the data from consumption and the Survey of Professional Forecasters. The process for technology in the model is independent of the degree of strategic complementarities or risk aversion, so its parameters can be estimated independently. The remaining parameters can then be estimated conditional on the process fitted to technology.

There are three parameters to estimate in the exogenous process for technology, the variance of the transitory component and the innovation variance and persistence of the persistent component. We collect these in the vector $\Phi_1 = (\rho \ \sigma_\eta \ \sigma_\varepsilon)$. The properties of the technology process are inherited by consumption through an equilibrium relationship that only depends on calibrated parameter values, hence we can estimate the parameters of interest from the moments of the detrended consumption data. Our estimator targets the variance and first two autocovariances of detrended log consumption and the variance of detrended log consumption growth.[21] We are able to match the moments almost perfectly with a computationally efficient procedure that exploits closed form expressions to calculate the moments of log consumption in the model.[22]

[20] The share of labour compensation in GDP is taken from the Penn World Table 9.0 and has declined from around 64% at the beginning of the sample to values close to 60% from 2009 onwards. We detrend when necessary using the Hodrick-Prescott filter with $\lambda = 1600$ for quarterly data.

[21] Adding the second autocovariance is required to achieve separate identification of ρ and σ_ε.

[22] The equilibrium relationship is $\ln c_t = \ln c_{ss} + (\chi + 1)/(\chi + \alpha)a_t$ from Appendix A.6.

Having estimated Φ_1 we proceed to estimate $\Phi_2 = (\omega\ \sigma_\xi\ \sigma_\zeta\ \sigma)$, the innovation variances of common and idiosyncratic noise, the degree of strategic complementarity, and the coefficient of relative risk aversion. Three moments from the Survey of Professional Forecasters are targeted. The first two are the variance and autocovariance of the median forecast $\hat{\gamma}_t^{50}$ of technology growth over the next ten years, which proves highly informative for the estimates of ω, σ_ξ and σ_ζ because these parameters affect the amount of noise incorporated in forecasts. The third moment to target is the mean interquartile range of forecasts $\hat{\gamma}_t^{75} - \hat{\gamma}_t^{25}$, which helps in allocating the volatility in noise into its idiosyncratic and common components.

The moments from the Survey of Professional Forecasters are not informative about the coefficient of relative risk aversion. We therefore target one additional moment, the mean term premium on a one-year nominal Treasury.[23] This does *not* mean that we are targeting the whole term structure in estimation. Instead, we are allowing the estimation of the coefficient of relative risk aversion to target the smallest term premium in the data. This aims at matching the term premium at a single point on the yield curve, so it is a legitimate test of the model to ask whether it can explain term premia at other points on the yield curve with longer horizons. The procedure yields an estimate of σ that is within the range typically considered.

We use the Generalised Method of Moments to estimate Φ_1.[24] The moments are weighted by the identity matrix so our estimates minimise the sum of squared deviations of the model moments from the data. Autocorrelation robust standard errors of Newey and West (1987) are reported. Φ_2 is estimated using the Simulated Method of Moments. We calculate the data moments from a bootstrap sample and the model moments using a simulation of 500 households over 20,000 periods. The estimates are derived by minimising a weighted sum of the squared deviations between the data and model moments, repeating the steps 500 times for new bootstrap samples to obtain the distribution of parameter estimates. The weighting matrix is constructed in the standard way as the inverse of a bootstrap estimate of the variance-covariance matrix of the moments from the Survey of Professional Forecasters. It implies that less weight is placed on matching the mean interquartile range, since that particular moment is measured with less precision than the others.

5.5.4 Estimation results

The parameter estimates are presented in Table 2. The persistent component of the technology process is highly persistent, with an innovation standard deviation about twice that of the i.i.d. transitory component. The idiosyncratic and common components of noise share the same persistence as the persistent component of technology, so their volatilities are proportional to the standard deviations of their innovation variances. Idiosyncratic noise is about as volatile as

[23] The size and even the sign of the inflation risk premium are subject to debate, as we signposted in Footnote 1. In light of this, we treat all the nominal term premium on one-year Treasuries as compensation for real risk. Interpreting the data in this way is appropriate because inflation term premia are likely to be small at the short end of the yield curve.

[24] We obtain very similar distributions if we estimate the parameters by full information maximum likelihood.

the persistent component of technology; the volatility of common noise is considerably smaller. We find a role for strategic complementarities in explaining the survey data, by estimating that forecasters place a positive weight on their forecasts having the same sign as the expected average forecast of others. The model is able to match the mean term premium on a one-year nominal Treasury without excessive levels of risk aversion. Figure 7 shows bootstrapped distributions for the estimates of $\Phi_2 = (\omega \; \sigma_\xi \; \sigma_\zeta \; \sigma)$, with the median in red and 95% confidence intervals in blue. The distributions are well-behaved and the confidence intervals are tightly centred on the median estimates.

Parameter	Estimate	95% Confidence Interval	Description
ρ	0.90	$[0.81, 0.99]$	Shock persistence
σ_ε	2.0×10^{-3}	$[9.7 \times 10^{-4}, 3.1 \times 10^{-3}]$	SD of innovation to persistent component of technology
σ_η	8.0×10^{-4}	$[0, 2.4 \times 10^{-3}]$	SD of i.i.d. transitory component of technology
σ_ξ	9.9×10^{-5}	$[9.8 \times 10^{-5}, 1.0 \times 10^{-4}]$	SD of innovation to common component of noise
σ_ζ	2.2×10^{-3}	$[1.9 \times 10^{-3}, 2.5 \times 10^{-3}]$	SD of innovation to idiosyncratic component of noise
ω	0.80	$[0.78, 0.82]$	Strategic complementarity
σ	6.0	$[5.7, 6.3]$	Coefficient of relative risk aversion

Table 2: Parameter estimates

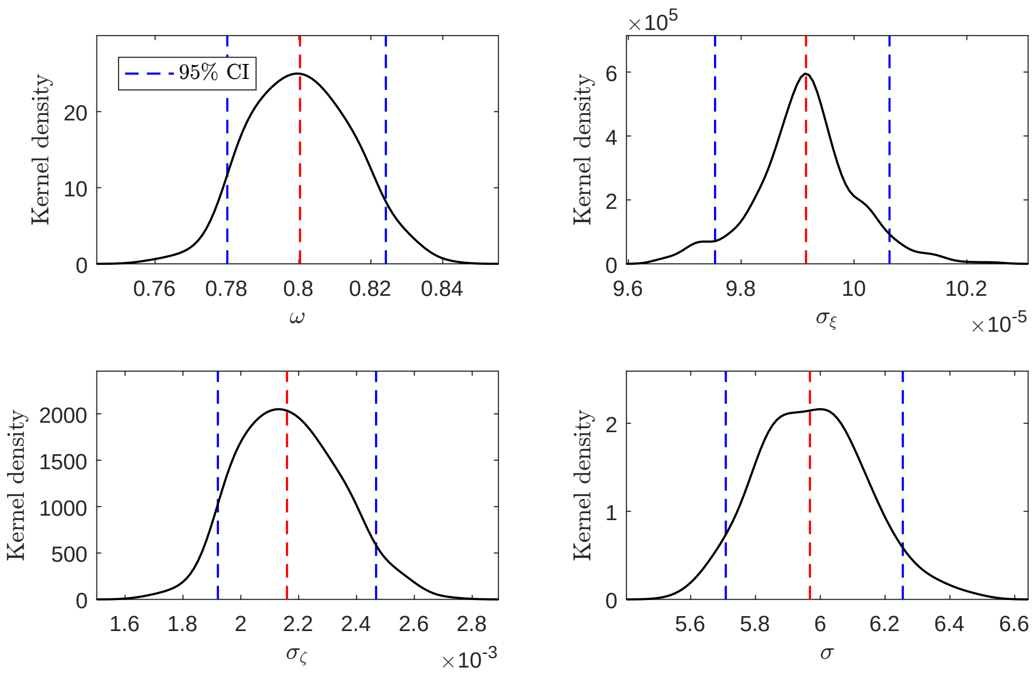

Figure 7: Bootstrapped parameter distributions

The estimated model matches the moments relating to consumption dynamics almost perfectly. The variance and first two autocovariances of log consumption are 5.5×10^{-5}, 4.8×10^{-5} and 4.3×10^{-5} respectively, compared to the target values of 5.6×10^{-5}, 5.0×10^{-5} and 4.2×10^{-5} from the data. The variance of log consumption growth is 1.4×10^{-5} in the estimated model and 1.1×10^{-5} in the data. It follows that the volatility in hours worked must match the data almost perfectly too, since the inverse Frisch elasticity is calibrated such that the relative volatility of hours and consumption is equal to that in the data.

Table 3 shows how the estimated model performs relative to other moments in the data. From the second and third columns it can be seen that it does well in capturing the moments targeted from the Survey of Professional Forecasters. All three moments line up closely with the data, as does the mean term premium on one-year nominal Treasuries.

Moment	US data 1999q1-2017q2	Estimated model	Model with full information	Model with $\omega = 0$
Targeted				
Var $\left(\hat{\gamma}_t^{50}\right)$	1.52×10^{-5}	1.42×10^{-5}	2.11×10^{-7}	4.68×10^{-8}
Cov $\left(\hat{\gamma}_t^{50}, \hat{\gamma}_{t-4}^{50}\right)$	1.25×10^{-5}	9.29×10^{-6}	1.38×10^{-7}	3.06×10^{-8}
E $\left(\hat{\gamma}_t^{75} - \hat{\gamma}_t^{25}\right)$	5.32×10^{-3}	5.40×10^{-3}	0	3.10×10^{-4}
$\mathrm{E}\psi_t^{(4)}$	8.2 bps	8.2 bps	2.6 bps	1.0 bps
Not targeted				
$\mathrm{E}\psi_t^{(8)}$	21.2 bps	16.0 bps	5.4 bps	2.0 bps
$\mathrm{E}\psi_t^{(12)}$	34.5 bps	21.1 bps	7.6 bps	2.7 bps
$\mathrm{E}\psi_t^{(16)}$	46.7 bps	24.4 bps	9.3 bps	3.3 bps
$\mathrm{E}\psi_t^{(20)}$	57.2 bps	26.7 bps	10.7 bps	3.7 bps

Table 3: Moments of the data and the models

The final four rows in Table 3 assess the fit of the model to moments that are *not* targeted in estimation. The model delivers an annualised real term premium of 16.0, 21.1, 24.4 and 26.7 bps at 2, 3, 4 and 5 year horizons, between 47% and 75% of the corresponding annualised term premia in the data. To understand the success in this dimension we add two more models in the fourth and fifth columns of Table 3. The model with full information takes the estimates from Table 2 and assumes that households observe the current level of technology and its decomposition into persistent and transitory components. Making technology observable renders estimates of the noise processes and the degree of strategic complementarities irrelevant. The model with $\omega = 0$ also adopts the parameter estimates in Table 2, but abstracts from strategic complementarities by removing the incentive for forecasters to coordinate their forecasts. The informational assumptions are otherwise as in the estimated model.

The term premia in the full information and $\omega = 0$ models fall badly short of those in the data and the estimated model. The level of the term premia can be brought closer to the data by estimating the alternative models rather than taking the estimates from Table 2, but doing so comes at the cost of significantly higher estimates for the coefficient of relative risk aversion. The reason why the full information and $\omega = 0$ models fail is their inability to generate sizeable variance and autocovariance in forecasts of technology. Forecasts are anchored to technology in the full information model and to rational expectations in the $\omega = 0$ model, neither of which leaves much scope for a large expectations component in term premia. The variances and autocovariances in these models are at least two orders of magnitude less than in the Survey of Professional Forecasters. It is only in the estimated model with strategic complementarities that movements in forecasts are sufficient to rationalise the term premia observed at longer horizons.

6 Conclusion

Our findings stress the importance of informational assumptions in models of the term structure of interest rates, as starkly visible in Figure 8 when comparing the term structure in our models to estimates from US data. The estimated model exactly matches the term premium at the short end since it is targeted in estimation, but in doing so the data prefers strategic complementarities over an unrealistically high coefficient of relative risk aversion. With these strategic complementarities in place, the estimated model can explain a significant proportion of term premia at all maturities, even though they have not been targeted. The full information model is singularly unable to do this with realistic levels of risk aversion, failing to deliver sizeable term premia.

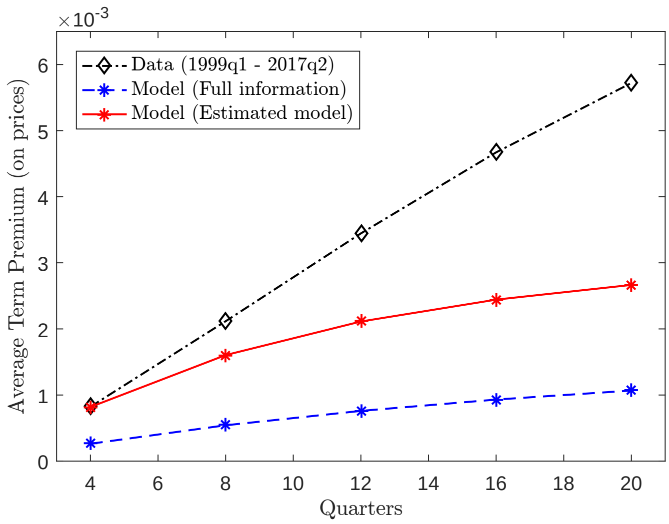

Figure 8: Term premia in the data and in the models

The success of the estimated model is driven by a beauty contest in forecasting, which rewards households for being not only accurate but also close to the average forecast of others. Even though we discipline the strength of the beauty contest by requiring the distribution of forecasts in the model to be consistent with that in the Survey of Professional Forecasters, we are still able to justify almost half of the term premium in the data at the 5-year horizon. The remaining half may be due to inflation risk premia or stronger strategic complementarities than can be supported in equilibrium in our models. Another possibility is that trading frictions create a liquidity spread that increases with maturity, as long bonds are likely to be traded over their lifetime more often than short bonds. Kozlowski (2018) estimates that the liquidity premium on corporate bonds issued in the US increases by 5 basis points for each year of their maturity. If US Treasuries are subject to similar trading frictions then liquidity premia could account for a significant proportion of the remaining term premium.

References

Abel, A. B. (1999). Risk premia and term premia in general equilibrium. *Journal of Monetary Economics*, 43(1):3–33.

Adrian, T., Crump, R. K., and Moench, E. (2013). Pricing the term structure with linear regressions. *Journal of Financial Economics*, 110(1):110–138.

Albuquerque, R., Eichenbaum, M., Luo, V. X., and Rebelo, S. (2016). Valuation Risk and Asset Pricing. *Journal of Finance*, 71(6):2861–2903.

Allen, F., Morris, S., and Shin, H. S. (2006). Beauty Contests and Iterated Expectations in Asset Markets. *Review of Financial Studies*, 19(3):719–752.

Andrade, P., Crump, R. K., Eusepi, S., and Moench, E. (2016). Fundamental disagreement. *Journal of Monetary Economics*, 83(C):106–128.

Andreasen, M. (2012). On the Effects of Rare Disasters and Uncertainty Shocks for Risk Premia in Non-Linear DSGE Models. *Review of Economic Dynamics*, 15(3):295–316.

Angeletos, G. and La'O, J. (2013). Sentiments. *Econometrica*, 81(2):739–779.

Angeletos, G.-M., Collard, F., and Dellas, H. (2018). Quantifying Confidence. *Econometrica*, 86(5):1689–1726.

Angeletos, G.-M. and Pavan, A. (2007). Efficient Use of Information and Social Value of Information. *Econometrica*, 75(4):1103–1142.

Bacchetta, P. and Wincoop, E. V. (2008). Higher Order Expectations in Asset Pricing. *Journal of Money, Credit and Banking*, 40(5):837–866.

Backus, D. K., Gregory, A. W., and Zin, S. E. (1989). Risk premiums in the term structure : Evidence from artificial economies. *Journal of Monetary Economics*, 24(3):371–399.

Bansal, R. and Yaron, A. (2004). Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles. *Journal of Finance*, 59(4):1481–1509.

Barillas, F., Hansen, L. P., and Sargent, T. J. (2009). Doubts or variability? *Journal of Economic Theory*, 144(6):2388–2418.

Barillas, F. and Nimark, K. P. (2018). Speculation and the bond market: An empirical no-arbitrage framework. *Management Science (forthcoming)*.

Barro, R. J. (2006). Rare Disasters and Asset Markets in the Twentieth Century. *The Quarterly Journal of Economics*, 121(3):823–866.

Benhabib, J., Wang, P., and Wen, Y. (2015). Sentiments and Aggregate Demand Fluctuations. *Econometrica*, 83:549–585.

Bergemann, D., Heumann, T., and Morris, S. (2015). Information and volatility. *Journal of Economic Theory*, 158(B):427–465.

Campbell, J. Y. and Cochrane, J. H. (1999). By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior. *Journal of Political Economy*, 107(2):205–251.

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.

Cogley, T. and Sargent, T. J. (2008). The market price of risk and the equity premium: A legacy of the Great Depression? *Journal of Monetary Economics*, 55(3):454–476.

Collard, F., Mukerji, S., Sheppard, K., and Tallon, J.-M. (2018). Ambiguity and the historical equity premium. *Quantitative Economics*, 9(2):945–993.

Constantinides, G. M. (1990). Habit Formation: A Resolution of the Equity Premium Puzzle. *Journal of Political Economy*, 98(3):519–543.

Croce, M. M. (2014). Long-run productivity risk: A new hope for production-based asset pricing? *Journal of Monetary Economics*, 66(C):13–31.

Dessein, W., Galeotti, A., and Santos, T. (2016). Rational Inattention and Organizational Focus. *American Economic Review*, 106(6):1522–1536.

Dessein, W. and Santos, T. (2006). Adaptive Organizations. *Journal of Political Economy*, 114(5):956–994.

Epstein, L. G. and Zin, S. E. (1989). Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework. *Econometrica*, 57(4):937–969.

Fuerst, T. S. (2015). Monetary policy and the term premium. *Journal of Economic Dynamics & Control*, 52(C):1–10.

Gertler, M., Kiyotaki, N., and Queralto, A. (2012). Financial crises, bank risk exposure and government financial policy. *Journal of Monetary Economics*, 59:S17 – S34.

Greenwood, J., Hercowitz, Z., and Huffman, G. W. (1988). Investment, Capacity Utilization, and the Real Business Cycle. *American Economic Review*, 78(3):402–417.

Gürkaynak, R. S. and Wright, J. H. (2012). Macroeconomics and the Term Structure. *Journal of Economic Literature*, 50(2):331–367.

Haltiwanger, J. and Waldman, M. (1989). Limited Rationality and Strategic Complements: The Implications for Macroeconomics. *Quarterly Journal of Economics*, 104(3):463–483.

Hansen, L. P. and Jagannathan, R. (1991). Implications of Security Market Data for Models of Dynamic Economies. *Journal of Political Economy*, 99(2):225–262.

Hördahl, P. and Tristani, O. (2012). Inflation Risk Premia in the Term Structure of Interest Rates. *Journal of European Economic Association*, 10(3):634–657.

Jaimovich, N. and Rebelo, S. (2009). Can News about the Future Drive the Business Cycle? *American Economic Review*, 99(4):1097–1118.

Jermann, U. J. (1998). Asset pricing in production economies. *Journal of Monetary Economics*, 41(2):257–275.

Kasa, K., Walker, T. B., and Whiteman, C. H. (2014). Heterogeneous Beliefs and Tests of Present Value Models. *Review of Economic Studies*, 81(3):1137–1163.

Kozlowski, J. (2018). Long-Term Finance and Investment with Frictional AssetMarkets. *Federal Reserve Bank of St Louis Working Paper*, 2018-012B.

Kung, H. (2015). Macroeconomic linkages between monetary policy and the term structure of interest rates. *Journal of Financial Economis*, 115:42–57.

Luo, Y. (2010). Rational Inattention, Long-run Consumption Risk, and Portfolio Choice. *Review of Economic Dynamics*, 13(4):843–860.

Luo, Y. and Young, E. R. (2014). Signal Extraction and Rational Inattention. *Economic Inquiry*, 52(2):811–829.

Luo, Y. and Young, E. R. (2016). Long-Run Consumption Risk and Asset Allocation under Recursive Utility and Rational Inattention. *Journal of Money, Credit and Banking*, 48(2-3):325–362.

Mehra, R. and Prescott, E. C. (1985). The equity premium: A puzzle. *Journal of Monetary Economics*, 15(2):145–161.

Milgrom, P. and Stokey, N. (1982). Information, Trade and Common Knowledge. *Journal of Economic Theory*, 26(1):17–27.

Morris, S. and Shin, H. S. (2002). Social Value of Public Information. *The American Economic Review*, 92(5):1521–1534.

Nakamura, E., Steinsson, J., Barro, R., and Ursúa, J. (2013). Crises and Recoveries in an Empirical Model of Consumption Disasters. *American Economic Journal: Macroeconomics*, 5(3):35–74.

Newey, W. K. and West, K. D. (1987). A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica*, 55(3):703–708.

Piazzesi, M. and Schneider, M. (2007). Equilibrium Yield Curves. *NBER Macroeconomics Annual*, 21:389–472.

Rietz, T. A. (1988). The equity risk premium a solution. *Journal of Monetary Economics*, 22(1):117–131.

Rudebusch, G. D. and Swanson, E. T. (2008). Examining the bond premium puzzle with a DSGE model. *Journal of Monetary Economics*, 55:S111–S126.

Rudebusch, G. D. and Swanson, E. T. (2012). The Bond Premium in a DSGE Model with Long-Run Real and Nominal Risks. *American Economic Journal: Macroeconomics*, 4(1):105–143.

Schmitt-Grohé, S. and Uribe, M. (2012). What's News in Business Cycles. *Econometrica*, 80(6):2733–2764.

Shiller, R. J. (1984). Stock Prices and Social Dynamics. *Brookings Papers on Economic Activity*, 2:457–498.

Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics*, 50(3):665–690.

Smets, F. and Wouters, R. (2003). An estimated dynamic stochastic general equilibrium model of the euro area. *Journal of the European Economic Association*, 1(5):1123 – 1175.

Smets, F. and Wouters, R. (2007). Shocks and frictions in us business cycles: A bayesian dsge approach. *American Economic Review*, 97(3):586 – 606.

Tallarini, T. D. (2000). Risk-sensitive real business cycles. *Journal of Monetary Economics*, 45(3):507–532.

Tauchen, G. (1986). Finite state markov-chain approximations to univariate and vector autoregressions. *Economics Letters*, 20(2):177–181.

van Binsbergen, J. H., Fernández-Villaverde, J., Koijen, R. S., and Rubio-Ramírez, J. (2012). The term structure of interest rates in a DSGE model with recursive preferences. *Journal of Monetary Economics*, 59(7):634–648.

Vives, X. (2014). Strategic Complementarity, Fragility, and Regulation. *Review of Financial Studies*, 27(12):3547–3592.

Weil, P. (1989). The equity premium puzzle and the risk-free rate puzzle. *Journal of Monetary Economics*, 24(3):401–421.

Appendix

A.1 Recursive formulation of bond prices

The consumption Euler equation for pricing a two-period bond (3) expands to

$$p_t^{(2)} = \mathrm{E}_t\left(m_{t+1}\mathrm{E}_{t+1}m_{t+2}\right) \tag{51}$$

The assumption that the stochastic discount factor is in the contemporaneous information set implies that $m_{t+1} \in \mathcal{I}_{t+1}$ and means the price satisfies $p_t^{(2)} = \mathrm{E}_t\left(\mathrm{E}_{t+1}m_{t+1}m_{t+2}\right)$. Imposing the law of iterated expectations then defines the two-period bond price as

$$p_t^{(2)} = \mathrm{E}_t\left(m_{t+1}m_{t+2}\right) \tag{52}$$

The general form for pricing bonds of longer horizons (5) is derived by applying the same steps to the bond pricing equation (3) for successively higher n.

A.2 Derivation of the decomposition for two-period bonds

The key step in simplifying the decomposition is to show that the one-step ahead forecast error in the stochastic discount factor $\mathrm{E}_{t+1}m_{t+2} - m_{t+2}$ is orthogonal to the stochastic discount factor m_{t+1}, in which case the second covariance term in (14) is zero. To begin, expand the second covariance term to

$$\mathrm{Cov}_t\left(m_{t+1}, \mathrm{E}_{t+1}m_{t+2} - m_{t+2}\right) = \mathrm{E}_t\left[m_{t+1}\left(\mathrm{E}_{t+1}m_{t+2} - m_{t+2}\right)\right] - \left(\mathrm{E}_t m_{t+1}\right)\mathrm{E}_t\left(\mathrm{E}_{t+1}m_{t+2} - m_{t+2}\right)$$

With $m_{t+1} \in \mathcal{I}_{t+1}$ we have

$$= \mathrm{E}_t\mathrm{E}_{t+1}m_{t+1}m_{t+2} - \mathrm{E}_t m_{t+1}m_{t+2} - \left(\mathrm{E}_t m_{t+1}\right)\mathrm{E}_t\mathrm{E}_{t+1}m_{t+2} + \left(\mathrm{E}_t m_{t+1}\right)\mathrm{E}_t m_{t+2}$$

which by the law of iterated expectations is zero, as required. Applying the law of total covariance to the conditional term premium (14) implies

$$\mathrm{E}\psi_t^{(2)} = \frac{1}{2}\left[-\mathrm{Cov}(m_{t+1}, m_{t+2}) + \mathrm{Cov}(\mathrm{E}_t m_{t+1}, \mathrm{E}_t m_{t+2})\right]$$

By the definition of conditional covariance and the law of iterated expectations

$$\begin{aligned}\mathrm{Cov}\left(\mathrm{E}_t m_{t+1}, \mathrm{E}_{t+1}m_{t+2}\right) &= \mathrm{E}\mathrm{Cov}_t\left(\mathrm{E}_t m_{t+1}, \mathrm{E}_{t+1}m_{t+2}\right) + \mathrm{Cov}\left(\mathrm{E}_t\mathrm{E}_t m_{t+1}, \mathrm{E}_t\mathrm{E}_{t+1}m_{t+2}\right)\\ &= \mathrm{Cov}\left(\mathrm{E}_t m_{t+1}, \mathrm{E}_t m_{t+2}\right)\end{aligned} \tag{53}$$

and the unconditional term premium is given by equation (15).

A.3 Generalisation of the decomposition

Begin by expanding the price of an n-period bond (5)

$$
\begin{aligned}
p_t^{(n)} &= \mathrm{E}_t \prod_{j=1}^{n} m_{t+j} \\
&= (\mathrm{E}_t m_{t+1})(\mathrm{E}_t m_{t+2}) \dots (\mathrm{E}_t m_{t+n}) + \mathrm{Cov}_t(m_{t+1}, m_{t+2} m_{t+3} \dots m_{t+n}) \\
&\quad + (\mathrm{E}_t m_{t+1}) \mathrm{Cov}_t(m_{t+2}, m_{t+3} m_{t+4} \dots m_{t+n}) \\
&\quad + (\mathrm{E}_t m_{t+1})(\mathrm{E}_t m_{t+2}) \mathrm{Cov}_t(m_{t+3}, m_{t+4} m_{t+5} \dots m_{t+n}) \\
&\quad + \dots \\
&\quad + (\mathrm{E}_t m_{t+1})(\mathrm{E}_t m_{t+2}) \dots (\mathrm{E}_t m_{t+n-2}) \mathrm{Cov}_t(m_{t+n-1}, m_{t+n}) \\
&= \prod_{j=1}^{n} \mathrm{E}_t m_{t+j} + \mathrm{Cov}_t \left(m_{t+1}, \prod_{j=2}^{n} m_{t+j} \right) + \sum_{k=1}^{n-2} \left(\prod_{j=1}^{k} \mathrm{E}_t m_{t+j} \right) \mathrm{Cov}_t \left(m_{t+k+1}, \prod_{j=k}^{n-2} m_{t+j+2} \right)
\end{aligned}
$$

Since $p_t^{(1)} = \mathrm{E}_t m_{t+1}$ from (5) it follows from (6) that $\mathrm{E}_t m_{t+1} = \exp(-i_t)$ and by the law of iterated expectations $\mathrm{E}_t m_{t+j} = \mathrm{E}_t \exp(-i_{t+j-1})$ for $j > 0$. Substituting into the above

$$
\begin{aligned}
p_t^{(n)} = &\prod_{j=0}^{n-1} \mathrm{E}_t \mathrm{e}^{-i_{t+j}} + \mathrm{Cov}_t \left(m_{t+1}, \prod_{j=2}^{n} m_{t+j} \right) \\
&+ \sum_{k=1}^{n-2} \left(\prod_{j=0}^{k-1} \mathrm{E}_t \mathrm{e}^{-i_{t+j}} \right) \mathrm{Cov}_t \left(m_{t+k+1}, \prod_{j=k}^{n-2} m_{t+j+2} \right)
\end{aligned}
$$

The risk-neutral price (7) can analogously be expanded as

$$
\begin{aligned}
\tilde{p}_t^{(n)} &= \mathrm{E}_t \prod_{j=0}^{n-1} \mathrm{e}^{-i_{t+j}} \\
&= \prod_{j=0}^{n-1} \mathrm{E}_t \mathrm{e}^{-i_{t+j}} + \mathrm{Cov}_t \left(\mathrm{e}^{-i_t}, \prod_{j=2}^{n} \mathrm{e}^{-i_{t+j-1}} \right) \\
&\quad + \sum_{k=1}^{n-2} \left(\prod_{j=0}^{k-1} \mathrm{E}_t \mathrm{e}^{-i_{t+j}} \right) \mathrm{Cov}_t \left(\mathrm{e}^{-i_{t+k}}, \prod_{j=k}^{n-2} \mathrm{e}^{-i_{t+j+1}} \right)
\end{aligned}
$$

Substituting for $\mathrm{E}_t m_{t+1} = \exp(-i_t)$

$$
\begin{aligned}
\tilde{p}_t^{(n)} = &\prod_{j=0}^{n-1} \mathrm{E}_t \mathrm{e}^{-i_{t+j}} + \mathrm{Cov}_t \left(\mathrm{E}_t m_{t+1}, \prod_{j=2}^{n} \mathrm{E}_{t+j-1} m_{t+j} \right) \\
&+ \sum_{k=1}^{n-2} \left(\prod_{j=0}^{k-1} \mathrm{E}_t \mathrm{e}^{-i_{t+j}} \right) \mathrm{Cov}_t \left(\mathrm{E}_{t+k} m_{t+k+1}, \prod_{j=k}^{n-2} \mathrm{E}_{t+j+1} m_{t+j+2} \right)
\end{aligned}
$$

The general decomposition of the per-period real term premium in bond prices (9) into components depending on realisations and expectations follows as Proposition 1 for the real term premium and Corollary 1 for the mean real term premium.

Proposition 1. *The real term premium* $\psi_t^{(n)}$ *for* $n \in \{2, 3, \ldots\}$ *is given by*

$$\psi_t^{(n)} = \frac{1}{n}\sum_{k=0}^{n-2} \iota_t(k) \left[-\mathrm{Cov}_t \left(m_{t+k+1}, \prod_{j=k}^{n-2} m_{t+j+2} \right) + \mathrm{Cov}_t \left(\mathrm{E}_{t+k} m_{t+k+1}, \prod_{j=k}^{n-2} \mathrm{E}_{t+j+1} m_{t+j+2} \right) \right]$$

where

$$\iota_t(k) \equiv \begin{cases} 1 & \textit{for } k = 0 \\ \prod_{j=0}^{k-1} E_t \mathrm{e}^{-i_{t+j}} & \textit{otherwise} \end{cases}$$

Corollary 1. *The real term premium* $\mathrm{E}\psi^{(n)}$ *for* $n \in \{2, 3, \ldots\}$ *is given by*

$$\mathrm{E}\psi^{(n)} = \frac{1}{n}\sum_{k=0}^{n-2} \Bigg\{ \mathrm{E}\left[\iota_t(k)\right] \Bigg[-\mathrm{Cov}\left(m_{t+k+1}, \prod_{j=k}^{n-2} m_{t+j+2} \right) + \mathrm{Cov}\left(\mathrm{E}_t m_{t+k+1}, \mathrm{E}_t \prod_{j=k}^{n-2} m_{t+j+2} \right) + \mathrm{Cov}\left(\mathrm{E}_{t+k} m_{t+k+1}, \prod_{j=k}^{n-2} \mathrm{E}_{t+j+1} m_{t+j+2} \right) - \mathrm{Cov}\left(\mathrm{E}_t m_{t+k+1}, \mathrm{E}_t \prod_{j=k}^{n-2} \mathrm{E}_{t+j+1} m_{t+j+2} \right) \Bigg] + \mathrm{Cov}\left[\iota_t(k), -\mathrm{Cov}_t \left(m_{t+k+1}, \prod_{j=k}^{n-2} m_{t+j+2} \right) + \mathrm{Cov}_t \left(\mathrm{E}_{t+k} m_{t+k+1}, \prod_{j=k}^{n-2} \mathrm{E}_{t+j+1} m_{t+j+2} \right) \right] \Bigg\}$$

A.4 A simple foundation of strategic complementarity in forecasting

Suppose that each period an investor requires z units of external funding to operate in financial markets. The funds are supplied inelastically by a price-setting financial intermediary that demands a rental rate $r_{i,t}^z$ from investor i, with the rental fee $r_{i,t}^z z$ rebated to the investor as a dividend at the end of the period. The financial intermediary observes the forecasts made by investors but does directly not observe technology, so can only condition $r_{i,t}^z$ on the forecast made by investor i and the distribution of forecasts made by all investors. One possibility is to set the rental rate as a function of investor i's forecast and the average forecasts of all investors.

$$r_{i,t}^z = r^z \left(\hat{\mathrm{E}}\left(x_t | \mathcal{I}_{i,t}\right), \int_0^1 \hat{\mathrm{E}}\left(x_t | \mathcal{I}_{j,t}\right) dj \right) \tag{54}$$

The way that forecasts map into the rental rate depends on the objectives and behaviour of the financial intermediary. We assume that they have an aversion to funding contrarian investors. A *contrarian* investor buys and sells financial assets in contrast to the sentiment prevailing at the time, i.e. buys when markets are bearish and sells when they are bullish.[25] To discourage investors from taking contrarian positions, the financial intermediary needs to raise the rental rate for optimistic investors when forecasts are on average pessimistic and raise the rental rate

[25] Famous invetsors with a reputation for being contrarian include Warren Buffet and Paul Tudor Jones. Less successful contrarian investors seldom make the headlines.

for pessimistic investors when forecasts are on average optimistic. A simple way to achieve this is to set

$$r^z_{i,t} = \delta_0 + \delta_1 \left(\int_0^1 \hat{\mathrm{E}}\left(x_t|\mathcal{I}_{j,t}\right) dj - x_0 \right) \left(x_0 - \hat{\mathrm{E}}\left(x_t|\mathcal{I}_{i,t}\right) \right) \tag{55}$$

where $\delta_0,\ \delta_1 > 0$ are parameters and $x_0 = \mathrm{E}x_t$ is the average forecast of investors in steady-state. If the average forecast is optimistic then $\left(\int_0^1 \hat{\mathrm{E}}\left(x_t|\mathcal{I}_{j,t}\right) dj - x_0 \right) > 0$ and the rental rate is increasing investor i's pessimism $\left(x_0 - \hat{\mathrm{E}}\left(x_t|\mathcal{I}_{i,t}\right) \right) > 0$ as required. Conversely, the rental rate is increasing in an investor's optimism when the average forecast is pessimistic.

The forecasting problem of the investor is to minimise the sum of its funding cost and the monetary cost of forecast errors.

$$\min\ \mathrm{E}\left[(\hat{\mathrm{E}}\left(x_t|\mathcal{I}_{i,t}\right) - x_t)^2 + \tilde{\omega} r^z_{i,t} z \middle| \mathcal{I}_{i,t} \right] \tag{56}$$

or

$$\min\ \mathrm{E}\left[(\hat{\mathrm{E}}\left(x_t|\mathcal{I}_{i,t}\right) - x_t)^2 \middle| \mathcal{I}_{i,t} \right] + \mathrm{E}\left\{ \tilde{\omega} \left[\delta_0 + \delta_1 \left(\int_0^1 \hat{\mathrm{E}}\left(x_t|\mathcal{I}_{j,t}\right) dj - x_0 \right) \left(x_0 - \hat{\mathrm{E}}\left(x_t|\mathcal{I}_{i,t}\right) \right) \right] z \middle| \mathcal{I}_{i,t} \right\}$$

where $\tilde{\omega} > 0$ is a scale parameter that converts both types of costs into comparable units. Noting that the unconditional mean of x_t is zero and simplifying yields

$$\min\ \mathrm{E}\left[(\hat{\mathrm{E}}\left(x_t|\mathcal{I}_{i,t}\right) - x_t)^2 \middle| \mathcal{I}_{i,t} \right] - \tilde{\omega}\delta_1 z \mathrm{E}\left(\int_0^1 \hat{\mathrm{E}}\left(x_t|\mathcal{I}_{j,t}\right) dj \middle| \mathcal{I}_{i,t} \right) \hat{\mathrm{E}}\left(x_t|\mathcal{I}_{i,t}\right) \tag{57}$$

which can be rewritten as

$$\min\ (1-\omega)\mathrm{E}\left[(\hat{\mathrm{E}}\left(x_t|\mathcal{I}_{i,t}\right) - x_t)^2 \middle| \mathcal{I}_{i,t} \right] - \omega \mathrm{E}\left(\int_0^1 \hat{\mathrm{E}}\left(x_t|\mathcal{I}_{j,t}\right) dj \middle| \mathcal{I}_{i,t} \right) \hat{\mathrm{E}}\left(x_t|\mathcal{I}_{i,t}\right), \tag{58}$$

the expression in Section 4.5.2 with $\omega \equiv \tilde{\omega}\delta_1 z/(1 + \tilde{\omega}\delta_1 z)$.

A.5 Decomposing the term premium in the beauty contest

The decomposition of the bond term premium in Section 3.3 relies on the law of iterated expectations, as explained in Appendix A.2. The law only partially holds in the beauty contest environment, in which case there is an additional convexity adjustment term in the decomposition. We derive this additional term below and show that it is negligible for relevant parameterisations of our model.

We begin by noting that the law of iterated expectations holds in the beauty contest when expectations are taken over linear terms. In equilibrium

$$\hat{\mathrm{E}}\left(\hat{\mathrm{E}}\left(x_{t+1}|\mathcal{I}_{i,t+1}\right) \middle| \mathcal{I}_{i,t} \right) = \hat{\mathrm{E}}\left(\theta x^n_{i,t+1} \middle| \mathcal{I}_{i,t} \right) = \theta\rho x^n_{i,t} = \hat{\mathrm{E}}\left(x_{t+1} \middle| \mathcal{I}_{i,t} \right)$$

and similarly for unconditional expectations

$$\mathrm{E}\hat{\mathrm{E}}\left(x_{t+1} \middle| \mathcal{I}_{i,t} \right) = \mathrm{E}\hat{\mathrm{E}}\left(\theta x^n_{i,t+1} \middle| \mathcal{I}_{i,t} \right) = 0 = \mathrm{E}x_{t+1}$$

so the law of iterated expectations holds for expectations taken over linear transforms of x_t. Since m_{t+1} is linear in x_t it also holds for linear transforms of the stochastic discount factor. The conditional term premium demanded by household i in the beauty contest (59) is analogous to equation (14) in Section 3.3.

$$\psi_{i,t}^{(2)} = -\frac{1}{2}\left[\hat{\text{Cov}}_t\left(m_{t+1}, \hat{\text{E}}\left(m_{t+2}|\mathcal{I}_{i,t+1}\right)\right)\right] \tag{59}$$

The covariance appears with a hat because it is formed in the beauty contest. The conditional covariance in (59) is

$$\begin{aligned}\hat{\text{Cov}}_t\left(m_{t+1}, \hat{\text{E}}\left(m_{t+2}|\mathcal{I}_{i,t+1}\right)\right) &= \hat{\text{E}}\left(m_{t+1}\hat{\text{E}}\left(m_{t+2}|\mathcal{I}_{i,t+1}\right)\Big|\mathcal{I}_{i,t}\right)\\ &\quad - \hat{\text{E}}\left(m_{t+1}|\mathcal{I}_{i,t}\right)\hat{\text{E}}\left(\hat{\text{E}}\left(m_{t+2}|\mathcal{I}_{i,t+1}\right)\Big|\mathcal{I}_{i,t}\right)\end{aligned}$$

from which it follows that the unconditional mean term premium can be written

$$\text{E}\psi_{i,t}^{(2)} = -\frac{1}{2}\left[\text{E}\hat{\text{E}}\left(m_{t+1}\hat{\text{E}}\left(m_{t+2}|\mathcal{I}_{i,t+1}\right)\Big|\mathcal{I}_{i,t}\right) - \text{E}\left(\hat{\text{E}}\left(m_{t+1}|\mathcal{I}_{i,t}\right)\hat{\text{E}}\left(m_{t+2}|\mathcal{I}_{i,t}\right)\right)\right]$$

To simplify this expression observe that

$$\begin{aligned}\text{Cov}\left(\hat{\text{E}}\left(m_{t+1}|\mathcal{I}_{i,t}\right), \hat{\text{E}}\left(m_{t+2}|\mathcal{I}_{i,t+1}\right)\right) &= \text{Cov}\left(\hat{\text{E}}\left(m_{t+1}|\mathcal{I}_{i,t}\right), \hat{\text{E}}\left(m_{t+2}|\mathcal{I}_{i,t}\right)\right)\\ &= \text{E}\left(\hat{\text{E}}\left(m_{t+1}|\mathcal{I}_{i,t}\right)\hat{\text{E}}\left(m_{t+2}|\mathcal{I}_{i,t}\right)\right)\\ &\quad - \text{E}\hat{\text{E}}\left(m_{t+1}|\mathcal{I}_{i,t}\right)\text{E}\hat{\text{E}}\left(m_{t+2}|\mathcal{I}_{i,t}\right)\end{aligned}$$

where the first equality follows from the same logic as (53) and the final term is equal to $-\text{E}\left(m_{t+1}\right)\text{E}\left(m_{t+2}\right) = \text{Cov}\left(m_{t+1}, m_{t+2}\right) - \text{E}m_{t+1}m_{t+2}$ by the law of iterated expectations holding for linear functions of the stochastic discount factor. After suitable algebraic manipulation, the expression for the mean unconditional term premium simplifies to

$$\text{E}\psi_{i,t}^{(2)} = \frac{1}{2}\left[-\text{Cov}\left(m_{t+1}m_{t+2}\right) + \text{Cov}\left(\hat{\text{E}}\left(m_{t+1}|\mathcal{I}_{i,t+1}\right), \hat{\text{E}}\left(m_{t+2}|\mathcal{I}_{i,t+1}\right)\right) - \tau\right]$$

with

$$\tau = \text{E}\hat{\text{E}}\left(m_{t+1}\hat{\text{E}}\left(m_{t+2}|\mathcal{I}_{i,t+1}\right)\Big|\mathcal{I}_{i,t}\right) - \text{E}m_{t+1}m_{t+2} \tag{60}$$

The convexity adjustment term τ reduces to $\text{E}\hat{\text{E}}\left(\hat{\text{E}}\left(m_{t+1}m_{t+2}|\mathcal{I}_{i,t+1}\right)\Big|\mathcal{I}_{i,t}\right) - \text{E}m_{t+1}m_{t+2}$ in case $m_{t+1} \in \mathcal{I}_{i,t+1}$. In general $\tau \neq 0$ in beauty contests because the law of iterated expectations only holds over linear functions of the stochastic discount factor.

We use (60) to confirm that deviations from the law of iterated expectations are not fundamental to our results. Substituting in for the stochastic discount factors, taking appropriate expectations and simplifying yields

$$\tau = \beta^2(1-\rho)\left\{\frac{\rho}{1+\rho}\left[\theta^2(\sigma_\varepsilon^2 + \sigma_\xi^2 + \sigma_\zeta^2) - \sigma_\varepsilon^2\right] + (1-\theta)\sigma_\varepsilon^2\right\}$$

Figure 9 plots the mean term premium with and without the convexity adjustment term τ, using the same calibration of the technology process and parameters as in Section 4.5.4. The differences for the empirically relevant values of ρ are negligible. Note that the algorithm used to calculate term premia in the quantitative beauty contest model in Section 5 fully accounts for deviations from the law of iterated expectations.

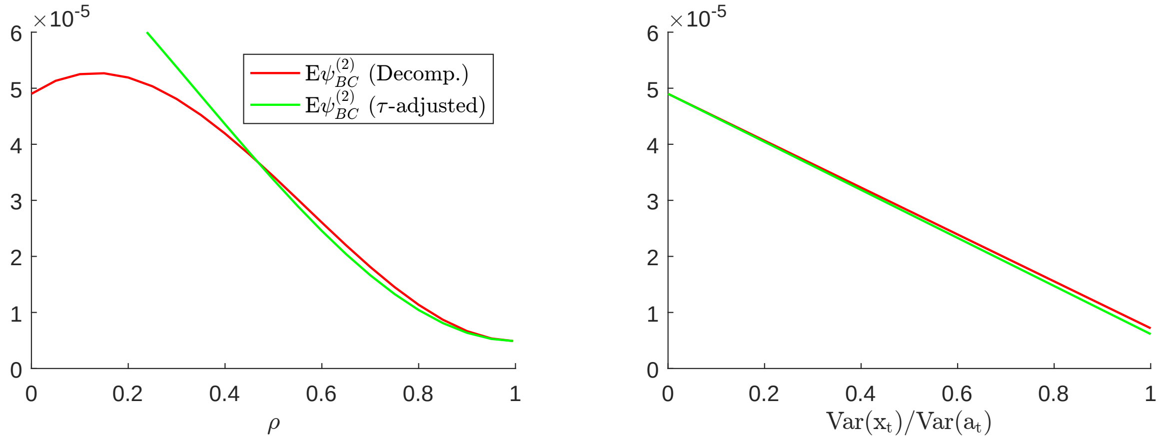

Figure 9: Term premium and law of iterated expectations

A.6 Equilibrium in the general model

The Lagrangian of the representative consumer's utility maximisation problem is

$$\mathcal{L} = \sum_{t=0}^{\infty} \beta^t \mathrm{E}_t \left(\left(u(c_t, l_t) + \lambda_t \left(w_t l_t + d_t + \sum_{n=1}^{N} p_t^{(n-1)} b_{t-1}^{(n)} - c_t - \sum_{n=1}^{N} p_t^{(n)} b_t^{(n)} \right) \right) \middle| \mathcal{I}_t^* \right)$$

where I_t^* is their information set. The first order conditions are

$$0 = u_c(c_t, l_t) - \lambda_t$$
$$0 = -\lambda_t p_t^{(n)} + \beta \mathrm{E}_t \left(\lambda_{t+1} p_{t+1}^{(n-1)} \middle| \mathcal{I}_t^* \right)$$
$$0 = -u_l(c_t, l_t) + \lambda_t w_t$$

Profit maximisation by the firm sets $w_t = (1-\alpha) A_t L_t^{-\alpha}$, bonds are in zero net supply, and all markets clear. Combining the first and third first order conditions gives $w_t = l_t^{\chi}$ and consumption drops out of the intratemporal condition for labour supply, as expected with GHH preferences. Labour market clearing defines

$$l_t = \left(\frac{(1-\alpha) A_t}{\chi_0} \right)^{\frac{1}{\chi+\alpha}}$$

Consumption is

$$c_t = \left(\frac{1-\alpha}{\chi_0}\right)^{\frac{1-\alpha}{\chi+\alpha}} A_t^{\frac{\chi+1}{\chi+\alpha}}$$

and household i's valuation of an n-period bond satisfies

$$p_{i,t}^{(n)} = \mathrm{E}_t\left(m_{t+1}p_{i,t+1}^{(n-1)}\middle|\mathcal{I}_{i,t}\right)$$

for $n \in \{1, 2, \ldots, N\}$, $p_t^{(0)} = 1$, and

$$m_{t+1} = \beta\left(\frac{\gamma_1 A_{t+1}^{\gamma_2} - \frac{\chi_0}{1+\chi}\left(\frac{(1-\alpha)}{\chi_0}A_{t+1}\right)^{\gamma_2}}{\gamma_1 A_t^{\gamma_2} - \frac{\chi_0}{1+\chi}\left(\frac{(1-\alpha)}{\chi_0}A_t\right)^{\gamma_2}}\right)^{-\sigma}$$

with $\gamma_1 \equiv ((1-\alpha)/\chi_0)^{(1-\alpha)/(\chi+\alpha)}$ and $\gamma_2 \equiv (\chi+1)/(\chi+\alpha)$.

A.7 Computational algorithm

Before describing the algorithm, it is necessary to identify the joint conditional distributions of current and future endogenous variables that households use to price the term premium. Under full information these are the usual conditional distributions under rational expectations that can be characterised by standard signal extraction results. Things are more involved when information is heterogeneous and there are strategic complementarities, because households price the term premium according to forecasts rather than rational expectations. The key determinant of forecasts is the conditional distribution of household i's forecast of the current persistent component of technology, conditional on the signal they observe. The first moment of this conditional forecast is pinned down by the linear reaction of the forecast to the signal, but our procedure for setting forecasts is silent on its conditional second moment. We therefore fix the second moment at what it would be under rational expectations.

$$x_t|x_{i,t}^n \sim N\left[\theta x_{i,t}^n, \frac{\sigma_\varepsilon^2}{1-\rho^2}\left(1 - \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2+\sigma_\xi^2+\sigma_\zeta^2}\right)\right]$$

A detailed description of steps 1 to 4 follows.

1. The univariate exogenous processes for z_t where $z \in \{x, \eta, x_i^n, \varepsilon, \xi + \zeta_i\}$ are approximated by finite-state Markov chains with $\tilde{N}$ nodes using the Tauchen (1986) procedure. The vector of nodes $\tilde{x}_z = (\tilde{x}_{z,1}, \tilde{x}_{z,2}, \ldots, \tilde{x}_{z,\tilde{N}})$ is chosen so that the extremes $\tilde{x}_{z,\tilde{N}} = -\tilde{x}_{z,1}$ cover a multiple of the standard deviation of the unconditional distribution of z_t. The remaining nodes partition $[\tilde{x}_{z,1}, \tilde{x}_{z,\tilde{N}}]$ into equispaced intervals. The transition probabilities between nodes follow from the normal distribution of z_t.

2. To discretise the key conditional distribution $x_t|x_{i,t}^n$ we adapt and adopt methods of Gauss-Hermite quadrature. Let $\Pi_{x|x^n}$ be an $\tilde{N} \times \tilde{N}$ matrix to be populated by conditional

probabilities $\left[\Pi_{x|x^n}\right]_{i,j} = \Pr(\tilde{x}_{x,j}|\tilde{x}_{x_i^n,i}) \equiv \pi_{i,j}$ over the nodes $\tilde{x}_x$ and $\tilde{x}_{x_i^n}$ calculated in Step 1. The probabilities $\pi_i \equiv (\pi_{i,1}, \pi_{i,2}, \ldots, \pi_{i,\tilde{N}})$ in row i of the matrix condition on $\tilde{x}_{x_i^n,i}$ and are chosen to minimise the Euclidean distance between the first $\tilde{N}-1$ central moments implied by the discretisation and their theoretical counterparts.

$$\min_{\pi_i} \left\| \left\{ \sum_{j=1}^{\tilde{N}} \pi_{i,j} \left(\tilde{x}_{x,j} - \theta \tilde{x}_{x_i^n,i}\right)^k \right\}_{k=0}^{\tilde{N}-1} - \left\{ \mathrm{E}_t \left((x_t - \theta \tilde{x}_{x_i^n,i})^k | x_{i,t}^n \right) \right\}_{k=0}^{\tilde{N}-1} \right\|_2 \text{ s.t. } \pi_{i,j} \in [0,1] \; \forall j$$

where the index k runs from 0 to $\tilde{N}-1$ to include the monopole zeroth moment so that probabilities sum to one. The advantage of our approach is that probabilities are calculated row-by-row, which means solving $\tilde{N}$ systems of $\tilde{N}$ equations to complete the approximation rather than the $\tilde{N}$ systems of $2\tilde{N}$ that have to be solved with standard Gaussian quadrature methods. The gain in computational efficiency allows to improve accuracy by selecting higher values of $\tilde{N}$.

3. It is easy to calculate conditional real term premia once the relevant conditional distributions have been discretised in Step 2. Household i's valuation in (x_i^n, η) space is

$$\begin{aligned} p_i^{(n)}(\tilde{x}_{x_i^n,i}, \tilde{x}_{\eta,j}) = & \sum_{k=1}^{\tilde{N}} \sum_{l=1}^{\tilde{N}} \sum_{m=1}^{\tilde{N}} \sum_{n=1}^{\tilde{N}} \pi_{i,k} \gamma_{i,l}^{\varepsilon} \gamma_{i,m}^{\eta} \gamma_{i,n}^{\xi+\zeta_i} \\ & \beta \left\{ \frac{\gamma_1 (\mathrm{e}^{\rho \tilde{x}_{x,k} + \tilde{x}_{\varepsilon,l} + \tilde{x}_{\eta,m}})^{\gamma_2} - \frac{\chi_0}{1+\chi} \left[\frac{(1-\alpha)}{\chi_0} \mathrm{e}^{\rho \tilde{x}_{x,k} + \tilde{x}_{\varepsilon,l} + \tilde{x}_{\eta,m}} \right]^{\gamma_2}}{\gamma_1 (\mathrm{e}^{\tilde{x}_{x,k} + \tilde{x}_{\eta,j}})^{\gamma_2} - \frac{\chi_0}{1+\chi} \left[\frac{(1-\alpha)}{\chi_0} \mathrm{e}^{\tilde{x}_{x,k} + \tilde{x}_{\eta,j}} \right]^{\gamma_2}} \right\}^{-\sigma} \times \\ & p^{(n-1)}(\rho \tilde{x}_{x_i^n,i} + \tilde{x}_{\varepsilon,l} + \tilde{x}_{\xi+\zeta_i,n}, \tilde{x}_{\eta,m}) \end{aligned}$$

where $\gamma_{i,j}^z \equiv [\Gamma_z]_{i,j}$ are the transition probabilities found in Step 1 and we use linear interpolation to evaluate $p_i^{(n-1)}(\rho \tilde{x}_{x_i^n,i} + \tilde{x}_{\varepsilon,l} + \tilde{x}_{\xi+\zeta_i,n}, \tilde{x}_{\eta,m})$ because its arguments will not typically lie on one of the nodes. The risk-neutral valuation can be calculated in an analogous way.

4. The unconditional real term premium is approximated by the difference in sample averages between the value of a bond and its risk-neutral counterpart in Monte Carlo simulations. The sample average valuation of the bond with full information can be calculated similarly, although there is no need for moment-matching in Step 2 or linear interpolation in Step 3 since with full information the bond is priced by rational expectations rather than forecasts. The sample average of the risk-neutral counterparts follow analogously.

To achieve precision in computations we set $\tilde{N} = 15$, which is significantly larger than the values typically used in the literature. The extremes of the vector of nodes in Step 1 are at ± 7 standard deviations for all the exogenous processes, with the exception of x_i^n for which a lower span of ±2 standard deviations is imposed to guarantee that the conditional forecast $\theta x_{i,t}^n$ remains in $[\tilde{x}_{x,1}, \tilde{x}_{x,\tilde{N}}]$ even for large θ. Figure 10 shows the Euclidean distance between discretised and theoretical moments in Step 2 for different values of ω. At all levels of strategic complementarity the mean squared distance on the vertical axis lies between 10^{-5} and 10^{-8}, implying a high degree of accuracy when approximating the $x_t|x_{i,t}^n$ distribution.

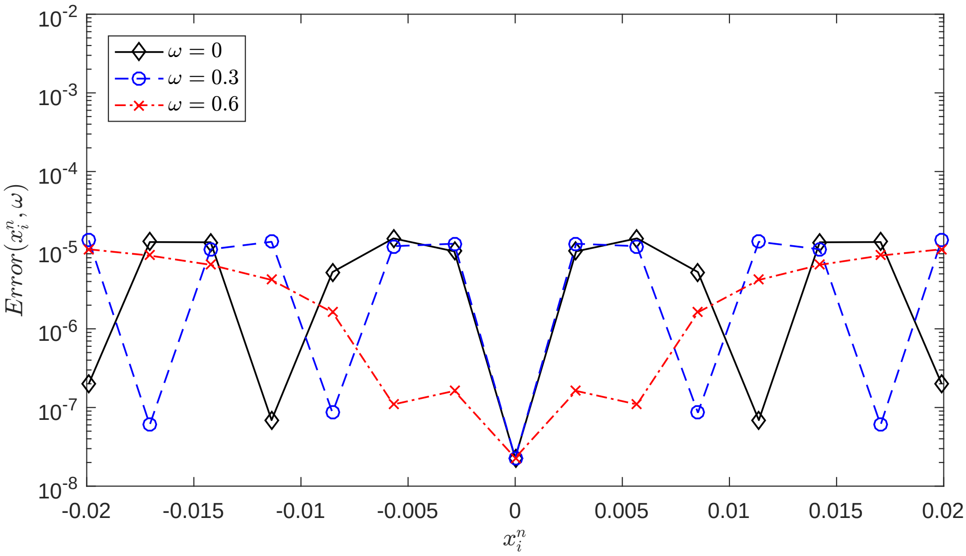

Figure 10: Approximation error in distribution of $x_t|x_{i,t}^n$