Primary Dealers, Subjective Bond Returns and the Term Structure of Interest Rates

Cem Tütüncü*

Stanford University

December 2023

Click here for the latest version.

Abstract

This paper shows how the subjective beliefs of large dealer banks help understand the excess volatility in bond markets, the large volatility of long-term interest rates. I document that the interest rate exposures of primary dealers comove systematically with the interest-rate forecasts of their research departments, both in the cross-section of dealers as well as over time. In particular, primary dealers choose higher interest-rate risk exposures when they are more optimistic about the returns on long Treasury bonds relative to short T-bills. I develop and estimate an equilibrium model with dealer banks that have heterogeneous interest-rate expectations. The quantitative model shows that the variation in dealers' beliefs about future interest rates is a strong mechanism to explain the volatility of long rates.

*E-mail address: ctutuncu@stanford.edu. I am deeply grateful to Monika Piazzesi, Martin Schneider, and Hanno Lustig for their invaluable guidance. For helpful suggestions and discussions, I thank Adrien Auclert, Yunus Can Aybas, Luigi Bocola, Jean-Felix Brouillette, Sebastian Di Tella, Refet Gürkaynak, Matteo Leombroni, Marcelo Sena, Chris Tonetti, and seminar participants at Stanford. I acknowledge financial support from the Stanford Institute for Economic Policy Research. Any errors are my own.

1 Introduction

Historically, long-term interest rates in the U.S. have displayed significant volatility when contrasted with the long-term averages of short-term rates. Moreover, there is documented evidence of profitable market-timing strategies within bond markets, suggesting that it is possible to anticipate excess returns on long-term bonds. For example, an investment strategy involving borrowing at the short-term rate to purchase long-term bonds at the onset of a recession and selling them right after the recession yields substantial returns. Nevertheless, it is unclear why this predictability evidence is not exploited in real-time; if investors were to capitalize on these opportunities, bond prices would adjust, and predictability would vanish.

Primary dealers are involved in substantial trading within Treasury bond markets as marketmakers and actively research bond market dynamics. It is particularly important to understand why such sophisticated investors, who play a crucial role within the United States financial system, have not yet taken advantage of the predictability of bond returns. As of June 2022, the five largest dealer banks held 45.5% of all assets in the U.S. banking industry.[1] Given the pivotal position that dealers hold in bond markets, it is remarkable that they have not seized the opportunities presented by the aforementioned market-timing strategies.

In this paper, I show that the subjective bond return expectations of the largest bond dealers, and their portfolio decisions based on these expectations, can account for the predictability of bond excess returns and time-variation in bond risk premia. If dealers' real-time expectations about bond returns do not coincide with what econometricians observe in hindsight, portfolio decisions based on these expectations need not exploit the return predictability documented with in-sample predictive regressions. I empirically document that large primary dealers' interest rate risk exposure and their research departments' bond excess return forecasts comove systematically, both over time and in the cross-section of dealers. Dealers who expect higher returns on long-term Treasury bonds bear more interest rate risk by increasing their holdings of long-term assets.

I develop a quantitative-term structure model with dealer banks who have heterogeneous beliefs about the dynamics of interest-rates. In the model, dealers choose their interest rate risk exposures according to their subjective beliefs, and bond prices reflect dealer beliefs as well as their risk exposures. I quantify my model by incorporating data on dealers' exposures and return forecasts, thus disciplining the subjective beliefs in the

[1] Source: https://www.chicagofed.org/banking/financial-institution-reports/top-banks-bhcs.

model to generate realistic portfolio positions. The quantified model suggests that time-variation in the subjective beliefs of dealers is a strong mechanism to explain the excess volatility of long-term rates and the predictability of excess bond returns. Hence, my model connects the triad of large dealers' interest rate risk exposures, their subjective bond return forecasts, and time-variation in expected excess returns.

For my main empirical specification, I construct a measure of interest rate risk exposure at the Bank Holding Company (BHC) level for the largest primary dealers in the U.S., using FR Y-9C filings administered by the Federal Reserve and the CRSP/-Compustat database.[2] I then compute the subjective expected excess bond returns of dealers at a quarterly frequency using their interest rate forecasts reported in Blue Chip Financial Forecasts. Running panel regressions of dealers' interest rate risk exposures on their expected excess return forecasts reveals that risk exposures comove with expected returns on long-term bonds. When dealers' subjective expected excess returns on long-term bonds are one standard deviation higher, their interest rate risk exposure is 2.5 percentage points higher on average. Dealers are heterogeneous in their expectations and exposures both over time and in the cross-section, with dealers who expect higher excess returns being more exposed to interest rate risk. Additionally, I demonstrate that the risk exposure of dealers significantly predicts realized bond excess returns, with the predictive power of exposures increasing when dealers' expected excess returns are higher. This finding suggests that large dealer banks' Euler equations hold under their subjective beliefs.

To assess the contribution of dealers' subjective beliefs in explaining time-variation in bond risk premia, I develop a quantitative term-structure model and use dealers' portfolio positions and survey forecasts to discipline the dynamics of subjective beliefs. The model is designed to parsimoniously capture the impact of dealers' subjective beliefs about interest rates, and their associated risk exposures, while ruling out alternative mechanisms that could generate time-varying bond risk premia. I find that dealer beliefs account for substantial time-variation in bond risk premia, and can explain the predictability of Treasury excess returns. My results demonstrate the significance of large financial intermediaries' subjective expectations for explaining bond price dynamics, complementing the evidence on the importance of their risk assessment.

In the model, I consider a single risk factor, representing interest-rate risk. The short-term interest rate moves around a stochastic mean, both varying with interest-rate risk,

[2] The dealers in my sample are Bank of America, Bank One, Citigroup, Goldman & Sachs, and J.P. Morgan & Chase.

capturing the impact of short-rate shocks on the expected path of future short-rates. Log-utility dealers have heterogeneous beliefs about the stochastic mean, characterized as a mean-reverting bias from its true value. I postulate that dealer beliefs also vary over time along with the single risk factor. Hence, dealers have time-varying subjective beliefs about the expected path of future short-rates, and accordingly about future bond prices. The subjective belief processes are exogenous, and based on my empirical findings, I use data on dealer exposures and forecasts to estimate their dynamics. The model does not impose a particular foundation for the formation of these beliefs (such as Bayesian learning or behavioral biases).

How do the subjective beliefs of each particular dealer impact bond prices? In the model, zero-coupon bond prices are wealth-share weighted averages of bond prices in hypothetical economies where there is only a single type of dealer. Beliefs about future short-rates both affect the hypothetical single-dealer bond prices, and also the distribution of wealth amongst dealers. Thus, each dealer's impact on prices is directly related to their wealth-share. I solve the single-dealer bond prices in closed-form, which allows me to evaluate the effects of the short-rate, the stochastic-mean, and the subjective beliefs on interest rates in a straightforward fashion. Dealers' relative wealth shares over time are entirely determined by their belief differences.

The model features two key mechanisms that generate time-varying risk premia. Subjective beliefs about future short-rates, and consequently about future bond excess returns, are the first drivers of risk premia dynamics. Dealers agree-to-disagree about the short-rate dynamics, specifically about its stochastic mean, and choose their interest rate risk exposures accordingly. Dealers who are *pessimistic* about long-term bond returns require a higher compensation for bearing interest rate risk, which is reflected in bond prices, and thus in the statistical risk premia observed by the econometrician. As dealers' subjective beliefs about future returns change over time, the compensation they require for bearing risk also varies over time. An econometrician who observes the equilibrium bond returns from this model concludes that they are predictable. The model successfully generates the bond return predictability evidence documented in data using actual bond returns.[3] In the estimated model, the time-variation in dealers' pessimism is the main force of bond risk premia. Dealers on average expect lower bond returns relative to the statistical expectation, and thus command higher risk premia to hold long-term bonds. Dealers' reluctance to take long positions in bonds (due to low subjective expected re-

[3] Specifically, I run Fama and Bliss (1987) regressions using model implied and actual bond prices. The model also produces the hump-shaped factor in Cochrane and Piazzesi (2005).

turns) can thus speak to why investors have not capitalized on the predictability patterns historically.

The second source of risk premia in the model is due to redistribution of wealth. Dealers engage in speculative trading based on their beliefs about future short-rates, and thus face the risk of making losses on their positions, which redistributes wealth from dealers with highly incorrect beliefs to more accurate dealers. I find that redistribution risk hardly matters for risk premia dynamics, because the wealth distribution among dealers only changes slowly over time. It therefore cannot explain the cyclical variation in bond risk premia.

Nevertheless, dealers' wealth distribution still matters for bond risk premia, because it determines the weights on dealers' subjective beliefs, and the beliefs of dealers with a larger share of aggregate wealth are reflected more in bond prices. The estimation shows that the dealers with the most pessimistic beliefs also have the largest wealth share. As a consequence, their pessimism receives a large weight in determining risk premia in equilibrium. During recessions, this pessimism is especially pronounced and raises the risk premium. Moreover, the model is able to produce the dealer wealth distribution patterns observed in data, such as J.P. Morgan & Chase taking over Bank of America to become the largest bank over time. Hence, my model suggests that subjective beliefs about asset returns, and the resulting portfolio positions, could provide a meaningful explanation for the evolution of the dealer wealth distribution, amongst many other possible explanations.[4]

Finally, the estimation of the model uses quarterly data on dealer exposures, survey forecasts, and actual interest rates during the years 2001-2021. Dealers' forecast errors for the 3-month interest rate are informative about the persistence of their subjective beliefs and the volatility of the short-rate. Time variation in dealers' interest-rate exposures is informative about the volatility of their subjective beliefs. The initial distribution of dealer wealth is measured from the distribution of dealers' assets in Federal Reserve filings at the beginning of the data sample. The remaining parameters are estimated with observations on long-term interest rates.

[4] My model does not speak to mergers and acquisitions, hence is unable to capture factors such as Bank One and J.P. Morgan's merger in the early part of the sample, or the redistribution of Lehman Brothers and Bear Stearns assets following 2008.

1.1 Related Literature

I contribute to several strands of the macro-finance literature. Many studies document potential drivers of bond return predictability, emphasizing time variation in bond market risks or the risk assessment of investors. Wachter (2006) and Bansal and Shaliastovich (2013) focus on the implications of variation in risk assessment for the term-structure of interest rates.[5] They argue that when expected excess returns are projected to be high, investors do not act on this predictability due to increasing risks associated with holding long-term bonds or a greater aversion to interest rate risk among investors. This paper proposes an alternative explanation based on subjective beliefs: when econometricians project high bond returns, bond investors are pessimistic about bond returns, and therefore do not want to buy long bonds.

There is a growing body of work that consider financial intermediaries as marginal agents for asset prices, rather than households, and my paper also follows this approach.[6] Vayanos and Vila (2021) and Schneider (2023) are examples of papers that model intermediaries as marginal agents in bond markets. Vayanos and Vila (2021) consider a model of preferred-habitat, where intermediaries act as arbitrageurs facing the demand from habitat-investors for various bond maturities. Schneider (2023) considers an economy with a financial sector along the lines of Brunnermeier and Sannikov (2014), and shows that occasionally-binding endogenous leverage constraints of intermediaries generates time-varying term premium. Numerous studies focus specifically on primary dealers due to their significant role in bond markets. Haddad and Sraer (2020) show that the average interest risk sensitivity of U.S. banks predicts future bond returns, and relate it to long-term bond holdings of banks. They develop an equilibrium model of the term-structure, where the interest risk sensitivity of banks is an exogenous state-variable that generates time-varying bond risk premia. Kekre et al. (2022) provide further evidence on primary dealers' portfolio holdings impacting bond yields, and show that it is connected to monetary policy. Du et al. (2022) document that primary dealers' positions in Treasury bonds have shifted from a net short to a net long following the Great Recession, and bond yields have reflected this change. They further relate these findings to interest swap spreads and balance sheet costs. I contribute to this literature by also incorporating the

[5]These papers build on the frameworks in Campbell and Cochrane (1999) and Bansal and Yaron (2004) respectively.

[6]See for example: He and Krishnamurthy (2013), Dell'Ariccia and Marquez (2013), Brunnermeier and Sannikov (2014), He et al. (2017), Adrian et al. (2014), Adrian et al. (2016), Drechsler et al. (2018), Greenwood et al. (2018), Haddad and Muir (2021).

heterogeneity in dealer positions and expectations, and demonstrating that it influences bond risk premia.[7]

My paper also contributes to the literature exploring the heterogeneity of intermediaries in various contexts. Gabaix and Maggiori (2015), Maggiori (2017) and Morelli et al. (2022) are examples of papers that study heterogenous intermediaries' role in international financial markets, focusing on exchange rate determination, international risk-sharing, and systemic debt crises respectively. Importantly, they connect heterogeneity in risk exposures of intermediaries with financial markets. Kargar (2021) and Coimbra and Rey (2021) explore the implications of heterogeneous intermediaries for risk premia. Kargar (2021) considers the implications of heterogeneity in financial intermediaries' risk aversion for asset prices. Coimbra and Rey (2021) model intermediaries who face different Value-at-Risk constraints, and show that heterogeneity in risk-taking matters for asset price and risk premia dynamics. My contribution is to introduce another dimension of heterogeneity amongst intermediaries by investigating the subjective beliefs of financial intermediaries.

Several papers use survey data on expectations to connect investor beliefs and asset prices. Piazzesi et al. (2015) use median survey forecasts to estimate investors' subjective bond risk premia using an affine-term structure model. They find that subjective risk premia is less volatile and not very cyclical compared to the statistical premia. Nagel and Xu (2023) explore the dynamics of subjective risk premia in many asset classes including Treasury bonds, also using survey forecast data. They also find that subjective premia is not very cyclical, and document a similar lack of cyclicality in out-of-sample statistical forecasts of excess returns. I contribute to the literature studying subjective bond risk premia by highlighting the role of primary dealer beliefs, and their associated risk exposures in a structural equilibrium model. Furthermore, I consider the potential impact of heterogeneity in dealers' forecasts instead of using the consensus or median forecasts.

Many papers present theoretical frameworks to study the implications of subjective beliefs for asset prices.[8] Fan (2006) and Xiong and Yan (2010) are studies that specifically focus on the Treasury bond interest rates and disagreement amongst forecasters. Similar

[7] Other examples of papers that explore dealer and bank balance sheets in various contexts include: Begenau et al. (2015), Boyarchenko et al. (2021) He et al. (2022), Allen and Wittwer (2023).

[8] See for example: Detemple and Murthy (1994), Zapatero (1998), Basak (2000), Chiarella and He (2002), Scheinkman and Xiong (2003) Basak (2005), Anderson et al. (2005), Buraschi and Jiltsov (2006), Jouini and Napp (2006), Jouini and Napp (2007), David (2008), Dumas et al. (2009), Chen et al. (2010), Cvitanić and Malamud (2011), Chen et al. (2012), Cvitanić et al. (2012), Bhamra and Uppal (2014), Chien et al. (2016), Andrei et al. (2019), Borovička (2020).

to my setup, in their settings time-variation in risk premia is entirely due to subjective beliefs. In Fan (2006), agents have subjective beliefs about their income shocks, and also about other agents' beliefs about their respective income shocks. In Xiong and Yan (2010), agents disagree about inflation dynamics, as they receive signals about shocks to the long-run mean of inflation, and have different beliefs about the informativeness of signals regarding economic fundamentals. Bond prices in their setup are also a wealth-weighted average of bond prices in hypothetical single-agent economies. However, their work considers only two types of investors rather than multiple financial intermediaries. Giacoletti et al. (2021) incorporate both the consensus forecast and disagreement amongst forecasters in an affine-term structure model, and consider bond risk premia from the perspective of a Bayesian econometrician. Yet, none of these studies incorporate data on portfolio holdings while disciplining the belief processes.

Finally, Leombroni et al. (2020), Giglio et al. (2021) Kekre and Lenel (2022) are recent studies that connect subjective investor expectations with portfolio holdings. My paper complements their findings, as I show that primary dealers' subjective beliefs and portfolio holdings can serve as a strong mechanism in explaining the excess volatility of long-term interest rates and time-variation in bond risk premia.

This paper is organized as follows. Section 2 describes the data and presents the main empirical results about exposures and subjective expected excess returns. Section 3 describes the model. Section 4 presents the results from the model estimation. Section 5 concludes.

2 Exposure & Subjective Bond Returns

2.1 Data

Subjective Expectations

The subjective expectations of dealers are from the Blue Chip Financial Forecasts. The sample starts in 2001:Q1 and ends in 2021:Q4. The data consists of surveys conducted at a monthly frequency, where survey participants report their forecasts for the averages of various macro-financial variables over a quarterly horizon up to five quarters ahead. Primary dealers who participate in the surveys over my sample are Bank of America, Bank One, Citigroup, Goldman & Sachs, J.P. Morgan & Chase, and Wells Fargo.

The forecasts I use to construct subjective bond excess returns are for the average Federal Funds Rate, the 3-month, 6-month, 1-year, 2-year, 5-year, 10-year, and 30-year Treasury bond interest rates, over the quarter during which the survey was conducted. To obtain quarterly observations on dealers' expectations, I use their forecasts at the beginning month of each quarter.

In some surveys, multiple subsidiaries of the same dealer participate. I take the average of their forecasts, and exclude those who are not directly subsidiaries of the dealer.

Expected Excess Returns

Let $y_t^{(\tau)}$ denote the (log) interest rate (or yield) on a zero-coupon bond with maturity of τ quarters at time t. The bond price is: $P_t^{(\tau)} = \exp\big(-\tau y_t^{(\tau)}\big)$. I use this relation to convert interest rate forecasts into implied bond price expectations. The notation $r_t = y_t^{(1)}$ denotes the one-quarter short-term interest rate, and $p_t^{(\tau)}$ denotes the log bond price. The return (in logs) on a bond with maturity τ, in excess of the short-rate, over a holding period from t to $t+1$ is defined as:

$$rx_{t+1}^{(\tau)} = p_{t+1}^{(\tau-1)} - p_t^{(\tau)} - r_t. \tag{1}$$

Let $\overline{y}_{t,t+1}^{(\tau)}$ denote the average interest rate on a τ period bond between t and $t+1$. This average rate corresponds to the forecasted interest rate variables in Blue Chip surveys. Dealer i's average expected returns from buying a bond of maturity τ between t and $t+1$ and selling it between $t+1$ and $t+2$, in excess of the average short-rate between t and $t+1$, is denoted with $E_t^i[\overline{rx}_{t+1,t+2}^{(\tau)}]$. Using the relation between log prices and log interest rates, these expectations relate to survey forecasts at time t as:

$$E_t^i[\overline{rx}_{t+1,t+2}^{(\tau)}] = -(\tau-1)\underbrace{E_t^i[\overline{y}_{t+1,t+2}^{(\tau-1)}]}_{\substack{\textit{forecast of } \tau-1 \\ \textit{quarter interest} \\ \textit{rate}}} + \tau \underbrace{E_t^i[\overline{y}_{t,t+1}^{(\tau)}]}_{\substack{\textit{forecast of } \tau \\ \textit{quarter interest} \\ \textit{rate}}} - \underbrace{E_t^i[\overline{r}_{t,t+1}]}_{\substack{\textit{forecast of} \\ \textit{3-month} \\ \textit{interest rate}}}. \tag{2}$$

The τ quarter and the 3-month interest rate forecasts are directly observed. However, $y_t^{(\tau-1)}$ is not reported for all the maturities available in the surveys. For the 1-year bond, I use the next-quarter ahead forecasts of the 6-month rate to recover the implied 9-month interest rate.

For longer maturities, I use term-structure equalities to approximate the implied forecasts. The quality of approximation relies on two key assumptions. First, I assume that each forecaster's beliefs satisfy the law of iterated expectations. Now, let $rpy_t^{i,(\tau)}$ denote the subjective term premium of dealer i at time t. No-arbitrage implies that the interest rate forecast of a τ period bond can be decomposed into the average expected future short-rates, and a term premium:

$$E_t^i\left[y_t^{(\tau)}\right] = \frac{1}{\tau}E_t^i\left[r_t + r_{t+1} + \ldots + r_{t+\tau-1}\right] + rpy_t^{i,(\tau)}. \tag{3}$$

The second assumption is that the difference between the subjective term premia on $\tau-1$ and τ period bonds is small relative to the subjective term premium on the $\tau-1$ period bond. This difference arises only due to the subjective premium commanded for holding a zero-coupon bond from $t+\tau-1$ to $t+\tau$. For a 2-year bond, the difference corresponds to the additional premium the dealer requires to hold the bond for 8 quarters as opposed to 7, and to the additional premium to hold the bond for 40 quarters as opposed to 39 quarters for a 10-year bond. As long as this additional component constitutes a small fraction of the term-premium, the approximation works well. In Appendix A.1, I formally derive why the two assumptions are needed, and discuss the accuracy of the approximation in detail. Appendix A.2 further shows that replacing the forecast of the 3-month interest rate in (2) with the actually observed 3-month interest rate does not alter the results of the paper.

Interest Rate Data

The data on interest rates is from Liu and Wu (2021), who use a non-parametric kernel-smoothing method to estimate rates on zero-coupon bonds from observed Treasury bond prices. The data is monthly with maturities ranging from 1 month to 30 years. I average their monthly data over a quarter to obtain a quarterly series.

Exposures

To construct my exposure measure, I closely follow Gomez et al. (2021). I use the quarterly Consolidated Financial Statements for Bank Holding Companies (form FR Y-9C) administered by the Federal Reserve Board, and the quarterly CRSP/Compustat database. Both datasets are obtained from WRDS. I focus on the Bank Holding Companies (BHC) of the primary dealers, for which granular data is available over the sample period.

The measure of exposure is:

$$Exposure^i = 1 - \frac{Assets^i_{\leqslant 1} - Liabilities^i_{\leqslant 1}}{Total\ Assets^i}, \tag{4}$$

where $Assets^i_{\leqslant 1}$ and $Liabilities^i_{\leqslant 1}$ indicate the assets and liabilities of dealer i that are repricing or maturing within one year respectively. The measure corresponds to $1 - Income\ Gap^i$ as defined by Mishkin and Eakins (2009) and Gomez et al. (2021). My measure reflects the portion of net long-term, fixed-rate asset holdings, and thus the interest-rate sensitivity of dealer portfolios. It is available at a quarterly frequency and for the entire cross-section of dealers in my sample, and it corresponds to how interest-rate sensitivity is measured in practice by market-participants. Appendix A.2 details the use of CRSP/Compustat database as an alternative to construct $Assets^i_{\leqslant 1}$ and $Liabilities^i_{\leqslant 1}$, in a way analogous with Gomez et al. (2021)'s construction, and reports the relevant results.

Core Deposits & Derivatives

There are two key issues in regard to using income gaps when constructing dealer exposure. The first concerns the treatment of "core" deposits, corresponding to transaction or savings deposits. Since I use BHC-level data, these deposits contribute to a large percentage of short-term liabilities through the dealers' commercial banking subsidiaries. However, despite their short-term contractual maturity, the interest rates on these deposits are known to adjust sluggishly to market rates, and thus these deposits are like longer term assets. In constructing the income gaps, and consequently the exposures, I treat these deposits as liabilities that reprice or mature within more than one year as in Gomez et al. (2021).

The second issue with measuring interest-rate risk exposure is the treatment of derivatives. Dealers can in principle trade in derivatives to manage their exposure to interest-rate risk, which would imply that the measure of exposures would overestimate the actual risk exposures. Yet, Begenau et al. (2015) find that banks who are large participants in the derivatives market do not hedge their exposure due to other business, and instead have derivative positions that gain in value when interest rates fall, just like long-term bonds. This finding suggests that dealers are even more exposed to interest rate risk. Since the exposure measure (4) does not contain derivatives, it understates the actual exposure by dealers.

Despite its limitations, the measure of exposures reveals an important degree of variation both over time and in the cross-section that is consistent with the findings of the literature on banks' risk-taking. Figure 1 depicts the time-series of the cross-sectional average of dealer exposures. There is a slight increase in dealer exposures leading up to the Great Financial Crisis (GFC) in 2008, and since then exposures have been declin-

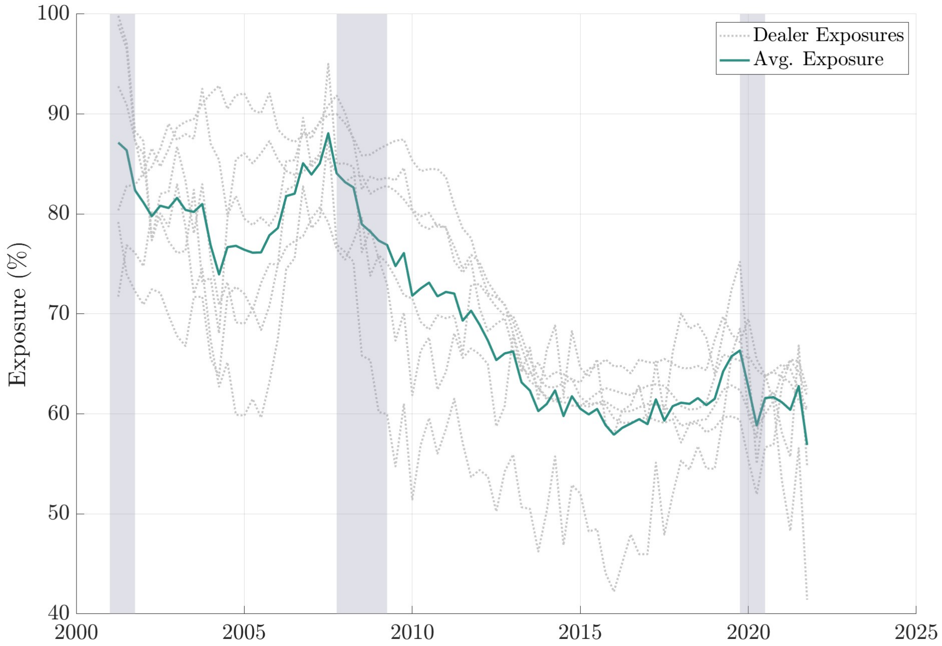

Figure 1: Avg. Dealer Exposure

Notes: The teal line plots the average exposure across dealers. The grey dotted lines plot the individual dealer exposures. Shaded areas correspond to NBER recessions.

ing. With the infusion of short-term reserves into the banking sector due to quantitative easing (QE), dealer exposures steadily decline in the post-GFC period. Similarly, during the Covid-19 recession, exposures declined due to Fed's QE policies.

The average exposure is not very volatile, with a standard deviation of 9.4% compared to its sample average of 70.5%. However, the individual dealer exposures show stark differences over the sample, and are much more volatile relative to the average. Figure 2 displays the dealer exposures after removing a time-trend to account for the effects of QE, and after removing dealer fixed-effects.[9] The two largest dealers, J.P. Morgan & Chase and Bank of America are highlighted for comparison. Dealer exposures are much more volatile relative to the average exposure, and display stark differences at times. For example, earlier in the sample Bank of America is more exposed to interest rate risk than J.P. Morgan, whereas this pattern switches directions following the GFC. During the COVID-19 crisis, J.P. Morgan responds by lowering its exposure and then increasing

[9]Dealers may have different business models or investment mandates that lead to consistent differences in exposures, which are not due to business cycle variables.

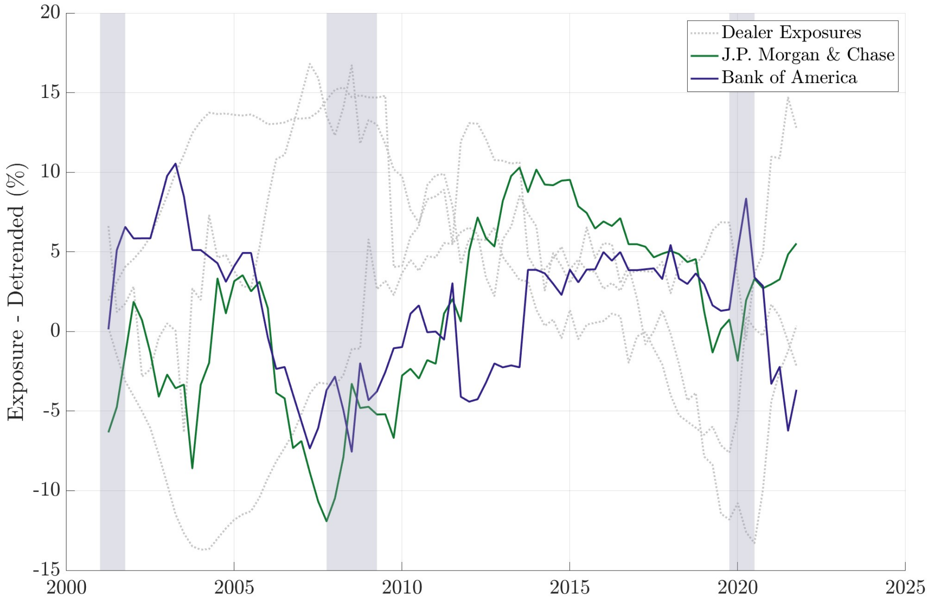

Figure 2: Comparison of Detrended Exposure

Notes: The dark green line plots the exposure of J.P. Morgan & Chase, and the dark blue line plots the exposure of Bank of America. The grey dotted lines plot the individual dealer exposures. Shaded areas correspond to NBER recessions.

it back up again, whereas Bank of America rapidly increases its exposure, and then proceeds to take on a much safer position following the crisis. These patterns suggest that dealers choose different strategies to make profits. The following section shows that these strategies are tightly related to dealers' beliefs about future bond returns.

2.2 Main Results

Baseline Regression

For my benchmark empirical analysis, I estimate the following regression using quarterly data:

$$Exposure_{t,i} = \alpha + fe_i + \kappa t + \beta^{(\tau)} E_t^i\left[\overline{rx}_{t+1,t+2}^{(\tau)}\right] + \varepsilon_{t,i}, \tag{5}$$

where i represents each individual dealer and fe_i are dealer-fixed effects. I include a time-trend κt to control for the downward trend in exposures due to QE. The dependent variable $Exposure_{t,i}$ is the interest-rate risk exposure measure in (4). I use the last available observation at the beginning of the quarter for the corresponding survey forecasts. That is, I regress the last available "snapshot" of a dealer's exposure on their subjective expected excess returns on long-term bonds over the following quarter.

Table 1: Exposures & Subjective Expected Excess Returns

	Exposure					
	3-Month Treasury			Fed Funds Rate		
	(1)	(2)	(3)	(1)	(2)	(3)
1 YR	−1.210	−2.538	−2.538*	−2.489	−2.489	−2.489*
	(3.602)	(1.91)	(1.556)	(2.719)	(3.501)	(1.774)
2 YR	1.289	2.216**	2.216**	0.766	2.539**	2.539**
	(1.600)	(1.139)	(1.157)	(1.204)	(1.454)	(1.163)
5 YR	0.610	0.796**	0.796***	0.447	0.780*	0.780*
	(0.573)	(0.359)	(0.338)	(0.385)	(0.512)	(0.534)
10 YR	0.502*	0.372**	0.372***	0.290*	0.548**	0.548**
	(0.310)	(0.166)	(0.150)	(0.183)	(0.323)	(0.345)
30 YR	0.255**	0.135**	0.135***	0.168**	0.091*	0.091*
	(0.127)	(0.055)	(0.037)	(0.083)	(0.063)	(0.057)
Avg. R^2	0.74	0.83	0.83	0.67	0.81	0.81
Observations	323	323	323	370	370	370
Dealer FE	*YES*	*YES*	*YES*	*YES*	*YES*	*YES*
Time Trend		*YES*	*YES*		*YES*	*YES*
Cluster Robust SE			*YES*			*YES*

Notes: This table displays the estimation results from the regression:

$$Exposure_{t,i} = \alpha + fe_i + \kappa t + \beta^{(\tau)} E_t^i\big[\overline{rx}^{(\tau)}_{t+1,t+2}\big] + \varepsilon_{t,i}$$

Standard errors are in parentheses. Stars indicate significance at ***: 99%, **: 95%, *: 90% confidence levels.

I consider two main specifications; one where I use the 3-month interest rate forecasts as the short-rate while constructing expected excess returns, and another where I instead use the Federal Funds Rate forecasts as the short-rate.[10] Table 1 reports the estimated coefficients. The standard errors for the baseline cases are corrected for heteroskedasticity using the Huber-White method, and I also consider Liang-Zeger cluster-robust standard errors to account for potential serial correlations.[11] The coefficient estimates in column

[10]Dealers have access to overnight lending markets of the Fed, so the relevant short-term borrowing rate is the Federal Funds Rate. In this case, the returns in (1) correspond to bond returns in excess of the average overnight borrowing rate over a quarter.

[11]See Abadie et al. (2023).

(2), which correspond to the estimates from (5), are all statistically significant except for the 1-year returns. The coefficients on two and longer maturity bonds are all positive, consistent with the notion that higher expected excess returns on long-term bonds are associated with higher exposure to these long-term assets, captured by my exposure measure. In contrast, the coefficients on the 1-year bond are negative, suggesting that dealers adjust their portfolios to hold short-term, low-risk assets when they expect high returns on short-term securities. The adjusted R^2s are surprisingly high, suggesting that after adjusting for dealer fixed effects and the time-trend, cyclical variation in subjective expected excess returns accounts for almost all of the remaining variation in exposures.[12] The results for the specification using the Federal Funds Rate are similar, displaying a similar pattern across maturities, albeit losing some statistical significance.

Finally, to account for potential serial correlation over time in dealer specific residuals, I compute cluster-robust standard errors treating each dealer as a single cluster. Columns labeled (3) on 1 report the resulting standard errors and significance levels. Cluster-robust standard errors for the 3-month Treasury short-rate specification are lower than the Huber-White errors for maturities 1, 5, 10, and 30 years, hinting towards some degree of negative serial correlation in dealer-specific residuals relating to these maturities. For the 2-year returns, Huber-White errors are slightly lower, suggesting a mild positive serial correlation. The directions of the effect for the Federal Funds Rate specification are mixed, with no change in significance levels except for the estimate for the 1-year returns.

The coefficients in Table 1 represent the response of exposures to a 1 percentage point (pp.) increase in the subjective expected excess returns for each maturity, and decline in magnitude as maturity increases. However, subjective expected excess returns on longer bonds are much more volatile, thus it is natural that a 1 pp. increase in the expected returns on a 30-year bond barely has a significant implication for risk exposures. To account for this disparity, I standardize the estimates and standard errors by scaling the subjective expected returns by their sample standard deviations. Figure 3 demonstrates the resulting coefficients, which now correspond to a one standard deviation increase in the subjective expected excess returns. Following the volatility adjustment, the coefficients for each maturity line up around 2.5 pp. This finding provides further support for the use of a single "interest rate risk exposure" measure constructed from the long-term holdings of dealers: exposure moves in the same way in response to a one standard deviation

[12]Note that I do not imply any causality, as variation in subjective expected excess returns might be reflecting variation in other macro-financial variables that matter for dealers' exposure decisions.

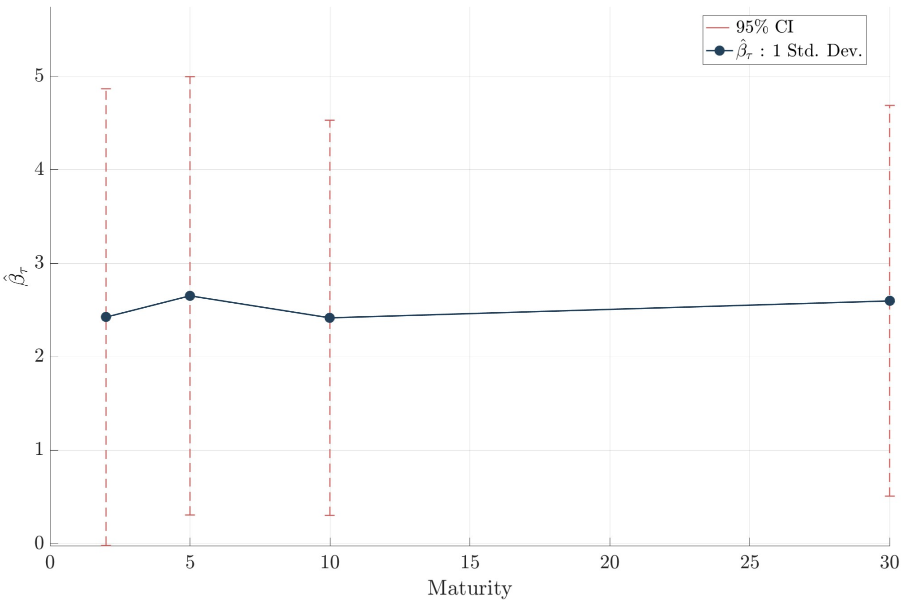

Figure 3: Standardized $\beta^{(\tau)}$ Estimates

Notes: The dark blue dots represent the regression coefficients from Table 1, scaled by the sample standard deviation of the subjective expected excess returns for the corresponding maturity. The vertical dashed red lines display the scaled 95% confidence intervals.

increase in long-term bond returns regardless of maturity. If bonds load on a single risk factor and the loadings are increasing with maturity, then these coefficients would correspond to the response of exposures to an expected upward movement in the risk factor.

Predictive Regressions

Do large dealers' interest rate risk exposures matter for realized excess returns? Haddad and Sraer (2020) document that average income gaps of banks significantly predict one-year excess returns on Treasury bonds. I first repeat their exercise using only the position data of the dealers in my sample, and I document that the resulting estimates are very similar to those of Haddad and Sraer (2020). I run the regressions

$$rx^{(\tau)}_{t,t+4} = a^{(\tau)} + b^{(\tau)} \times \overline{Exposure}_t + \epsilon^{(\tau)}_{t+4}, \tag{6}$$

where $rx^{(\tau)}_{t,t+4}$ is the one-year excess return on a bond with maturity τ quarters. $\overline{Exposure}_t$ corresponds to the cross-sectional average of dealers' exposures at the beginning of each period. Table 2 reports the estimates. I compute Newey and West (1987) standard errors with 4-lags correction, and adjust the estimates for potential bias as inStambaugh (1999) using the methodology of Lewellen (2004). Average exposures of dealers significantly predict one-year excess returns, with explanatory power declining in maturity. Except for the 30-year bonds, all coefficient estimates are statistically significant at 90% confidence level. The magnitude of coefficients increases in maturity in absolute terms, but the explanatory power declines, as the adjusted R^2 decreases from 0.41 for the 1-year bond to only 0.05 for the 10-year bond. Overall, the magnitudes of estimates are highly similar to those of Haddad and Sraer (2020), suggesting that large dealers' exposures are an important driver of the predictability evidence they document.[13]

Table 2: Exposure & Realized Bond Excess Returns

	(1) $rx^{(4)}$	(2) $rx^{(8)}$	(3) $rx^{(20)}$	(4) $rx^{(40)}$	(5) $rx^{(120)}$
$\overline{Exposure}_t$	0.23***	0.32***	0.44**	0.50*	0.15
	(0.07)	(0.12)	(0.23)	(0.33)	(0.75)
Constant	−0.12***	−0.17**	−0.19	−0.17	0.14
	(0.05)	(0.08)	(0.16)	(0.25)	(0.57)
Observations	83	83	83	83	83
Adjusted R^2	0.41	0.30	0.12	0.05	−0

Notes: This table displays the estimation results from the regression:

$$rx^{(\tau)}_{t,t+4} = a^{(\tau)} + b^{(\tau)} \times \overline{Exposure}_t + \epsilon^{(\tau)}_{t+4}$$

Newey-West standard errors with 4 lags are in parentheses. I correct for potential bias as in Stambaugh (1999) using the methodology of Lewellen (2004). Stars indicate significance at ***: 99%, **: 95%, *: 90% confidence levels.

Now that I establish that dealer exposures predict bond excess returns, the next question to ask is whether their subjective expectations are reflected in bond returns. To analyze the effect of dealers' subjective expected excess returns, I extend the regression in (6) to incorporate dealer forecasts:

$$rx^{(\tau)}_{t,t+4} = a^{(\tau)} + b^{(\tau)} \overline{Exposure}_t + c^{(\tau)} \overline{E^i_t\big[rx^{(\tau)}_{t,t+1}\big]} + \gamma^{(\tau)} \overline{Exposure}_t \times \overline{E^i_t\big[rx^{(\tau)}_{t,t+1}\big]} + \epsilon^{(\tau)}_{t+4}, \quad (7)$$

where $\overline{E^i_t\big[rx^{(\tau)}_{t,t+1}\big]}_t$ represents the cross-sectional average of the quarterly subjective expected excess returns of dealers at date t, for the maturities available in the surveys.

[13] Haddad and Sraer (2020) report results for maturities longer than 5-years in their Internet Appendix.

The main coefficient of interest is $\gamma^{(\tau)}$, the effect of the interaction term. I am interested in seeing whether the predictive power of dealer exposures depends on their subjective expected excess returns.

Table 3: Exposure, Subjective Beliefs & Realized Bond Excess Returns

	(1) $rx^{(8)}$	(2) $rx^{(20)}$	(3) $rx^{(40)}$	(4) $rx^{(120)}$
$\overline{Exposure}_t$	0.48***	0.95***	0.86**	0.49
	(0.10)	(0.22)	(0.39)	(0.89)
$\overline{E_t^i[rx_{t,t+1}^{(k)}]}_t$	−0.77**	−0.52***	−0.15*	−0.10
	(0.39)	(0.17)	(0.10)	(0.12)
$\overline{Exposure}_t \times \overline{E_t^i[rx_{t,t+1}^{(k)}]}$	1.29***	0.82***	0.25	0.11
	(0.51)	(0.22)	(0.21)	(0.14)
Constant	−0.27***	−0.52***	−0.39*	−0.14
	(0.07)	(0.16)	(0.30)	(0.71)
Observations	83	83	83	83
Adjusted R^2	0.49	0.30	0.08	-0

Notes: This table displays the estimation results from the regression:

$$rx_{t,t+4}^{(\tau)} = a^{(\tau)} + b^{(\tau)}\ \overline{Exposure}_t + c^{(\tau)}\ \overline{E_t^i[rx_{t,t+1}^{(\tau)}]} + \gamma^{(\tau)}\ \overline{Exposure}_t \ \times\ \overline{E_t^i[rx_{t,t+1}^{(\tau)}]} + \epsilon_{t+4}^{(\tau)}$$

Newey-West standard errors with 4 lags are in parentheses. I correct for potential bias as in Stambaugh (1999) using the methodology of Lewellen (2004). Stars indicate significance at ***: 99%, **: 95%, *: 90% confidence levels.

Table 3 reports the estimation results for this regression. The main effects of exposures are statistically significant except for the 30-year excess returns, and quarterly average expected excess returns of dealers positively predict realized one-year excess returns, as the coefficients in the second row of Table 3 are significant except for the 30-year returns. What provides even stronger evidence on the role of subjective beliefs are the coefficient estimates on the interaction terms. For 2 and 5-year excess returns, average dealer exposure more strongly predicts realized excess returns when the average expected excess returns of dealers are higher. This suggests that the Euler equation for large dealer banks holds under their subjective beliefs.

If in equilibrium dealers expect high excess returns on long-term bonds, and thus increase their holdings of long-term assets, this would lead them to be more exposed to interest rate risk. The market compensation for risk has to rise in order for them to take such positions. When the subjective expected excess returns, representing the compensation dealers expect for bearing interest rate risk, are higher, the market compensation for risk has to adjust more strongly. This mechanism is the main driving force for bond risk premia in my quantitative model.

3 Model

This section represents an equilibrium model that captures the main empirical findings presented in the previous section. The main focus of the model is interest rate risk, subjective expectations about interest rates, and risk exposures. The model describes heterogeneous dealers who trade in bonds with various maturities. The dealers have heterogeneous expectations about future interest rates, and choose their risk exposures accordingly.

3.1 Setup

I consider a continuous time, infinite horizon pure-exchange economy. Throughout I operate on a probability space $(\Omega, \mathcal{F}, \mathcal{P})$ augmented with the filtration $\mathbb{F} = \{\mathcal{F}_t : 0 \leqslant t\}$, satisfying the usual regularity conditions in Protter (2005).

The economy is populated by a continuum of infinitely lived dealers. There are N different types of dealers, indexed by $i : 1, \ldots N$, who differ in their beliefs about the interest rate r_t on a short-term bond. Each dealer discounts the future with a rate of time preference ρ, has log utility, and receives an endowment stream y_t^i.

Dealers trade with each other using contingent claims that are in zero net supply. There are enough contingent claims so that markets are dynamically complete. Dealers maximize their expected utility, which they evaluate using their subjective beliefs. All dealers of type i are identical with each other, thus in what follows I will directly proceed with dealer types instead of referring to individual dealers.

Short Rate

There is a single shock process in this economy: interest-rate risk. To introduce this shock, I first specify the equilibrium dynamics of the short-rate and subjective beliefs of different dealer types. Then I back out the exogenous aggregate endowment process that is consistent with the equilibrium short-rate and dealer beliefs.

The short-rate follows a mean-reverting Ornstein-Uhlenbeck process:

$$dr_t = \kappa_r(\mu_t - r_t) + \sigma_r dB_t, \tag{8}$$

where B_t is a standard Brownian Motion. Based on Balduzzi et al. (1998), I assume that the mean μ_t to which the short-rate reverts to is stochastic, and also follows an Ornstein-Uhlenbeck process:

$$d\mu_t = \kappa_\mu(\bar{\mu} - \mu_t) + \sigma_\mu dB_t. \tag{9}$$

The short-rate and its stochastic mean are driven by the same Brownian Motion B_t, which represents *interest-rate risk*. The volatilities σ_r and σ_μ can have positive or negative signs. For example, if $\sigma_r > 0$ and $\sigma_\mu < 0$, a positive short-rate shock today implies lower expected rates in the future.

Subjective Beliefs

At each time t, dealers have subjective expectations of the stochastic mean, described as a bias from the true value by ϵ^i_t:

$$\mu^i_{r,t} = \mu_{r,t} + \epsilon^i_t, \tag{10}$$

where the law of motion of ϵ^i_t is:

$$d\epsilon^i_t = -\kappa_\epsilon \epsilon^i_t dt + \sigma^i_\epsilon dB_t. \tag{11}$$

To keep the model parsimonious, I assume that the rate of mean-reversion κ_ϵ is the same across all dealers. However, dealers differ in the volatility of their beliefs, σ^i_ϵ. That is, each dealer's bias responds differently to interest rate risk, which is the only source of heterogeneity in dealers' beliefs. Yet, although σ^i_ϵ are time-invariant, because the belief processes ϵ^i_t vary over time, dealers will have heterogeneous beliefs at any point in time. I do not take a stance on why this bias arises. Specifically, although it is possible to generate such bias in a Bayesian-learning setting, I allow for the possibility that agents are non-Bayesian.[14]

Aggregate Endowment

The aggregate endowment follows a Geometric Brownian Motion with a time-varying drift:

$$\frac{dY_t}{Y_t} = (r_t + \mu_{Y,t})\, dt + \sigma_Y dB_t. \tag{12}$$

[14] In a similar setup, Dumas et al. (2009) show how the law of motion described in Eq. (11) can be the outcome of an imperfect information setting with Bayesian agents learning about correlated signals.

The component of the drift in excess of the short-rate, $\mu_{Y,t}$, is determined in equilibrium such that the short-rate process defined in (8) and the subjective beliefs defined in (10) are indeed the equilibrium processes. The volatility of the aggregate endowment σ_Y, on the other hand, is constant. This feature of my model is important to assess the role of subjective beliefs of dealers, as it implies that the *quantity of risk* is constant. Dealers have logarithmic utility with the coefficient of relative risk aversion set to 1, and thus *risk assessment* is also constant. Taken together, these assumptions would generate constant risk premium under rational expectations as in the standard Lucas-asset pricing model. The only potential source of time-variation in risk premium is via the subjective beliefs of dealers.

To summarize the model setup, dealers observe the short-rate, but agree-to-disagree about its dynamics. If σ_μ and σ_ϵ^i are of the same sign, then dealer i overreacts to interest rate shocks. If instead they are of the opposite sign, dealer i underreacts. Now suppose $\sigma_r > 0$, $\sigma_\mu < 0$. In this case, a dealer who underreacts expects higher future rates relative to the statistical dynamics. Because bond prices are inversely proportional to interest rates, this is a dealer who is *"pessimistic"* about bond returns. On the contrary, a dealer who overreacts expects very low rates and thus is *"optimistic"* about bond returns. Finally, degrees of dealer optimism and pessimism vary over time due to the stochastic volatility term in (51).

3.2 Equilibrium

Before I move on to defining the equilibrium in this economy, I first define the restriction on the aggregate endowment such that the resulting equilibria are consistent with the short-rate process and subjective beliefs.

Lemma 1. *Assume that the aggregate endowment process is as given in* (12)*. Then, if* $\mu_{Y,t}$ *satisfies:*

$$\mu_{Y,t} = -\rho + \sigma_Y^2 - \sum_{i=1}^{N} \frac{c_t^i}{Y_t} \left(\sigma_Y \epsilon_t^i \frac{\kappa_r}{\sigma_r} \right), \tag{13}$$

and the subjective beliefs are characterized by (10) *and* (11)*, then the short-rate process* (8) *is the equilibrium short-rate process.*

Proof. See B.6. □

Note that under rational expectations, $\epsilon_t^i = 0, \forall i \in I$ and this condition reduces down to $\mu_{Y,t} = -\rho + \sigma_Y^2$ for all $t > 0$. Then the drift of the aggregate endowment becomes

$r_t - \rho + \sigma_Y^2$, which is exactly the condition on the equilibrium short-rate satisfied in the standard Lucas-asset pricing model. Rearranging this expression yields further insights:

$$\sum_{i=1}^{N} \frac{c_t^i}{Y_t} \underbrace{\left(\mu_{Y,t} + \sigma_Y \epsilon_t^i \frac{\kappa_r}{\sigma_r} \right)}_{\text{subjective } \mu_{Y,t}} = -\rho + \sigma_Y^2. \tag{14}$$

The term in the parentheses is the subjective drift of the endowment in excess of the short-rate under dealer i's beliefs. The condition then states that the wealth-share weighted average of the subjective "excess" return on the aggregate endowment is constant.

Equipped with this condition, I can now define the equilibrium:

Equilibrium: Given a distribution of initial endowments $\{y_0^i\}_{i=1}^N$, an equilibrium in this economy is a collection of consumption processes $\{c_t^i\}_{i=1}^N$ and a set of contingent claims prices such that:

1. dealers optimize,
2. markets clear ; $\sum_{i=1}^{I} c_t^i = Y_t$, contingent claims holdings sum up to zero
3. $\mu_{Y,t}$ satisfies (14).

3.3 Interest Rate Risk Exposures

In this section, in order to connect the model to data on exposures and dealer forecasts, I express dealers' optimization problem in terms of bond returns and interest-rate risk exposures.

There is a single risk-factor, B_t. Then, standard arguments imply that the risk-free asset paying r_t and one risky asset are sufficient to span and complete asset markets.[15] Dealers choose their "exposure" to the risk-factor, therefore any dealer portfolio position can be perfectly replicated with a simple portfolio consisting of these two assets. Following Haddad and Sraer (2020), I introduce a "coupon bond" as the risky asset with a stream of dividend payments $\theta e^{-\theta\tau}$, at each date $\tau \geqslant t$. This coupon bond represents a consolidated portfolio of all long-term assets held by dealers, such as deposits, mortgage-backed securities, corporate bonds, loans, etc. The parameter θ controls the maturity profile of these assets, as the average maturity is θ^{-1}. I use x_t^i to denote type i dealers'

[15] See for example: Duffie (2010).

exposure to the risky asset, corresponding to my exposure measure in the data. Finally, I also introduce zero-coupon bonds of all maturities in zero-net supply.

With log-utility, standard arguments imply that optimal consumption is a constant fraction of wealth: $c_t^i = \rho w_t^i$ where w_t^i denotes the wealth of dealer type i. Moreover, since with complete markets, any portfolio position is redundant, I can express dealers' optimization problem in terms of their zero-coupon bond portfolios and their wealth:

$$V_t\left(w_t^i\right) = \max_{\left\{\left\{\alpha_{t+s}^{(\tau),i}\right\}_{\tau\in(0,\infty)}\right\}} \mathbb{E}_t^i\left[\int_t^\infty e^{-\rho s}\log\left(\rho w_s^i\right)ds\right]$$
$$s.t. \tag{15}$$
$$dw_t^i = w_t^i\left(r_t-\rho\right)dt + \int_0^\infty \alpha_t^{(\tau),i} w_t^i\left(\frac{dP_t^{(\tau)}}{P_t^{(\tau)}} - r_t dt\right)d\tau,$$

where $\alpha_t^{(\tau),i}$ denotes the portfolio share of bond with maturity τ as a fraction of wealth, and $P_t^{(\tau)}$ denotes the price of the bond with maturity τ. $\frac{dP_t^{(\tau)}}{P_t^{(\tau)}} - r_t$ is thus the instantaneous excess return on a bond with maturity τ.

Now, the portfolio of zero-coupon bonds can be perfectly replicated using the risky coupon bond. Thus the portfolio shares of bonds and "exposure" to the risky asset are related in the following way:

$$Exposure\ x_t^i = \frac{\alpha_t^{(\tau),i}}{\theta e^{-\theta\tau}} = \frac{Portfolio\ share\ of\ bond\ with\ maturity\ \tau}{dividend\ of\ long\ term\ asset}. \tag{16}$$

Then I can rewrite the dealers' problem in terms of exposures and bond returns as follows:

$$V_t\left(w_t^i\right) = \max_{\left\{\left\{x_t^i\right\}_{t=0}^\infty\right\}} \mathbb{E}_t^i\left[\int_t^\infty e^{-\rho s}\log\left(\rho w_s^i\right)ds\right]$$
$$s.t. \tag{17}$$
$$dw_t^i = w_t^i\left(r_t-\rho\right)dt + x_t^i\int_0^\infty \theta e^{-\theta\tau} w_t^i\left(\frac{dP_t^{(\tau)}}{P_t^{(\tau)}} - r_t dt\right)d\tau,$$

The optimality condition for dealers derived in Appendix B.2 is:

$$\underbrace{\mathbb{E}_t^i\left(\frac{dP_t^{(\tau)}}{P_t^{(\tau)}}\right) - r_t}_{subjective\ expected\ excess\ returns} = \underbrace{x_t^i\int_0^\infty \theta e^{-\theta u}\,\mathrm{Cov}_t^i\left(\frac{dP_t^{(\tau)}}{P_t^{(\tau)}}, \frac{dP_t^{(u)}}{P_t^{(u)}}\right)du}_{covariance\ of\ returns\ with\ portfolio}. \tag{18}$$

This condition is equivalent to the standard Euler equation for dealers, under dealer i's beliefs. The left-hand side of the equality is the dealer i's expected excess return, or required risk compensation for holding a bond with maturity τ. The right-hand side of the equality tells us that the compensation required depends on the covariance of the bond excess returns with the already accumulated exposure to interest-rate risk. Furthermore, Girsanov Theorem implies that $Cov_t^i\left(\frac{dP_t^{(\tau)}}{P_t^{(\tau)}}, \frac{dP_t^{(u)}}{P_t^{(u)}}\right) = Cov_t\left(\frac{dP_t^{(\tau)}}{P_t^{(\tau)}}, \frac{dP_t^{(u)}}{P_t^{(u)}}\right)$. Thus in the model, differences in exposures across dealers depend only on the subjective expected excess returns, not on differences in risk assessment.

3.4 Bond Prices

I use dealers' optimality condition (18) to derive zero-coupon bond prices. First, I show that a familiar condition in models with subjective beliefs and log-utility is also satisfied in my model.

Proposition 1. *Consider a hypothetical economy consisting of only type i dealers. Let* $P_t^{(\tau),i}$ *denote the price of a zero-coupon bond with maturity* τ *in this hypothetical single-dealer economy. Then, the bond price in the heterogeneous dealer economy is given by the wealth-share weighted average of single-dealer economy prices:*

$$P_t^{(\tau)} = \sum_{i=1}^{N} \frac{C_t^i}{Y_t} P_t^{(\tau),i}. \tag{19}$$

Proof. See Appendix B.4. □

With log-utility, each dealer's pricing impact on the actual bond price depends on their relative wealth (consumption). The richer a dealer is, the more influence it has on the market prices.

To obtain the single-dealer economy bond prices, I guess and verify an exponential-affine form, as is standard in the term-structure literature. I derive closed-form solutions for these prices, along with other equilibrium objects such as risk premia.

Proposition 2. *Bond prices in the hypothetical single-dealer economy are given by:*

$$P_t^{(\tau),i} = \exp\left(A(x_t^i, \tau) + C_r(\tau) r_t + C_\mu(\tau)\mu_t + C_\epsilon^i(\tau)\epsilon_t^i\right), \tag{20}$$

where

$$C_r(\tau) = -\frac{(1 - e^{-\kappa_r \tau})}{\kappa_r}, \tag{21}$$

$$C_\mu(\tau) = -\frac{1}{\kappa_\mu}\left(1 - \frac{\kappa_r e^{-\kappa_\mu \tau} - \kappa_\mu e^{-\kappa_r \tau}}{\kappa_r - \kappa_\mu}\right), \tag{22}$$

$$C_\epsilon^i(\tau) = \frac{\kappa_\mu\sigma_r + \kappa_r\sigma_\mu}{\kappa_\mu\left(\kappa_r\sigma_\epsilon^i - \kappa_\epsilon\sigma_r\right)}\left(1 - e^{-\left(\kappa_\epsilon - \kappa_r\frac{\sigma_\epsilon^i}{\sigma_r}\right)\tau}\right) - \frac{\kappa_\mu\sigma_r + \kappa_r\sigma_\mu}{\kappa_\mu\left(\kappa_r\sigma_\epsilon^i - \kappa_\epsilon\sigma_r + \kappa_\mu\sigma_r\right)}\left(e^{-\kappa_r\tau} - e^{-\left(\kappa_\epsilon - \kappa_r\frac{\sigma_\epsilon^i}{\sigma_r}\right)\tau}\right), \tag{23}$$

and

$$\begin{aligned} A(x_t^i,\tau) = \int \Big(& C_\mu(\tau)\kappa_{\mu_r}\bar{\mu} + \frac{1}{2}C_r(\tau)^2\sigma_r^2 + \frac{1}{2}C_\mu(\tau)^2\sigma_\mu^2 + \frac{1}{2}C_\epsilon^i(\tau)^2\sigma_\epsilon^{i2} + C_r(\tau)b_\mu(\tau)\sigma_r\sigma_{\mu_r} \\ & + C_r(\tau)C_\epsilon^i(\tau)\sigma_r\sigma_\epsilon^i - x_t^i\Big(C_r(\tau)\sigma_r + C_\mu(\tau)\sigma_\mu + C_\epsilon^i(\tau)\sigma_\epsilon^i\Big)Z_{\theta,r}^i\Big)d\tau, \end{aligned} \tag{24}$$

where

$$Z_{\theta,r}^i = \frac{\sigma_r\left(\kappa_\epsilon + \theta\right)\left(\kappa_r\sigma_\mu + \kappa_\mu\sigma_r + \sigma_r\theta\right)}{\left(\kappa_r + \theta\right)\left(\kappa_\mu + \theta\right)\left(\kappa_r\sigma_\epsilon^i - \kappa_\epsilon\sigma_r - \sigma_r\theta\right)}. \tag{25}$$

Proof. See Appendix B.4. □

The coefficients on the state-variables depend on the relative persistences and volatilities of the short-rate, its stochastic mean, and dealer beliefs. Note that the coefficient $A(x_t^i,\tau)$ depends on the exposure of dealer i. This is because the prices in the single-dealer representative agent economy should be consistent with the risk exposures under the heterogeneous dealer economy. x_t^i is not pinned down in the single-dealer economy, and hence must adjust until the hypothetical economy prices are such that the risk-factor exposure is identical to the one in the heterogeneous dealer economy. This has to be satisfied for each dealer, such that the optimal exposures are a fixed-point satisfying the bond price equation (19).

The interest rates on the bonds are:

$$y_t^{(\tau),i} = -\frac{1}{\tau}\Big(A(x_t^i,\tau) + C_r(\tau)r_t + C_\mu(\tau)\mu_t + C_\epsilon^i(\tau)\epsilon_t^i\Big). \tag{26}$$

Figure 4 demonstrates the interest rate loading coefficients, $-C_r(\tau)/\tau, -C_\mu(\tau)/\tau$ and $-C_\epsilon(\tau)/\tau$. The loading on r_t declines exponentially, indicating that the short-rate alone can potentially match the level of the yield curve but not the slope. The loadings on μ_t and ϵ_t^i on the other hand have a hump-shape. This hump-shape is important for matching the term-spreads of long and short-term interest rates, and the volatilities of long-term interest rates.

How do exposures affect expected excess returns? Figure 5 demonstrates how the

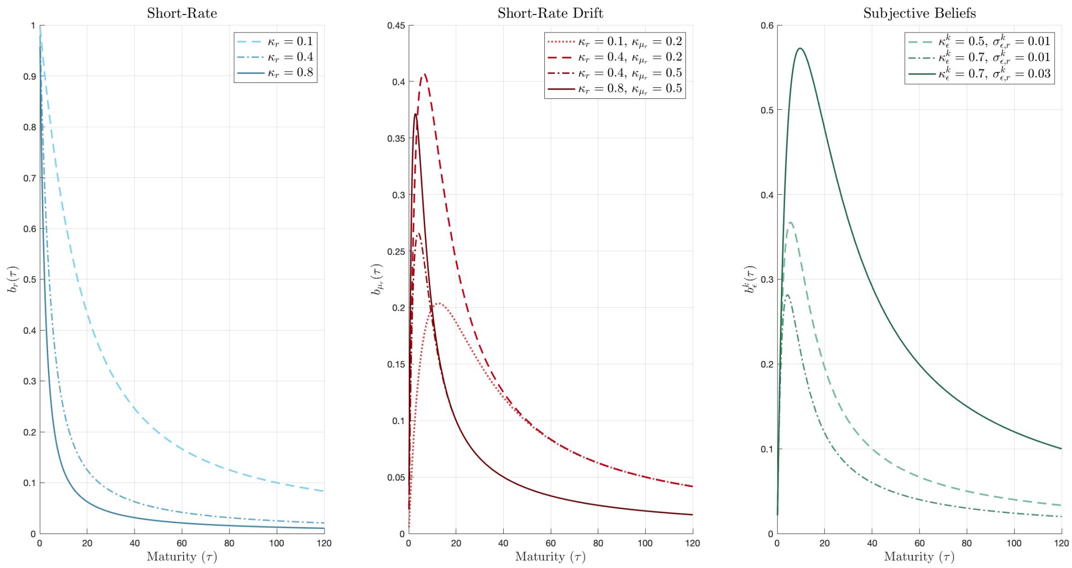

Figure 4: Interest Rate Loadings

Notes: *Left panel*: $-C_r(\tau)/\tau$ as a function of maturity for different values of κ_r. *Middle panel*: $-C_\mu(\tau)/\tau$ as a function of maturity for different values of κ_r and κ_μ. *Right panel*: $-C_\epsilon(\tau)/\tau$ as a function of maturity for different values of $\sigma_\epsilon^i,\ \kappa_\epsilon$.

expected excess bond returns in the single-dealer economy depend on exposures, from the perspective of the econometrician. Higher exposure increases the expected excess returns on bonds of all maturities, and the effect is stronger when short-rate volatility is higher. Intuitively, higher exposure increases the bond risk premium that the dealers command to hold long-term bonds. In an equilibrium with only type i dealers, bond returns adjust accordingly so that the portfolio positions of dealers are unchanged. This mechanism relating exposures and bond risk premia is identical to the one in Haddad and Sraer (2020) or Vayanos and Vila (2021). Yet, with heterogeneous dealers, there are now three forces that impact the (statistical) bond excess returns: *(i)* exposures of each dealer type, *(ii)* subjective beliefs of dealers, and *(iii)* the relative wealth-share of each dealer type. The risk premium is the required compensation for exposure to interest rate risk, the risk of being "wrong" about future interest rate dynamics, and the risk of wealth being redistributed amongst dealers given their positions.

Finally, to compute bond prices in the heterogeneous dealer economy I need the consumption shares. The following proposition shows that consumption shares can be computed directly using the exogenous laws of motion for the state variables.

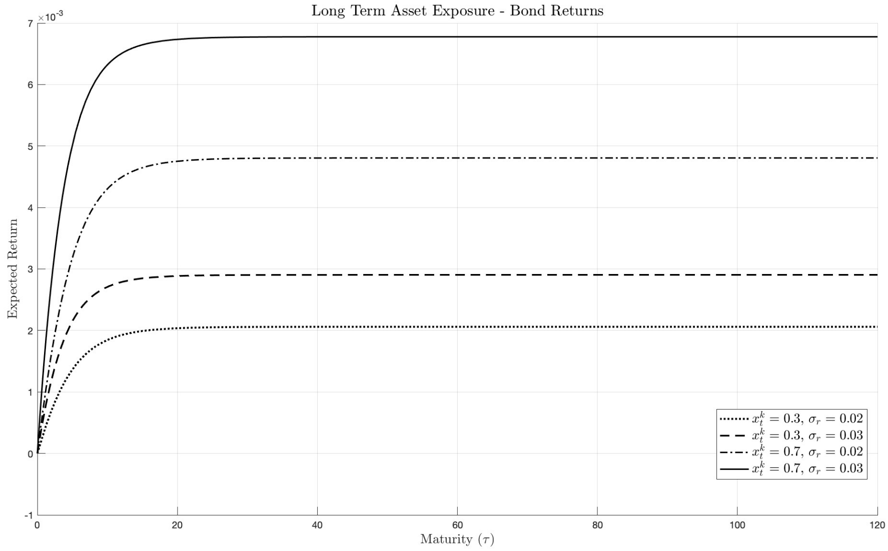

Figure 5: Exposures & Excess Returns

Notes: The plotted lines correspond to the drift of bond returns in the hypothetical single-dealer economy under the econometrician's measure. Each line corresponds to the expected excess returns for different levels of exposure to interest rate risk, and for different volatilities of the short-rate.

Proposition 3. *Define the relative likelihood ratio of dealer j's beliefs with respect to dealer i's beliefs as $\xi_t^{j,i}$. These likelihood ratios evolve as:*

$$\frac{d\xi_t^{j,i}}{\xi_t^{j,i}} = (\epsilon_t^j - \epsilon_t^i)\frac{\kappa_r}{\sigma_r} dB_t - (\epsilon_t^j - \epsilon_t^i)\epsilon_t^i \left(\frac{\kappa_r}{\sigma_r}\right)^2 dt \tag{27}$$

Moreover, relative consumption ratios also evolve as:

$$\frac{dc_t^j/c_t^i}{c_t^j/c_t^i} = \frac{d\xi_t^{j,i}}{\xi_t^{j,i}}, \tag{28}$$

starting from initial values λ^i/λ^j, where λ^i and λ^j depend only on the initial endowments $\{y_0^i\}_{i=1}^I$.

Proof. See Appendix B.6. □

Proposition 3 states that the relative consumption ratios of dealers can be characterized for the entire history given the evolution of exogenous subjective belief processes and the initial endowment distribution. I solve for c_t^1/Y_t at each period using the aggregate resource constraint, and use the relative likelihood ratios to recover the remaining consumption shares. Details of the solution method are provided in Appendix B.6.

Equation (27) provides useful insights about the evolution of wealth ratios. The first term $(\epsilon_t^j - \epsilon_t^i)\frac{\kappa_r}{\sigma_r}dB_t$ reflects the changes in wealth that are due to movements in interest rates. If $\epsilon_t^i < \epsilon_t^j$, that is dealer j expects higher interest rates compared to dealer i, then a positive shock to interest rates favors dealer j and wealth flows from i to j. The second term $-(\epsilon_t^j - \epsilon_t^i)\epsilon_t^i \left(\frac{\kappa_r}{\sigma_r}\right)^2 dt$ reflects the overall accuracy of beliefs. On average, the dealers with smaller σ_ϵ^i are more accurate about their forecasts of μ_t. The sign of this expression depends on the sign of ϵ_t^i and the relative magnitudes of ϵ_t^i and ϵ_t^j. Hence in the long run, the most accurate dealer (smallest $|\epsilon_t^i|$) accumulates all wealth by taking positions based on superior forecasts. These two forces characterize the evolution of wealth in the short and long-run. Shocks to interest rates can generate wealth flows towards less accurate dealers temporarily, but eventually the most accurate dealer is the sole survivor.[16]

4 Quantitative Analysis

4.1 Estimation

There are three key goals of my estimation strategy: *(i)* to match actual Treasury bond interest rate dynamics, *(ii)* to match the dynamics of dealers' forecasts, and *(iii)* to match the relation between exposures and forecasts. In order to do so, I proceed step-by-step.

Subjective Forecast Errors

Proposition 4. *Let Δt denote the time-subinterval in the discretized version of the model. Let $\Delta B_t = B_{t+\Delta t} - B_t$ denote the discrete-time version of the Brownian Motion intervals, that is, independent and identically distributed standard Normal variables. Then, the subjective forecast errors of the short-rate follow an ARMA(1,1) in discrete time:*

$$r_{t+\Delta t} - E_t^i\left[r_{t+\Delta t}\right] = e^{-\kappa_\epsilon \Delta t}\left(r_t - E_{t-\Delta t}^k\left[r_t\right]\right) - \left(\sigma_\epsilon^i\left(1 - e^{-\kappa_r \Delta t}\right) + e^{-\kappa_\epsilon \Delta t}\sigma_r\right)\sqrt{\Delta t}\ \Delta B_t + \sigma_r\sqrt{\Delta t}\ \Delta B_{t+\Delta t}. \tag{29}$$

Proof. See Appendix B.7. □

[16] Survival here refers to a positive wealth share as time goes to infinity, following Sandroni (2000), Kogan et al. (2006), Yan (2008), Borovička (2020) and others.

In my model, the time-subinterval Δt is one quarter. The autoregressive component of the ARMA(1,1) specification (29) identifies κ_ϵ, and the unconditional variance identifies σ_r. Moreover, since κ_ϵ is assumed to be the same across all dealers, dealers' forecast errors can be stacked together to estimate (29). I use the subjective forecast errors of the 3-month interest rate to estimate the parameters of this equation. The estimates are all statistically significant at 95% confidence level.

Exposures & Term Structure Estimation

Upon estimating κ_ϵ and σ_r via (29), there are $4 + K$ parameters left to estimate: $(\kappa_\mu, \sigma_\mu, \kappa_r, \bar{\mu})$ and each dealer's σ_ϵ^i. As derived formally in Appendix B.3, the model implies that dealers' exposures satisfy:

$$x_t^i \int_0^\infty \theta e^{-\theta\tau} \frac{\sum_{i=1}^{I} \left(\frac{c_t^i}{Y_t} P_t^{(\tau),i} \left(\zeta_{r,t}^{\tau,i} - \sum_{j \neq i} \frac{c_t^j}{Y_t} \left(\epsilon_t^j - \epsilon_t^i \right) \right) \right)}{\sum_{i=1}^{I} \frac{c_t^i}{Y_t} P_t^{(\tau),i}} d\tau = \sigma_{W,r} - \frac{\kappa_r}{\sigma_r} \sum_{i=1}^{I} \frac{c_t^i}{Y_t} \epsilon_t^i + \frac{\kappa_r}{\sigma_r} \epsilon_t^i. \quad (30)$$

The key observation in regards to (30) is that the differences in dealers' exposures are only due to the last term $\frac{\kappa_r}{\sigma_r}\epsilon_t^i$. The wealth-share weighted average of subjective beliefs and bond prices matter for the common "level" of exposures, yet the cross-sectional differences are entirely determined by individual biases. Then, the cross-sectional differences in the sample standard deviations of subjective beliefs perfectly mimic the cross-sectional differences in the sample standard deviations of exposures, scaled by a factor common to all dealers. This observation reduces the number of free parameters to estimate. Namely, I search over the parameter space for a single belief volatility σ_ϵ^1, and use the differences in the sample standard deviations of dealers' exposures to back out the remaining belief volatilities. This leaves 5 remaining parameters to estimate. My approach takes into account the connection between exposures and subjective beliefs, and restricts the estimated belief processes to be coherent with dealer positions in the data. These restrictions are highly important to discipline the model, as they prevent the subjective beliefs dynamics from generating counterfactual portfolios in order to match interest rate dynamics.

The methodology for estimating the parameter vector $(\kappa_\mu, \sigma_\mu, \kappa_r, \bar{\mu}, \sigma_\epsilon^1)$ is rather standard in the term-structure estimation literature. Given an initial guess for the parameters, I generate the subjective belief processes and accordingly the consumption shares. The exposures constructed in Section 2 are directly input to the model. Thus, the data on exposures both directly matches their model counterparts, and disciplines the subjective belief processes. Upon obtaining the consumption shares, beliefs, and exposures, it is straightforward to recover the component of interest rates that only depend on r_t and μ_t. I assume the 1-year and 5-year interest rates are perfectly observed without any measurement error, and subtract from them the model-implied components that depend on dealer-specific variables. The remaining components are affine in the common state vector, and can be inverted to recover r_t and μ_t, similar to an affine term-structure model estimation.[17]

I use the Generalized Method of Moments to search over the parameter space, targeting the following term-structure moments:

1. The time-series of the averages of 1-year, 2-year, 4-year, 5-year, 7-year and 10-year interest rates at each quarter.

2. Term-spreads of 2 and 5-year, and 2 and 10-year interest rates.

3. Sample standard deviations of 1-year, 2-year, 4-year, 5-year, 7-year and 10-year interest rates.

Finally, θ^{-1} is calibrated to match the average maturity of long-term assets to be 10 years following Haddad and Sraer (2020), the time-preference rate is set to 0.03, and σ_Y is calibrated to match the observed exposures. The initial distribution of endowments is calibrated to match the distribution of gross domestic assets amongst the dealer banks in my data, at the beginning of the sample in 2001:Q1.[18] I exclude Bank One from the model estimation, as they merge with J.P. Morgan early in the sample, and thus there are few observations for their exposures and forecasts.

Estimation Results

Table 4 reports the estimation results. The stochastic mean is highly persistent, almost like a random walk, as κ_μ is nearly zero. The short-rate is also highly persistent, yet less so than the stochastic mean. A noteworthy observation is that the volatility of the short-rate and the stochastic mean have opposite signs. A positive shock to the short-rate ($dB_t > 0$) leads to lower expected future rates since $\sigma_\mu < 0$, and therefore higher bond prices.

The value of κ_ϵ indicates that subjective biases are highly transitory; in discrete-time it corresponds to an autoregressive coefficient of 0.23 for quarterly data. There is substantial heterogeneity in the volatilities of beliefs, with estimates ranging from

[17] See for example: Piazzesi (2010).

[18] I compute the distribution using the reported domestic assets on the form https://www.federalreserve.gov/releases/lbr/20010331/lrg_bnk_lst.pdf.

0.005 to 0.047, which more than doubles the volatility of the stochastic mean in absolute terms. The estimates suggest that dealers are *"pessimistic"*, as the belief volatilities have opposing signs with the volatility of the stochastic mean. Two dealers, in particular, expect future short-rates to be higher following a positive short-rate shock today, as $\sigma_\epsilon^1 + \sigma_\mu$ and $\sigma_\epsilon^3 + \sigma_\mu$ are both positive. Since higher expected future rates signal lower bond prices, these two dealers expect low returns on long-term bonds following an increase in the short-rate, thereby requiring higher compensation for holding these securities. The overall pessimism of dealers about long bond returns explains why dealers do not perceive a market-timing strategy in real-time, as they expect lower excess returns compared to the econometrician.

Table 4: Estimated & Calibrated Parameters

Description	Value	Target
κ_μ : persistence of stochastic mean	0.0010	Interest rate moments
σ_μ : volatility of short-rate level	-0.0189	Interest rate moments
κ_r : persistence of short-rate	0.0302	Interest rate moments
σ_r : volatility of short-rate	0.0079	ARMA(1,1) estimates
$\bar{\mu}$: long-run mean of short-rate	0.0125	Interest rate moments
κ_ϵ : persistence of beliefs	3.0931	ARMA(1,1) estimates
σ_ϵ^i : volatilities of dealer beliefs	$0.030, 0.005, 0.047, 0.015, 0.016$	Differences in volatilities of dealer exposures
ρ : time-preference rate	0.03	Literature
σ_Y : volatility of aggregate endowment	-0.0034	Dealer exposures
θ : average maturity of long-term assets	0.1	10 years - (Haddad and Sraer (2020))

Notes: This table displays the estimated parameters of the model, and the calibrated parameters. The model is estimated using quarterly data, and the reported estimates are in quarterly terms.

4.2 Model Fit

In this section, I evaluate the goodness-of-fit of the model. Figure 6 displays the targeted data moments and the model counterparts. The model successfully matches the time-series of the average of 1, 2, 4, 5, 7, and 10-year interest rates. Model implied term-spreads of 5-2 year and 10-2 year interest rates replicate the dynamics of their data equivalents, yet the model implied spreads are slightly lower during certain parts of the sample. Finally, the model matches the 1-year and 10-year interest rate volatilities almost perfectly, yet slightly underestimates the volatilities of the remaining maturities. Still, the model-generated values are close to their data counterparts. Overall, the model fares well in matching the targeted moments.

4.3 Untargeted Moments

I now discuss how the model matches untargeted moments about excess return predictability. The model's ability to match the statistical properties of bond excess returns is an important indicator of the plausibility of the model-implied real-time bond risk premia.

Fama-Bliss Regressions

I first consider a classic test of excess return predictability based on Fama and Bliss (1987) (FB henceforth). Define the one year log forward rate as:

$$f_t^{(\tau)} = p_t^{(\tau-4)} - p_t^{(\tau)}. \tag{31}$$

FB document that the *"forward spread"* between the τ period forward rate and the one-year interest rate predicts the one-year excess returns on a τ period bond, by estimating the regression:

$$rx_{t,t+4}^{(\tau)} = a_{FB}^{(\tau)} + \beta_{FB}^{(\tau)}(f_t^{(\tau)} - y_t^{(4)}) + \upsilon_{t+4}^{(\tau)}. \tag{32}$$

If the Expectations Hypothesis holds in its weak form, then expected excess returns are constant, indicating $\beta_{FB}^{(\tau)} = 0$ and thus no predictability. In its strong form, it further implies $a_{FB}^{(\tau)} = 0$. Since the original work of FB, the null hypothesis of no-predictability is rejected by numerous studies.

I estimate 32 separately for model-generated and actual data, using 2, 3, 4, and 5-year bond excess returns and forward spreads. Figure 7 demonstrates the predicted values $\widehat{rx}_{t,t+4}^{(\tau)} = \widehat{a}_{FB}^{(\tau)} + \widehat{\beta}_{FB}^{(\tau)}(f_t^{(\tau)} - y_t^{(4)})$ from these regressions. These predicted values correspond to the in-sample statistical expected excess returns. Clearly both the model-implied and data series are time-varying, indicating return predictability. Expected excess returns are cyclical, rising swiftly following recessions and declining subsequently. My model overestimates the 2-year and 3-year expected excess returns, but for the 4 and 5-year bonds, the model successfully matches the data values. Moreover, despite overestimating the level of expected excess return on the 2 and 3-year bonds, the model successfully captures the cyclical behavior of these returns. Since these moments are not targeted during estimation, these results hint towards the model's ability to generate plausible time-variation in bond risk premium. Yet, one should keep in mind that the plotted

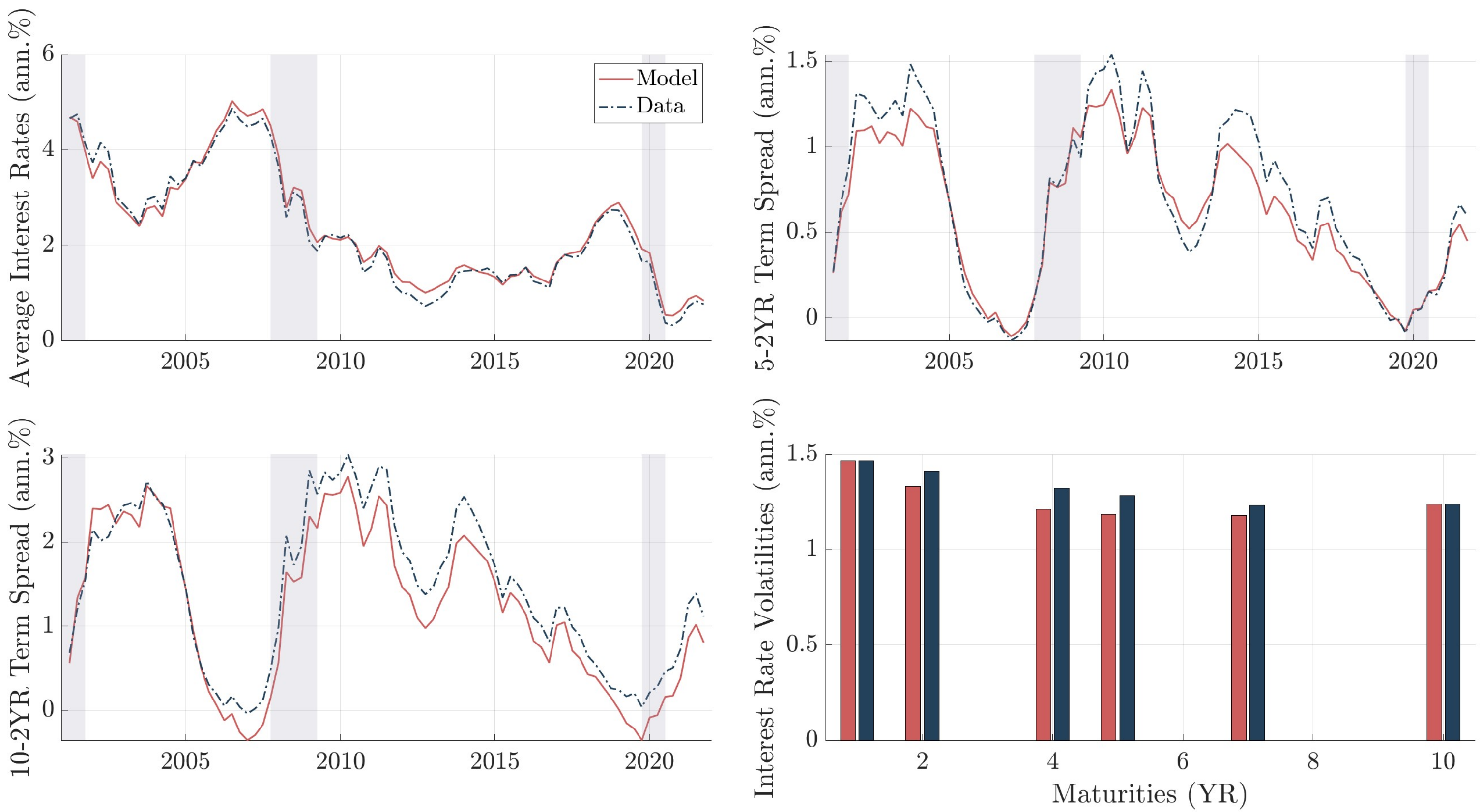

Figure 6: Model vs. Data

Notes: This figure displays the targeted moments in the data and the model counterparts. The red lines (bars) correspond to the model moments, and the blue lines (bars) correspond to the data moments. All values are expressed in annualized percentages. *Top Left:* this panel compares the time-series of the average of 1, 2, 4, 5, 7, and 10-year interest rates generated by the model with the actual data counterpart. *Top Right:* this panel compares the term-spread between the 5-year and 2-year interest rates generated by the model with the actual data counterpart. *Bottom Left:* this panel compares the term-spread between the 10-year and 2-year interest rates generated by the model with the actual data counterpart. *Bottom Right:* this panel compares sample standard deviations of the interest rates generated by the model with the actual data counterparts, for each maturity indicated on the horizontal axis.

series in Figure 7 are *statistical* risk premia, estimated using full-sample information. They do not necessarily reflect the real-time risk premium.

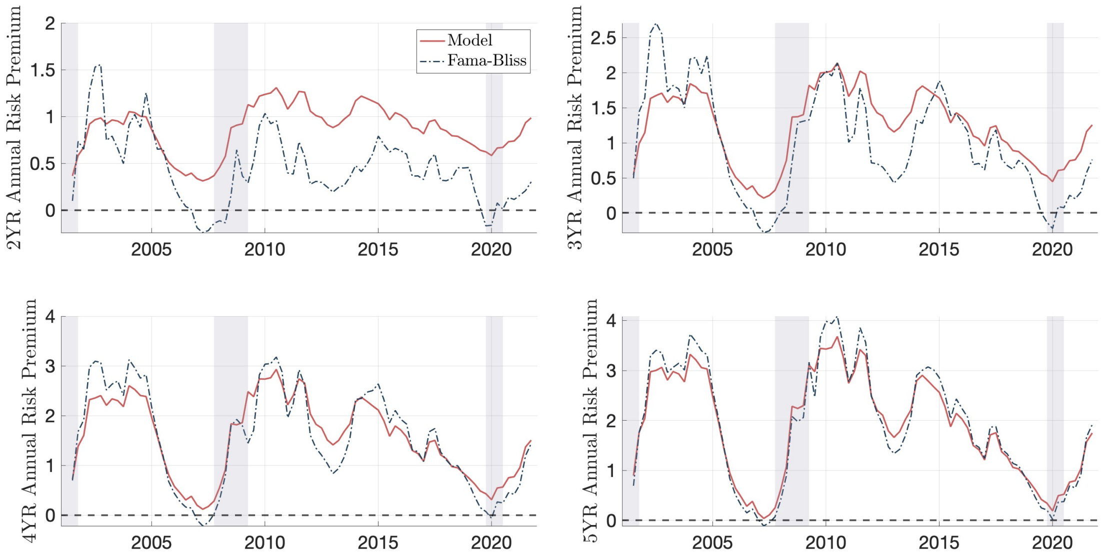

Figure 7: Fama-Bliss Predicted Values

Notes: The red lines correspond to the model-implied predicted values, and the blue lines correspond to the predicted values using actual data. All values are expressed in percentages.

Cochrane-Piazzesi Factor

Cochrane and Piazzesi (2005) (CP henceforth) extend FB's work by documenting that a "tent-shaped" linear combination of forward rates predicts bond excess returns, with much higher explanatory power than forward spreads. Following their approach, I regress the 2, 3, 4, and 5-year excess returns over one-year on the 1-year interest rate, and the 3 and 5-year forward rates using model-generated data:

$$rx_{t,t+4}^{(\tau)} = \alpha_{CP}^{(\tau)} + \beta_{CP,1}^{(\tau)} y_t^{(1)} + \beta_{CP,3}^{(\tau)} f_t^{(12)} + \beta_{CP,5}^{(\tau)} f_t^{(20)} + \upsilon_{t+4}^{(\tau)}. \quad (33)$$

Interest rates in the model are affine in the two-state variables r_t and μ_t, and in the wealth-share weighted average of ϵ_t^i. Thus, to avoid potential issues with multicollinearity, I only use two forward rates. CP further show that a single factor comprised of forward rates predicts excess returns. I replicate their two-stage "restricted" approach to see if the model generates a similar return-predicting factor. I first regress the average (across-maturity) excess returns on the forward rates, and recover the factor as the predicted values of this regression:

$$\sum_{\tau \in \{8,12,16,20\}} \frac{1}{4} rx_{t,t+4}^{(\tau)} = \gamma_0 + \gamma_1 y_t^{(1)} + \gamma_3 f_t^{(12)} + \gamma_5 f_t^{(20)} + \bar{v}_{t+4}. \tag{34}$$

I then regress each individual excess return on the recovered factor:

$$rx_{t,t+4}^{(\tau)} = b_{CP}^{(\tau)} \left(\widehat{\gamma}_0 + \widehat{\gamma}_1 y_t^{(1)} + \widehat{\gamma}_3 f_t^{(12)} + \widehat{\gamma}_5 f_t^{(20)} \right) + u_{t+4}^{(\tau)}, \tag{35}$$

where $\left(\widehat{\gamma}_0 + \widehat{\gamma}_1 y_t^{(1)} + \widehat{\gamma}_3 f_t^{(12)} + \widehat{\gamma}_5 f_t^{(20)} \right)$ are the predicted values from (34).

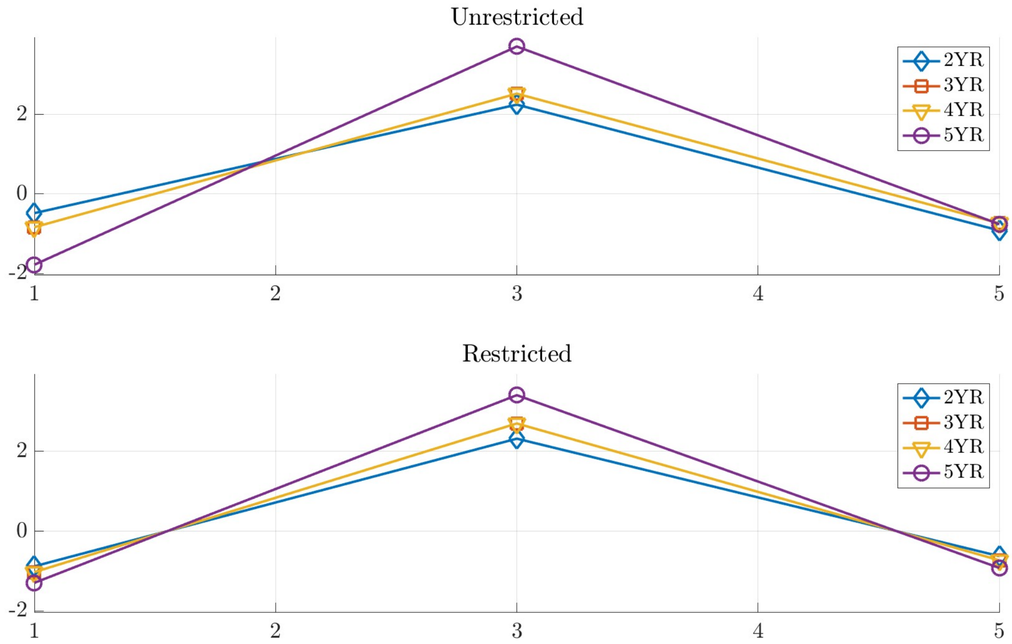

Figure 8: Cochrane-Piazzesi Factor

Notes: This figure plots the model-implied estimates of the CP regressions. Different markers and colors correspond to excess returns on bonds with different maturities. *Top panel:* the markers display the estimated $\beta_{CP,1}^{(\tau)}, \beta_{CP,3}^{(\tau)}, \beta_{CP,5}^{(\tau)}$ from the regression in (33). *Bottom panel:* the markers display the estimated $b_{CP}^{(\tau)}\gamma_1, b_{CP}^{(\tau)}\gamma_3, b_{CP}^{(\tau)}\gamma_5$ from the regressions in (34) and (35).

Figure 8 shows the results of the CP regressions using the model-generated data. Both the unrestricted and restricted estimates highlight a clear tent-shaped pattern, reproducing CP's findings. The model-generated CP factor significantly predicts excess returns, with much higher explanatory power than the FB regressions. The model's ability to produce a tent-shaped return predicting factor speaks to its success in generating excess return predictability that is consistent with empirical findings.

4.4 Bond Risk Premia

The results discussed in the previous section concern the "statistical" risk premia, estimated using full-sample information on interest rates. What about the real-time risk premia commanded by dealers? In this section, I examine the quarterly bond risk premia produced by the model.

First, I start by considering the model-implied risk premium in the rational expectations case. With rational expectations, since the aggregate endowment has constant volatility, the Expectations Hypothesis holds, and risk premia are constant. It is then natural to examine the contribution of subjective beliefs in generating time-varying risk premia.

Proposition 5. *Suppose all dealers have rational expectations, that is, $\epsilon_t^i = 0,\ i : 1, \dots, I$. Then the risk premium on a τ period bond is given by:*

$$rp_t^{(\tau),RE} = -C_r(\tau)C_\mu(\tau)\sigma_r\sigma_\mu + \sigma_Y\Big(C_r(\tau)\sigma_r + C_\mu(\tau)\sigma_\mu\Big). \tag{36}$$

Therefore, if dealers have rational expectations, the risk premium is constant. Moreover, the risk premium is increasing in maturity τ.

Proof. The result follows directly from setting $\epsilon_t^i = 0$ and computing the drift of bond returns given the bond prices in (20), which I formally show in B.5. □

With rational expectations, all dealers are identical and thus an identical representative dealer can be constructed. The market price of risk is then simply the covariance of the aggregate endowment with the pricing factors r_t and μ_t. The risk premium on a zero-coupon bond then depends on its loading on the pricing factors, as well as an adjustment term for the correlation of the factors.

The next proposition characterizes the impact of subjective beliefs on risk premia, relative to the rational expectations benchmark.

Proposition 6. *The subjective risk premium commanded by dealer i on a τ period bond,* *expressed under the econometrician's probability measure is given by:*

$$
\begin{aligned}
rp_t^{i,(\tau)} = rp_t^{(\tau),RE} & \underbrace{-\left(\sigma_\mu C_\mu(\tau) + \sigma_{\epsilon,r}^i C_\epsilon^i(\tau)\right)\frac{\kappa_r}{\sigma_r}\epsilon_t^i}_{\geqslant 0} \quad \underbrace{-\, C_r(\tau) C_\epsilon^i(\tau)\sigma_r\sigma_\epsilon^i}_{\text{sign}(\sigma_r\sigma_\epsilon^i)} \\
& \underbrace{+\, x_t^i C_\epsilon^i(\tau)\sigma_\epsilon^i \left(\frac{(\kappa_r\sigma_\mu + \kappa_\mu\sigma_r + \sigma_r\theta)}{(\kappa_r+\theta)(\kappa_\mu+\theta)}\right)\left(\frac{\sigma_r(\kappa_\epsilon+\theta)}{(\kappa_r\sigma_\epsilon^i - \kappa_\epsilon\sigma_r - \sigma_r\theta)}\right)}_{\geqslant 0}.
\end{aligned} \tag{37}
$$

Proof. Once again, the result follows directly from computing the drift of bond returns given the bond prices in (20), which I formally show in B.5. □

Comparing equations (36) and (37), it is straightforward to see that the only source of time-variation in risk premia is subjective beliefs. Beliefs affect risk premia both directly through the compensation required for the risk of having wrong beliefs about interest rate dynamics, and also indirectly through exposures. Unlike the rational expectations case, exposures are no longer constant or identical across agents. As can be seen from equation (30), they depend on the wealth-share weighted averages of beliefs and bond prices, as well as directly on dealers' individual beliefs, all of which vary over time.

Equation (37) describes the instantaneous risk compensation that a single type of dealer i to hold a long-term bond. The economy-wide risk premium under the econometrician's measure is given by a highly involved expression, as it also depends on the covariances of single-dealer economy bond returns with dealers' consumption shares. Figure 9 compares the bond risk premia on 2, 5, and 10-year bonds when dealers have subjective beliefs to the risk premia when dealers have rational expectations. All series are quarterly and are expressed in annualized percentages. When dealers have subjective beliefs, risk premia are time-varying and cyclical: risk premium on a bond starts to increase at the onset of recessions, and falls afterward. During the zero-lower bound (ZLB) period, risk premia are almost constant, similar to the premia under rational expectations.

Table 5 reports the summary statistics of risk premia on bonds with maturities 1, 2, 4, 5, 7, and 10-years, as well as the premia in the rational expectations case. The average risk premia with subjective beliefs are only slightly larger than the rational expectations counterparts, suggesting that subjective beliefs of dealers do not matter much for the average risk premium over the sample. This finding is not very surprising, as the estimation results suggest that subjective biases are highly transitory and quickly revert back to their long-run mean of zero. Accordingly, risk premia also quickly revert back to their rational expectations counterparts. However, subjective beliefs generate substantial

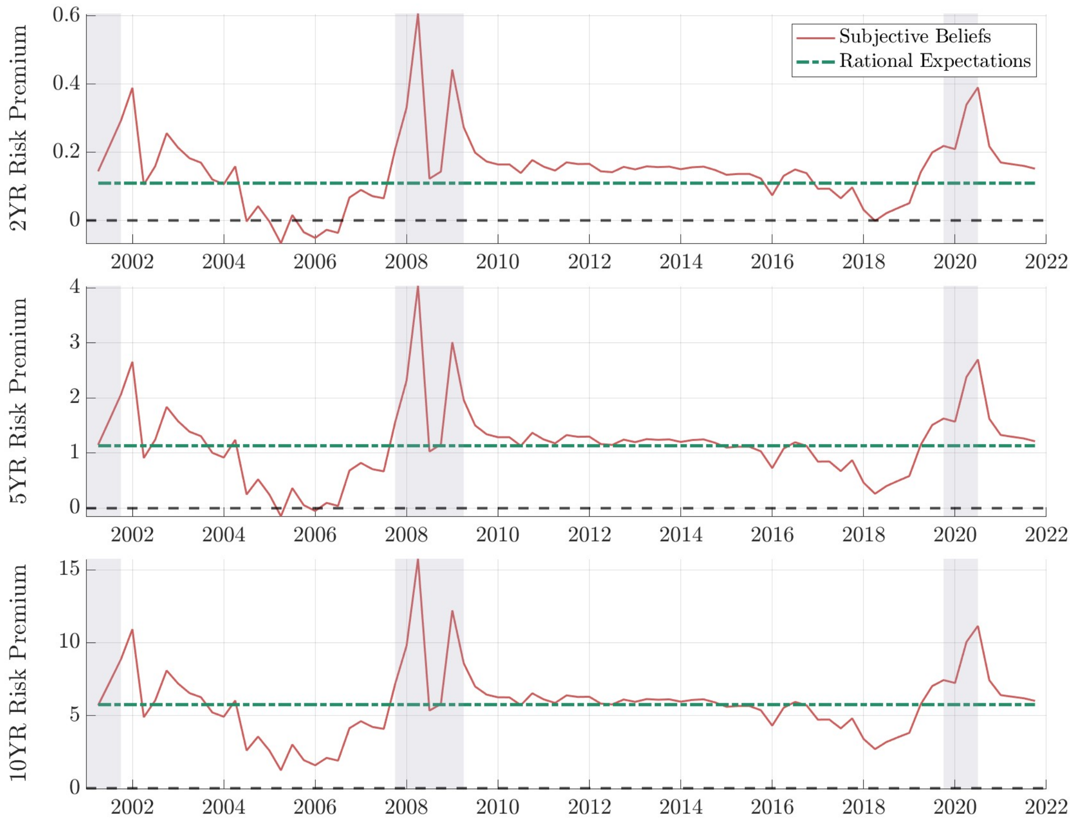

Figure 9: Bond Risk Premia

Notes: This figure displays the model-generated real-time bond risk premia, for bonds with maturities of 2, 5, and 10-years. Risk premia series are quarterly, and all values are expressed in annualized percentages. The red straight lines correspond to the risk premia of the baseline heterogeneous-dealer economy, computed under the econometrician's measure. The dashed green horizontal lines correspond to the risk premia with rational expectations, also computed under the econometrician's measure.

time-variation in bond risk premia, as their standard deviations are nearly as high as their average values for shorter maturities, and about 40% of its average for the 10-year bond. Due to the ZLB episode the distribution of premia is negatively skewed as evident by the comparison of the 25'th and 75'th percentiles.

In summary, the subjective beliefs of the largest bond dealers generate sizable time-variation in bond risk premia, the large volatility of long-term interest rates, and reproduce the empirical findings on return predictability. The main novelty underlying these results is that the subjective belief processes are disciplined by the portfolio positions of

Table 5: Summary Statistics of Bond Risk Premia

	Bond Maturity					
Summary Statistic (% ann.)	1-Year	2-Year	4-Year	5-Year	7-Year	10-Year
Mean	0.05	0.15	0.69	1.17	2.62	5.85
Std. Deviation	0.02	0.11	0.43	0.66	1.23	2.30
Median	0.05	0.15	0.70	1.20	2.66	5.94
25'th Percentile	0.04	0.09	0.47	0.85	2.01	4.72
75'th Percentile	0.05	0.17	0.70	1.33	2.91	6.40
Rational Expectations	0.02	0.11	0.64	1.13	2.57	5.76

Notes: This table displays the summary statistics of the model-generated real-time bond risk premia, for bonds with maturities of 1, 2, 4, 5, 7, and 10-years. Risk premia series are quarterly, and all values are expressed in annualized percentages. The bottom row displays the constant risk premia in the rational expectations economy.

dealers. My model connects the triad of large dealers' interest rate risk exposures, their subjective bond return forecasts, and time-variation in expected excess returns. I show that dealers' subjective beliefs are a sufficiently potent force to drive the empirical results on excess return predictability without implying counterfactual portfolio positions.

4.5 Average Dealers & Distribution of Wealth

Average Dealer Counterfactuals

Does heterogeneity in beliefs and exposures matter? Can the same risk premia dynamics be generated by an "average dealer", whose beliefs and exposures are the cross-sectional averages of dealer beliefs and exposures? To answer these questions, I consider two counterfactual exercises where there is only a single type of dealer in the economy.

First, I construct a "*weighted-average dealer*", whose beliefs and exposures at each date are the wealth-share weighted averages of dealers' subjective beliefs and interest rate risk exposures. That is, at time t, the weighted-average dealers have the beliefs and exposures defined as:

$$\epsilon_t^{w.avg.} = \sum_{i=1}^{N} \frac{c_t^i}{Y_t} \epsilon_t^i, \quad x_t^{w.avg.} = \sum_{i=1}^{N} \frac{c_t^i}{Y_t} x_t^i. \tag{38}$$

For the second counterfactual, I directly construct an "*average dealer*" in the same manner, but without weighting by wealth-shares.

Heterogenous dealers trade on their beliefs with each other, and in theory, require compensation for bearing the risk that their bond positions might redistribute some of their wealth-share to other dealers. This mechanism is missing from the weighted-average dealer economy. Therefore, the comparison of bond risk premia in an economy with only weighted-average dealers to the heterogeneous dealer baseline sheds light on whether *redistribution risk* plays a significant role in risk premia dynamics.

Even without redistribution risk, the distribution of dealers' wealth-shares, or equivalently wealth, might matter for risk premia. Recall from (19) that bond prices are wealth-share weighted averages of hypothetical single-dealer economy bond prices. Then, dealers with larger wealth-shares also have a larger influence on bond prices, and as a consequence on expected excess returns. Comparison of the weighted-average dealers economy with the average dealers economy reflects whether the distribution of dealer wealth impacts bond risk premia.

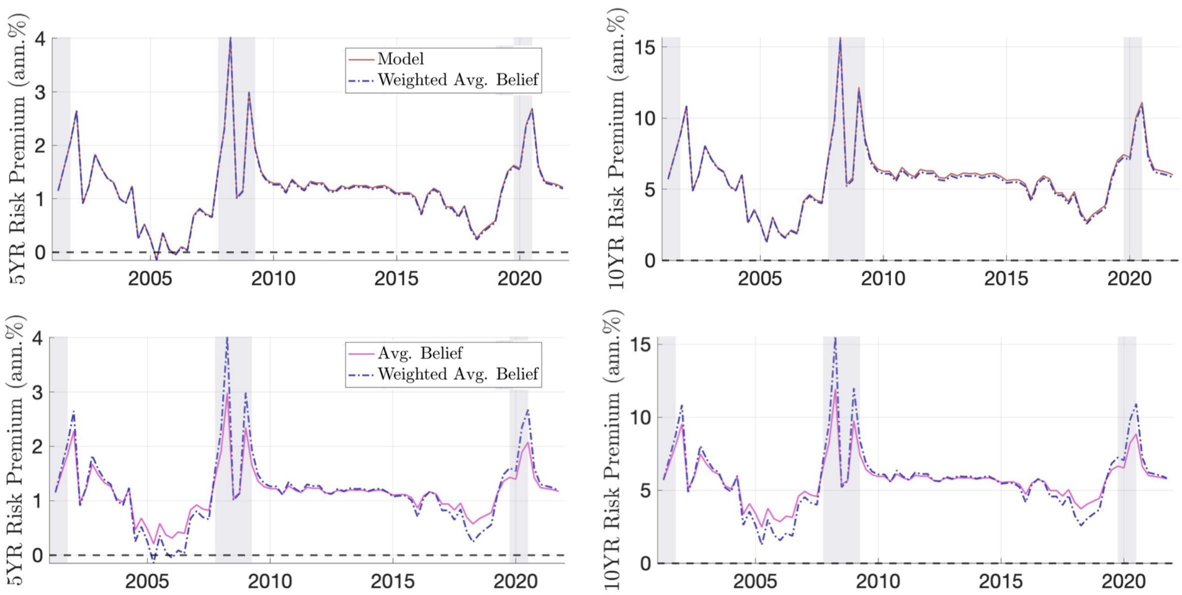

Figure 10: Bond Risk Premia: Weighted-Average Dealers vs. Average Dealers

Notes: This figure displays the risk premia on 5 and 10-year bonds, for the baseline case with heterogeneous dealers, for the weighted-average dealers economy, and for the average dealers economy. Risk premia series are quarterly, and all values are expressed in annualized percentages. *Top panel:* The red straight lines correspond to the risk premia of the baseline heterogeneous-dealer economy. The purple dashed lines correspond to the risk premia in the weighted-average dealers economy. Both series are computed under the econometrician's measure. *Bottom panel:* The purple dashed lines correspond to the risk premia in the weighted-average dealers economy. The magenta straight lines correspond to the risk premia in the average dealers economy. Both series are computed under the econometrician's measure.

Figure 10 demonstrates the comparison of the baseline model with heterogeneous dealers, and the two counterfactual cases. I focus on 5-year and 10-year bonds since for shorter maturities risk premia is small, making any comparison difficult to interpret. Looking at the top two panels reveals that redistribution risk essentially plays no role: the weighted-average dealer risk premia is nearly identical to the heterogeneous dealer premia. In other words, time-variation in bond risk premia is mainly driven by weighted-average beliefs and exposures, and cross-sectional heterogeneity hardly matters.

Yet, the bottom two panels of Figure 10 show that the distribution of wealth matters for bond risk premia. Premia are less volatile when the average dealer is constructed without taking wealth distribution into account: the annualized standard deviation of the 5-year (10-year) risk premium in the weighted-average dealers economy is 0.67% (2.30%), compared to 0.45% (1.53%) in the average dealers economy. The difference in premia is especially stark during recessions, as the average dealers' premia increases much less than those of weighted-average dealers.

How to interpret these findings? The decline in the volatility of bond risk premia in the average dealers economy suggests that the wealthier dealers are also the ones with more volatile subjective beliefs. This is indeed what I find when I estimate the model, as the largest belief volatilities reported on Table 4 are those of Bank of America and J.P. Morgan & Chase, the two largest dealers. The wealthiest dealers are also the most "pessimistic" ones, hence the large risk compensation they require is disproportionately reflected in the economy-wide risk premium.

Wealth Distribution

Finally, I plot the model implied evolution of the dealer wealth distribution in Figure 11. The model implies that J.P. Morgan earns high returns on its bond portfolios, and becomes the largest dealer surpassing Bank of America early in the sample. Wells Fargo starts with a slightly lower share than Citigroup, and eventually ends up with a slightly higher share. These patterns resemble the evolution of the gross asset holdings of these dealers over my sample period, so another success of the model is generating realistic wealth distribution patterns. However, in reality, there are many other factors that affect dealers' wealth distribution, such as J.P. Morgan's merger with Bank One in 2004, whereas my model attributes all of it to subjective beliefs. Thus, I do not compare the model-generated evolution of dealers' wealth distribution with the data counterpart, as it would be an unfair assessment of the model's capabilities.

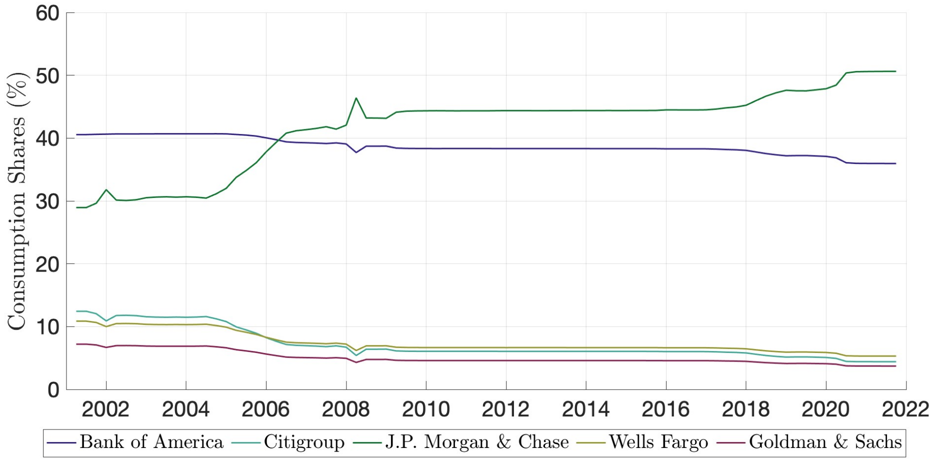

Figure 11: Wealth Distribution

Notes: This figure displays the evolution of dealers' consumption shares over time. Relative wealth-shares evolve according to the subjective belief differences of dealers, as described in equations (27) and (28).

5 Conclusion

In conclusion, this paper explores the dynamics of bond excess returns, particularly focusing on the role of large dealer banks and their subjective beliefs. I empirically document the comovement of primary dealers' interest-rate risk exposures and bond excess return forecasts, and show that forecasts and exposures are heterogeneous both in the cross-section of dealers and also over time.

By developing a quantitative-term structure model, that incorporates the heterogeneity of dealers' beliefs, this paper underscores the crucial link between subjective beliefs and the volatility of long-term interest rates. The model not only provides a theoretical framework for understanding the behavior of large dealers but also aligns these theoretical propositions with empirical data to generate realistic portfolio positions. One of the key findings of this paper is that the time-variation in the subjective beliefs of dealers serves as a potent mechanism for explaining the excess volatility of long-term rates and the predictability of excess bond returns. By running counterfactual exercises, I further find that the wealth distribution of dealers is important for the pass-through of subjective risk premia into market risk premia.

The empirical findings and the theoretical framework developed in this paper highlight the interplay between the subjective beliefs and actions of large dealer banks, contributing to a deeper understanding of the bond market's excess volatility and risk premia fluctuations. These findings have important implications for investors, policymakers, and market-participants seeking to make more informed decisions in bond markets, as well as for future studies that aim to explore the implications of subjective beliefs while incorporating portfolio evidence.

References

Abadie, A., S. Athey, G. W. Imbens, and J. M. Wooldridge (2023): "When should you adjust standard errors for clustering?" *The Quarterly Journal of Economics*, 138, 1–35.

Adrian, T., E. Etula, and T. Muir (2014): "Financial intermediaries and the cross-section of asset returns," *The Journal of Finance*, 69, 2557–2596.

Adrian, T., E. Moench, and H. S. Shin (2016): "Dynamic leverage asset pricing," .

Allen, J. and M. Wittwer (2023): "Intermediary capitalization and asset demand," Tech. rep.

Anderson, E. W., E. Ghysels, and J. L. Juergens (2005): "Do heterogeneous beliefs matter for asset pricing?" *The Review of Financial Studies*, 18, 875–924.

Andrei, D., B. Carlin, and M. Hasler (2019): "Asset pricing with disagreement and uncertainty about the length of business cycles," *Management Science*, 65, 2900–2923.

Balduzzi, P., S. R. Das, and S. Foresi (1998): "The central tendency: A second factor in bond yields," *Review of Economics and Statistics*, 80, 62–72.

Bansal, R. and I. Shaliastovich (2013): "A long-run risks explanation of predictability puzzles in bond and currency markets," *The Review of Financial Studies*, 26, 1–33.

Bansal, R. and A. Yaron (2004): "Risks for the long run: A potential resolution of asset pricing puzzles," *The journal of Finance*, 59, 1481–1509.

Basak, S. (2000): "A model of dynamic equilibrium asset pricing with heterogeneous beliefs and extraneous risk," *Journal of Economic Dynamics and Control*, 24, 63–95.

——— (2005): "Asset pricing with heterogeneous beliefs," *Journal of Banking & Finance*, 29, 2849–2881.

Begenau, J., M. Piazzesi, and M. Schneider (2015): "Banks' risk exposures," Tech. rep., National Bureau of Economic Research.

Bhamra, H. S. and R. Uppal (2014): "Asset prices with heterogeneity in preferences and beliefs," *The Review of Financial Studies*, 27, 519–580.

Borovička, J. (2020): "Survival and long-run dynamics with heterogeneous beliefs under recursive preferences," *Journal of Political Economy*, 128, 206–251.

Boyarchenko, N., D. O. Lucca, and L. Veldkamp (2021): "Taking orders and taking notes: Dealer information sharing in treasury auctions," *Journal of Political Economy*, 129, 607–645.

Brunnermeier, M. K. and Y. Sannikov (2014): "A macroeconomic model with a financial sector," *American Economic Review*, 104, 379–421.

Buraschi, A. and A. Jiltsov (2006): "Model uncertainty and option markets with heterogeneous beliefs," *The Journal of Finance*, 61, 2841–2897.

Campbell, J. Y. and J. H. Cochrane (1999): "By force of habit: A consumption-based explanation of aggregate stock market behavior," *Journal of political Economy*, 107, 205–251.

Chen, H., S. Joslin, and N.-K. Tran (2010): "Affine disagreement and asset pricing," *American Economic Review*, 100, 522–526.

——— (2012): "Rare disasters and risk sharing with heterogeneous beliefs," *The Review of Financial Studies*, 25, 2189–2224.

Chiarella, C. and X.-Z. He (2002): "Heterogeneous beliefs, risk and learning in a simple asset pricing model," *Computational Economics*, 19, 95–132.

Chien, Y., H. Cole, and H. Lustig (2016): "Implications of heterogeneity in preferences, beliefs and asset trading technologies in an endowment economy," *Review of Economic Dynamics*, 20, 215–239.

Cochrane, J. H. and M. Piazzesi (2005): "Bond risk premia," *American economic review*, 95, 138–160.

Coimbra, N. and H. Rey (2021): "Financial Cycles with Heterogeneous Intermediaries," .

Cox, J. C. and C.-f. Huang (1989): "Optimal consumption and portfolio policies when asset prices follow a diffusion process," *Journal of economic theory*, 49, 33–83.

Cvitanić, J., E. Jouini, S. Malamud, and C. Napp (2012): "Financial markets equilibrium with heterogeneous agents," *Review of Finance*, 16, 285–321.

Cvitanić, J. and S. Malamud (2011): "Price impact and portfolio impact," *Journal of Financial Economics*, 100, 201–225.

David, A. (2008): "Heterogeneous beliefs, speculation, and the equity premium," *The Journal of Finance*, 63, 41–83.

Dell'Ariccia, G. and R. Marquez (2013): "Interest rates and the bank risk-taking channel," *Annu. Rev. Financ. Econ.*, 5, 123–141.

Detemple, J. and S. Murthy (1994): "Intertemporal asset pricing with heterogeneous beliefs," *Journal of Economic Theory*, 62, 294–320.

Drechsler, I., A. Savov, and P. Schnabl (2018): "A model of monetary policy and risk premia," *The Journal of Finance*, 73, 317–373.

Du, W., B. M. Hébert, and W. Li (2022): "Intermediary balance sheets and the treasury yield curve," Tech. rep., National Bureau of Economic Research.

Duffie, D. (2010): *Dynamic asset pricing theory*, Princeton University Press.

Dumas, B., A. Kurshev, and R. Uppal (2009): "Equilibrium portfolio strategies in the presence of sentiment risk and excess volatility," *The Journal of Finance*, 64, 579–629.

Fama, E. F. and R. R. Bliss (1987): "The information in long-maturity forward rates," *The American Economic Review*, 680–692.

Fan, M. (2006): "Heterogeneous beliefs, the term structure and time-varying risk premia," *Annals of Finance*, 2, 259–285.

Gabaix, X. and M. Maggiori (2015): "International liquidity and exchange rate dynamics," *The Quarterly Journal of Economics*, 130, 1369–1420.

Giacoletti, M., K. T. Laursen, and K. J. Singleton (2021): "Learning from disagreement in the us treasury bond market," *The Journal of Finance*, 76, 395–441.

Giglio, S., M. Maggiori, J. Stroebel, and S. Utkus (2021): "Five facts about beliefs and portfolios," *American Economic Review*, 111, 1481–1522.

Gomez, M., A. Landier, D. Sraer, and D. Thesmar (2021): "Banks' exposure to interest rate risk and the transmission of monetary policy," *Journal of Monetary Economics*, 117, 543–570.

Greenwood, R., S. G. Hanson, and G. Y. Liao (2018): "Asset price dynamics in partially segmented markets," *The Review of Financial Studies*, 31, 3307–3343.

Haddad, V. and T. Muir (2021): "Do intermediaries matter for aggregate asset prices?" *The Journal of Finance*, 76, 2719–2761.

Haddad, V. and D. Sraer (2020): "The banking view of bond risk premia," *The Journal of Finance*, 75, 2465–2502.

He, Z., B. Kelly, and A. Manela (2017): "Intermediary asset pricing: New evidence from many asset classes," *Journal of Financial Economics*, 126, 1–35.

He, Z. and A. Krishnamurthy (2013): "Intermediary asset pricing," *American Economic Review*, 103, 732–770.

He, Z., S. Nagel, and Z. Song (2022): "Treasury inconvenience yields during the COVID-19 crisis," *Journal of Financial Economics*, 143, 57–79.

Jouini, E. and C. Napp (2006): "Heterogeneous beliefs and asset pricing in discrete time: An analysis of pessimism and doubt," *Journal of Economic Dynamics and Control*, 30, 1233–1260.

——— (2007): "Consensus consumer and intertemporal asset pricing with heterogeneous beliefs," *The Review of Economic Studies*, 74, 1149–1174.

Kargar, M. (2021): "Heterogeneous intermediary asset pricing," *Journal of Financial Economics*, 141, 505–532.

Kekre, R. and M. Lenel (2022): "Monetary policy, redistribution, and risk premia," *Econometrica*, 90, 2249–2282.

Kekre, R., M. Lenel, and F. Mainardi (2022): "Monetary Policy, Segmentation, and the Term Structure," .

Kim, D. H. and J. H. Wright (2005): "An arbitrage-free three-factor term structure model and the recent behavior of long-term yields and distant-horizon forward rates," .

Kogan, L., S. A. Ross, J. Wang, and M. M. Westerfield (2006): "The price impact and survival of irrational traders," *The Journal of Finance*, 61, 195–229.

Leombroni, M., M. Piazzesi, M. Schneider, and C. Rogers (2020): "Inflation and the price of real assets," Tech. rep., National Bureau of Economic Research.

Lewellen, J. (2004): "Predicting returns with financial ratios," *Journal of Financial Economics*, 74, 209–235.

Liu, Y. and J. C. Wu (2021): "Reconstructing the yield curve," *Journal of Financial Economics*, 142, 1395–1425.

Maggiori, M. (2017): "Financial intermediation, international risk sharing, and reserve currencies," *American Economic Review*, 107, 3038–3071.

Mishkin, F. and S. Eakins (2009): "Financial Markets and Institutions," .

Morelli, J. M., P. Ottonello, and D. J. Perez (2022): "Global banks and systemic debt crises," *Econometrica*, 90, 749–798.

Nagel, S. and Z. Xu (2023): "Dynamics of subjective risk premia," *Journal of Financial Economics*, 150, 103713.

Newey, W. K. and K. D. West (1987): "Hypothesis testing with efficient method of moments estimation," *International Economic Review*, 777–787.

Piazzesi, M. (2010): "Affine term structure models," in *Handbook of financial econometrics: Tools and Techniques*, Elsevier, 691–766.

Piazzesi, M., J. Salomao, and M. Schneider (2015): "Trend and Cycle in Bond Premia," .

Protter, P. E. (2005): *Stochastic differential equations*, Springer.

Sandroni, A. (2000): "Do markets favor agents able to make accurate predictions?" *Econometrica*, 68, 1303–1341.

Scheinkman, J. A. and W. Xiong (2003): "Overconfidence and speculative bubbles," *Journal of political Economy*, 111, 1183–1220.

Schneider, A. (2023): "Financial Intermediaries and the Yield Curve," .

Stambaugh, R. F. (1999): "Predictive regressions," *Journal of financial economics*, 54, 375–421.

Vayanos, D. and J.-L. Vila (2021): "A preferred-habitat model of the term structure of interest rates," *Econometrica*, 89, 77–112.

Wachter, J. A. (2006): "A consumption-based model of the term structure of interest rates," *Journal of Financial economics*, 79, 365–399.

Xiong, W. and H. Yan (2010): "Heterogeneous expectations and bond markets," *The Review of Financial Studies*, 23, 1433–1466.

Yan, H. (2008): "Natural selection in financial markets: Does it work?" *Management Science*, 54, 1935–1950.

Zapatero, F. (1998): "Effects of financial innovations on market volatility when beliefs are heterogeneous," *Journal of Economic Dynamics and Control*, 22, 597–626.

A Data Appendix

A.1 Approximating Expected Excess Returns

Let $rpy_t^{i,(\tau)}$ denote the subjective term premium of dealer i at time t. No-arbitrage implies that the interest rate forecast of a τ period bond can be decomposed as:

$$E_t^i\left[y_t^{(\tau)}\right] = \frac{1}{\tau}E_t^i\left[r_t + r_{t+1} + \ldots + r_{t+\tau-1}\right] + rpy_t^{i,(\tau)} \tag{39}$$

Then the forecast for the one-quarter ahead rate on the same maturity is:

$$E_t^i\left[y_{t+1}^{(\tau)}\right] = \frac{1}{\tau}E_t^i\left[r_{t+1} + \ldots + r_{t+\tau-1} + r_{t+\tau}\right] + E_t^i\left[rpy_{t+1}^{i,(\tau)}\right] \tag{40}$$

and

$$E_t^i\left[y_{t+1}^{(\tau-1)}\right] = \frac{1}{\tau-1}E_t^i\left[r_{t+1} + \ldots + r_{t+\tau-1}\right] + E_t^i\left[rpy_{t+1}^{i,(\tau-1)}\right] \tag{41}$$

Now, let r_{t+H} denote the farthest short-rate forecast available in the surveys. That is, $E_t^i[r_{t+H}]$ is dealer i's short-rate forecast for 5-quarters ahead. Denote $e_t^{(\tau),i} = E_t^i[r_{t+\tau}] - E_t^i[r_{t+H}]$. This residual represents the difference between the farthest available forecast of dealer i, and their expectation of the τ-quarters ahead short-rate. Then (39) and (40) can be rewritten as:

$$E_t^i\left[y_t^{(\tau)}\right] - \frac{1}{\tau}E_t^i\left[r_t\right] = \frac{1}{\tau}E_t^i\left[r_{t+1} + \ldots + r_{t+\tau-1}\right] + rpy_t^{i,(\tau)} \tag{42}$$

$$E_t^i\left[y_{t+1}^{(\tau)}\right] - \frac{1}{\tau}E_t^i[r_{t+H}] = \frac{1}{\tau}E_t^i\left[r_{t+1} + \ldots + r_{t+\tau-1}\right] + \frac{1}{\tau}e_t^{(\tau),i} + E_t^i\left[rpy_{t+1}^{i,(\tau)}\right] \tag{43}$$

Subtracting both sides of (42) from (43), I obtain:

$$\left(E_t^i\left[y_{t+1}^{(\tau)}\right] - \frac{1}{\tau}E_t^i[r_{t+H}]\right) - \left(E_t^i\left[y_t^{(\tau)}\right] - \frac{1}{\tau}E_t^i\left[r_t\right]\right) = \frac{1}{\tau}e_t^{(\tau),i} + E_t^i\left[rpy_{t+1}^{i,(\tau)}\right] - rpy_t^{i,(\tau)} \tag{44}$$

The assumption that the law of iterated expectations holds under each forecaster's beliefs then implies $E_t^i\left[rpy_{t+1}^{i,(\tau)}\right] - rpy_t^{i,(\tau)} = 0$, hence using (44), I recover the residual $e_t^{(\tau),i}$ and construct $E_t^i[r_{t+\tau}] = E_t^i[r_{t+H}] + e_t^{(\tau),i}$. Subtracting $\frac{1}{\tau}E_t^i[r_{t+\tau}]$ from $E_t^i\left[y_{t+1}^{(\tau)}\right]$ and multiplying by $\frac{\tau}{\tau-1}$ yields:

$$\begin{aligned} &\frac{1}{\tau-1}E_t^i\left[r_{t+1}+\ldots+r_{t+\tau-1}\right]+\frac{\tau}{\tau-1}E_t^i\left[rpy_{t+1}^{i,(\tau)}\right] \\ &=E_t^i\left[y_{t+1}^{(\tau-1)}\right]-E_t^i\left[rpy_{t+1}^{i,(\tau-1)}\right]+\frac{\tau}{\tau-1}E_t^i\left[rpy_{t+1}^{i,(\tau)}\right] \end{aligned} \tag{45}$$

With the law of iterated expectations, the term-premia terms on the right-hand side of this equality equals:

$$-E_t^i\left[rpy_{t+1}^{i,(\tau-1)}\right]+\frac{\tau}{\tau-1}E_t^i\left[rpy_{t+1}^{i,(\tau)}\right]=-rpy_t^{i,(\tau-1)}+\frac{\tau}{\tau-1}rpy_t^{i,(\tau)} \tag{46}$$

Straightforward algebraic manipulation yields the second assumption:

$$\left(\frac{\tau}{\tau-1}\frac{rpy_t^{i,(\tau)}}{rpy_t^{i,(\tau-1)}}-1\right)rpy_t^{i,(\tau-1)}\approx 0 \tag{47}$$

For longer maturities, both the contribution to term premium from this holding period and the ratio $\frac{\tau}{\tau-1}$ are small, hence the approximation improves in accuracy as bond maturity increases. For smaller maturities, the quarterly term premium on the bond itself is likely to be small. As a quantified example, the average magnitude of the 2-year statistical quarterly term-premium estimated using the methodology in Kim and Wright (2005) over the period 2001:Q1-2021:Q4 is approximately 0.21%.[19] Even if $\frac{rpy_t^{i,(4)}}{rpy_t^{i,(3)}}=1.5$, that is, just the term-premia due to the holding period of $t+3$ to $t+4$ equals half of the premium on a 1.75-year bond, the approximation error in (47) would only be 0.004.

A.2 Additional Results

Actual Short-Rate

In this section, instead of using the forecasts of the average 3-month Treasury interest rate as the short-term interest rate for computing excess returns as in (1), I use the actual 3-month Treasury interest rates. The 3-month interest rate data comes from Liu and Wu (2021), and I use the last available observation for the month preceding each survey.

[19]Data comes from the series [THREEFYTP2], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/THREEFYTP2

Thus, subjective expected excess returns computed this way are expressed as:

$$E_t^i[\overline{rx}_{t+1,t+2}^{(\tau)}] = -(\tau-1)\underbrace{E_t^i[\overline{y}_{t+1,t+2}^{(\tau-1)}]}_{\substack{\textit{forecast of } \tau-1\\ \textit{quarter interest}\\ \textit{rate}}} + \tau \underbrace{E_t^i[\overline{y}_{t,t+1}^{(\tau)}]}_{\substack{\textit{forecast of } \tau\\ \textit{quarter interest}\\ \textit{rate}}} - \underbrace{r_t}_{\substack{\textit{3-month}\\ \textit{interest rate}}}. \tag{48}$$

I then run the regression (5) again using the subjective expected excess returns in (48). Table 6 reports the estimates. The coefficients are slightly larger than those in Table 1, with the magnitude of increase declining in maturity. Coefficients in columns (2) and (3) for long maturities are all statistically significant. Thus, using actual short-rates instead of forecasts of the 3-month interest rates does not alter the conclusions of the empirical exercise in this paper.

Table 6: Exposures & Subjective Expected Excess Returns – Actual Short-Rate

	Exposure		
	3-Month Treasury		
	(1)	(2)	(3)
1 YR	−1.524	−1.612	−1.612*
	(2.120)	(2.064)	(1.812)
2 YR	2.592	2.860***	2.860***
	(2.408)	(1.082)	(1.357)
5 YR	1.188	1.056***	1.056**
	(0.520)	(0.448)	(0.436)
10 YR	0.544**	0.428**	0.428*
	(0.212)	(0.184)	(0.164)
30 YR	0.212**	0.140***	0.140**
	(0.068)	(0.056)	(0.040)
Avg. R^2	0.83	0.87	0.88
Observations	323	323	323
Dealer FE	*YES*	*YES*	*YES*
Time Trend		*YES*	*YES*
Cluster Robust SE			*YES*

Notes: This table displays the estimation results from the regression: $Exposure_{t,i} = \alpha + fe_i + \kappa t + \beta^{(\tau)} E_t^i[\overline{rx}_{t+1,t+2}^{(\tau)}] + \varepsilon_{t,i}$. Standard errors are in parentheses. Stars indicate significance at ***: 99%, **: 95%, *: 90% confidence levels.

Figure 12 displays the coefficient estimates from this regression after scaling by the standard deviation of expected excess returns. The stronger increase in the shorter-term coefficients is reflected here in the figure, as the coefficient on the 2-year expected excess returns is considerably larger than those for other maturities, unlike the pattern in 3. Hence, replacing the forecasted short-rate with the actual one mainly has an impact on the results concerning short-term maturities.

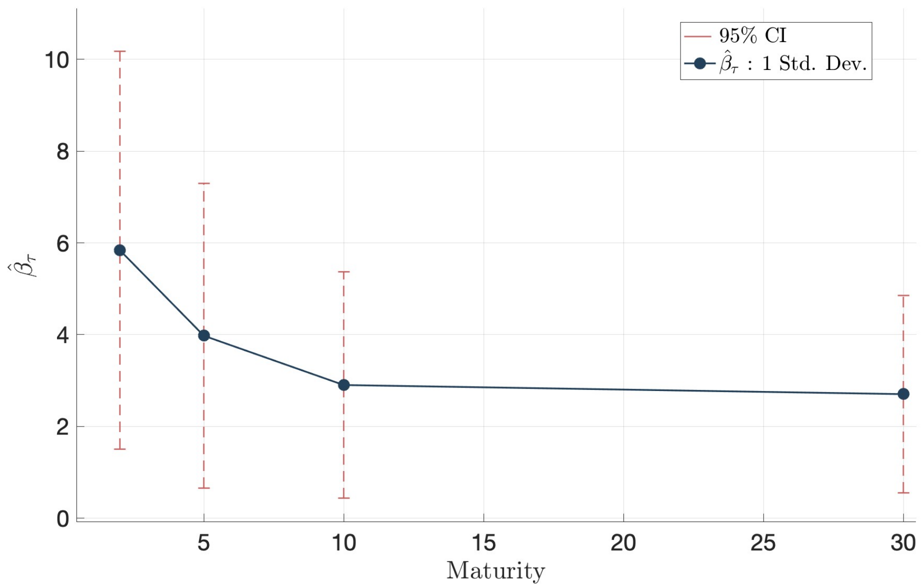

Figure 12: Standardized $\beta^{(\tau)}$ Estimates – Actual Short-Rate

Notes: The dark blue dots represent the regression coefficients from Table 1, scaled by the sample standard deviation of the subjective expected excess returns for the corresponding maturity. The vertical dashed red lines display the scaled 95% confidence intervals.

CRSP/Compustat Data

In addition to the FR Y-9C filings, I alternatively use the CRSP/Compustat database obtained from Wharton Research Data Services (WRDS).[20] CRSP/Compustat is a comprehensive financial database that provides a wide range of financial, statistical, and market information about publicly traded companies. I construct the exposure measure (4) using the analogous balance sheet items reported in CRSP/Compustat as follows. $Assets^i_{\leqslant 1}$ is the item "Cash and Short-Term Investments". $Liabilities^i_{\leqslant 1}$ comprises "Debt in Current Liabilities", "Long-Term Debt Due in One Year" and "Preferred/Preference Stock (Capital) - Total". I again scale them by "Total Assets".

[20]See https://wrds-www.wharton.upenn.edu/pages/about/data-vendors/center-for-research-in-security-prices-crsp/.

Table 7 presents the estimates from the main regression (5) using exposures constructed with CRSP/Compustat data. The slope coefficients in column (2) are slightly larger than their counterparts in Table 1, albeit less statistically significant. The coefficients for the 2-year expected excess returns lose significance at 90% confidence level when using Liang-Zeger standard errors or the Federal Funds Rate as the short-term borrowing rate. Yet, overall the estimates display a similar pattern as in Table 1, and suggest positive co-movement between risk exposures and subjective expected excess returns.

Table 7: Exposures & Subjective Expected Excess Returns – CRSP/Compustat Data

	Exposure					
	3-Month Treasury			Fed Funds Rate		
	(1)	(2)	(3)	(1)	(2)	(3)
1 YR	−0.705	−1.748	−1.748*	−1.640	−2.496	−2.496*
	(5.426)	(5.582)	(5.483)	(5.527)	(5.619)	(6.230)
2 YR	1.456	3.097*	3.097	−2.390	3.039	3.039
	(4.517)	(2.420)	(4.029)	(4.731)	(2.585)	(4.433)
5 YR	1.374*	0.954*	0.954**	0.400	0.978**	0.978*
	(0.984)	(0.588)	(0.681)	(1.002)	(0.572)	(0.732)
10 YR	0.990**	0.453**	0.453*	0.795**	0.490**	0.490*
	(0.449)	(0.263)	(0.344)	(0.452)	(0.256)	(0.367)
30 YR	0.718***	0.189**	0.189**	0.582***	0.191*	0.191**
	(0.146)	(0.101)	(0.125)	(0.156)	(0.090)	(0.115)
Avg. R^2	0.68	0.83	0.83	0.66	0.86	0.86
Observations	337	337	337	396	396	396
Dealer FE	*YES*	*YES*	*YES*	*YES*	*YES*	*YES*
Time Trend		*YES*	*YES*		*YES*	*YES*
Cluster Robust SE			*YES*			*YES*

Notes: This table displays the estimation results from the regression:

$$Exposure_{t,i} = \alpha + fe_i + \kappa t + \beta^{(\tau)} E_t^i\big[\overline{rx}_{t+1,t+2}^{(\tau)}\big] + \varepsilon_{t,i}$$

Standard errors are in parentheses. Stars indicate significance at ***: 99%, **: 95%, *: 90% confidence levels.

B Model Derivations

B.1 Belief Disagreement

Let $(\mathcal{I}, \mathcal{F}^{\mathcal{I}}, d\iota)$ denote the measured-space of dealers. The effective probability space of dealers of type i is $\left(I \times \Omega, \mathcal{I} \boxtimes \mathcal{F}_t^i, d\iota \boxtimes \mathcal{P}_t^i\right)$,[21] with $\mathcal{P}_t^i$ representing the probability measure implied by the information filtration of type i dealers. Define a random variable $\frac{d\mathcal{P}^i}{\mathcal{P}}$ such that $\mathbb{E}^i[\frac{d\mathcal{P}^i}{\mathcal{P}}] = 1$ and

$$\mathbb{E}^i[X] = \mathbb{E}\left[\frac{d\mathcal{P}^i}{d\mathcal{P}} X\right] \tag{49}$$

The random variable $\frac{d\mathcal{P}^i}{d\mathcal{P}}$ is the Radon-Nikodym derivative of dealer i's probability measure with respect to the "true" statistical measure. Define the density process $\xi_t^i = \frac{d\mathcal{P}^i}{d\mathcal{P}}\Big|_{\mathcal{F}_t^i}$ such that:

$$\mathbb{E}_t^i[X_s] = \frac{\mathbb{E}_t[\xi_s^i X_s]}{\mathbb{E}_t[\xi_s^i]} = \frac{\mathbb{E}_t[\xi_s^i X_s]}{\xi_t}, \quad s > t \tag{50}$$

Then by Girsanov Theorem, I have:

$$\frac{d\xi_t^i}{\xi_t^i} = \epsilon_t^i \frac{\kappa_r}{\sigma_r} dB_t \tag{51}$$

Thus, belief disagreement is characterized by the Radon-Nikodym derivative ξ_t^i, and ϵ_t^i acts as its stochastic volatility component.

B.2 Individual Optimization & HJB

In what follows, I will drop the i superscripts when convenient for ease of exposition, since all dealers of a particular type solve the same problem given their beliefs. First of all, define portfolio shares

$$\alpha_t^{(\tau)} = \frac{h_t^{(\tau)}}{w_t} \tag{52}$$

With log-utility, the consumption-wealth ratio is constant: $c_t = \varsigma w_t$. Then the problem can be rewritten as:

$$V_t\left(w_t^i\right) = \max_{\left\{\left\{\alpha_{t+s}^{(\tau),i}\right\}_{\tau\in(0,\infty)}\right\},} \mathbb{E}_t^i\left[\int_t^\infty e^{-\rho s} \log(\varsigma w_s^i) ds\right] \tag{53}$$

[21] I consider the Fubini extension of the probability space as in Sun (2006).

subject to

$$dw_t^i = w_t^i(r_t - \varsigma)dt + \int_0^\infty \alpha_t^{(\tau),i} w_t^i \left(\frac{dP_t^{(\tau)}}{P_t^{(\tau)}} - r_t dt \right) d\tau \tag{54}$$

and

$$dr_t = -\kappa_r(r_t - \mu_t^i)dt + \sigma_r dB_t \tag{55}$$

Now, conjecture an equilibrium pricing function that follows the Itô process:

$$\frac{dP_t^{(\tau)}}{P_t^{(\tau)}} = \ell_t(\tau)dt + \zeta_{r,t}^\tau dB_t \tag{56}$$

for some functions $\ell(\tau), \zeta_{r,t}^\tau, \zeta_{\epsilon,t}^\tau$ which will be determined in equilibrium. I will use $\widehat{\cdot}$ to indicate that a variable is expressed under the beliefs of type i. Conjecture that the value function of dealer i is:

$$J(\tau, w_t, \boldsymbol{X}_t, r_t, \mu_t^i) = \log w_t + G_t\Big(\tau, \boldsymbol{X}_t, r_t, \mu_t^i\Big) \tag{57}$$

Due to heterogeneous beliefs, each dealer's wealth is an endogenous state variable of the model. For ease of notation, I use $\boldsymbol{X}_t$ to denote the $S \times 1$ vector of (endogenous) state variables other than the dealer wealth, the real short-rate, and its stochastic drift, and x_t^s to denote a particular state-variable.

Then I can write down the HJB as:

$$\sup_{\alpha_{t+s}^{(\tau)}} \Bigg\{ \frac{\partial J}{\partial w}\Bigg(w_t(r_t - \varsigma) + \int_0^\infty \alpha_t^{(\tau)} w_t \left(\widehat{\ell}_t(\tau) - r_t\right) d\tau \Bigg) + \sum_{s=1}^S \frac{\partial J}{\partial x_t^s}\mu_{x,t}^s - \frac{\partial J}{\partial r}\kappa_r(r_t - \mu_t^i) - \frac{\partial J}{\partial \mu_t^i}\kappa_\mu(\mu_t^i - \bar{\mu})$$
$$+ \frac{1}{2}\frac{\partial^2 J}{\partial w^2} w_t^2 \left(\int_0^\infty \alpha_t^{(\tau)} \widehat{\zeta}_{r,t}^\tau d\tau \right)^2 + \sum_{s=1}^S \sum_{s=1}^S \frac{1}{2}\frac{\partial^2 J}{\partial x_t^{s2}} \sigma_{x,t}^{s2} + \frac{1}{2}\frac{\partial^2 J}{\partial r^2}\sigma_r^2 + \frac{1}{2}\frac{\partial^2 J}{\partial \mu_t^{i2}} + u(c) \Bigg\} = -\frac{\partial J}{\partial t} \tag{58}$$

where $u(c) = \log(c)$. Local optimization with respect to $\alpha_{t+s}^{(\tau)}$ yields the first-order conditions:

$$w_t\Big(\widehat{\ell}_t(\tau) - r_t\Big)\frac{\partial J}{\partial w} = -w_t^2 \left(\int_0^\infty \alpha_t^{(u)} \widehat{\zeta}_{r,t}^\tau \widehat{\zeta}_{r,t}^u du \right) \frac{\partial^2 J}{\partial w^2}, \quad \forall \tau \in (0, \infty) \tag{59}$$

Plugging in my conjecture for J, I get:

$$\widehat{\ell}_t(\tau) - r_t = \left(\int_0^\infty \alpha_t^{(u)} \widehat{\zeta}_{r,t}^\tau \widehat{\zeta}_{r,t}^u du \right), \quad \forall \tau \in (0, \infty) \tag{60}$$

Optimal portfolio shares do not directly depend on dealer wealth. The dependence only arises from how the wealth distribution affects the bond return process: given the return process, portfolio shares are independent of the individual dealer wealth.

Using the definitions of $\ell_t(\tau)$ and $\zeta^{\tau}_{r,t}$, one can rewrite the first-order condition as:

$$\mathbb{E}^i_t\left(\frac{dP^{(\tau)}_t}{P^{(\tau)}_t}\right) - r_t = \int_0^\infty \alpha^{(u),i}_t Cov^i_t\left(\frac{dP^{(\tau)}_t}{P^{(\tau)}_t}, \frac{dP^{(u)}_t}{P^{(u)}_t}\right) du \tag{61}$$

or equivalently

$$\mathbb{E}^i_t\left(\frac{dP^{(\tau)}_t}{P^{(\tau)}_t}\right) - r_t = \frac{1}{w^i_t}\int_0^\infty h^{(u),i}_t Cov^i_t\left(\frac{dP^{(\tau)}_t}{P^{(\tau)}_t}, \frac{dP^{(u)}_t}{P^{(u)}_t}\right) du \tag{62}$$

That is, the portfolio share of the bond with maturity τ is determined by the *subjective expected excess return* on the bond,[22] and the *subjective covariance* of the bond return with total bond-portfolio return.[23]

At optimum, the envelope condition implies $\varsigma = \rho$. That is the time-preference rate ρ is also equal to the dividend payout ratio. Plugging in the first order conditions back into the HJB and canceling out the $\log(w_t)$ terms I get:

$$\begin{aligned} &G_t\left(r_t - \rho + \int_0^\infty \alpha^{(\tau)}_t \left(\widehat{\ell}_t(\tau) - r_t\right) d\tau\right) + \sum_{s=1}^{S} \frac{\partial G}{\partial x^s_t}\mu^s_{x,t} - \frac{\partial G}{\partial r}\kappa_r(r_t - \mu^i_t) - \frac{\partial G}{\partial \mu^i_t}\kappa_\mu(\mu^i_t - \bar{\mu}) \\ &+ \frac{\partial G}{\partial t} - \frac{1}{2}\left(\int_0^\infty \alpha^{(\tau)}_t \widehat{\zeta}^{\tau}_{r,t} d\tau\right)^2 + \sum_{s=1}^{S}\sum_{s=1}^{S} \frac{1}{2}\frac{\partial^2 G}{\partial x^{s2}_t}\sigma^{s2}_{x,t} + \frac{1}{2}\frac{\partial^2 G}{\partial r^2}\sigma^2_r + \frac{1}{2}\frac{\partial^2 G}{\partial \mu^{i2}_t} + \log(\rho) = 0 \end{aligned} \tag{63}$$

G_t solves this equation, which does not depend on individual dealer wealth, verifying my value function conjecture. Finally, the state price density takes the form:

$$\frac{d\pi_t}{\pi_t} = -r_t dt - \left(\int_0^\infty \alpha^{(\tau)}_t \zeta^{\tau}_{r,t} d\tau\right) dB_t \tag{64}$$

B.3 Long-Term Asset Exposure

Markets are complete, thus the long-term asset is redundant. Zero-coupon Treasury bonds are in zero net supply. Long-term assets are in finite net supply. A portfolio of

[22] In excess of the short-rate, plus a dealer-specific spread on bond with maturity τ.
[23] This follows from the linearity of covariance.

$\theta e^{-\theta\tau}$ bonds of each maturity τ replicates a unit position in the long-term assets. Then, one can define the "*net exposure*" to the long-term assets, x_t^i. The relation between bond holdings and net exposure is then:

$$x_t^i = \frac{\alpha_t^{(\tau),i}}{\theta e^{-\theta\tau}} \tag{65}$$

That is, the net exposure denotes type i's exposure to the long-term assets' payoff at time τ by holding $\alpha_t^{(\tau),i}$ share of bonds.

Then, the optimality condition becomes:

$$\mathbb{E}_t^i\left(\frac{dP_t^{(\tau)}}{P_t^{(\tau)}}\right) - r_t = x_t^i \int_0^\infty \theta e^{-\theta u} Cov_t^i\left(\frac{dP_t^{(\tau)}}{P_t^{(\tau)}}, \frac{dP_t^{(u)}}{P_t^{(u)}}\right) du \tag{66}$$

and the wealth evolution is:

$$dw_t^i = w_t^i(r_t - \varsigma)dt + x_t^i w_t^i \int_0^\infty \theta e^{-\theta\tau} \left(\frac{dP_t^{(\tau)}}{P_t^{(\tau)}} - r_t dt\right) d\tau \tag{67}$$

Recall the conjectured price process:

$$\frac{dP_t^{(\tau)}}{P_t^{(\tau)}} = \ell_t(\tau)dt + \zeta_{r,t}^\tau dB_t \tag{68}$$

By Girsanov Theorem, the subjective price process is then:

$$\frac{dP_t^{(\tau)}}{P_t^{(\tau)}} = \widehat{\ell}_t^i(\tau)dt + \zeta_{r,t}^\tau d\widetilde{B}_t^i \tag{69}$$

where

$$d\widetilde{B}_t^i = dB_t - \epsilon_t^i \frac{\kappa_r}{\sigma_r} dt \tag{70}$$

and

$$\begin{aligned} \mathbb{E}_t^i\left(\frac{dP_t^{(\tau)}}{P_t^{(\tau)}}\right) &= \mathbb{E}_t\left(\frac{dP_t^{(\tau)}}{P_t^{(\tau)}}\right) + \left[\frac{dP_t^{(\tau)}}{P_t^{(\tau)}}, \frac{d\xi_t^i}{\xi_t^i}\right]_t \\ &= \ell_t(\tau)dt + \zeta_{r,t}^\tau \epsilon_t^i \frac{\kappa_r}{\sigma_r} dt \end{aligned} \tag{71}$$

and

$$Cov_t^i\left(\frac{dP_t^{(\tau)}}{P_t^{(\tau)}}, \frac{dP_t^{(u)}}{P_t^{(u)}}\right) = Cov_t\left(\frac{dP_t^{(\tau)}}{P_t^{(\tau)}}, \frac{dP_t^{(u)}}{P_t^{(u)}}\right) \tag{72}$$

Then, one can rewrite the subjective optimality condition of dealer type i under the objective measure as:

$$\ell_t(\tau) + \zeta_{r,t}^{\tau}\epsilon_t^i\frac{\kappa_r}{\sigma_r} - r_t = x_t^i\int_0^{\infty}\theta e^{-\theta u}\zeta_{r,t}^{\tau}\zeta_{r,t}^{u}du \tag{73}$$

Now, rearranging this expression yields:

$$\ell_t(\tau) - r_t = x_t^i\int_0^{\infty}\theta e^{-\theta u}\zeta_{r,t}^{\tau}\zeta_{r,t}^{u}du - \zeta_{r,t}^{\tau}\epsilon_t^i\frac{\kappa_r}{\sigma_r} \tag{74}$$

This equation has to hold for each dealer. The left-hand side of the equality contains no dealer-specific terms. Isolating exposures, I find:

$$x_t^i = \frac{\ell_t(\tau) - r_t}{\int_0^{\infty}\theta e^{-\theta u}\zeta_{r,t}^{\tau}\zeta_{r,t}^{u}du} + \frac{\zeta_{r,t}^{\tau}\frac{\kappa_r}{\sigma_r}}{\int_0^{\infty}\theta e^{-\theta u}\zeta_{r,t}^{\tau}\zeta_{r,t}^{u}du}\epsilon_t^i \tag{75}$$

Finally, define

$$Z_{r,t}^i = x_t^i\int_0^{\infty}\theta e^{-\theta\tau}\zeta_{r,t}^{\tau}d\tau, \tag{76}$$

Recall that by market completeness a portfolio of $\theta e^{-\theta\tau}$ bonds replicate a unit position in the long-term assets. Then, as $\tau\to\infty$, the payoff of the long-term assets goes to zero: $\theta e^{-\theta\tau}\to 0$. Thus, the price of a bond of maturity τ as $\tau\to\infty$ must *almost surely* be zero over the entire state-space. Hence:

$$\lim_{\tau\to\infty}\zeta_{r,t}^{\tau} = 0 \tag{77}$$

Further, by the boundary condition for bond prices, I must have that the return on a bond of maturity $\tau = 0$ must be equal to the short-rate:

$$\lim_{\tau\to 0}\ell_t(\tau) = -\kappa_r(r_t - \mu_t), \quad \lim_{\tau\to 0}\zeta_{r,t}^{\tau} = \sigma_r \tag{78}$$

These conditions guarantee that $Z_{r,t}^i$ is well-defined.

B.4 Zero-Coupon Bond Prices

Proof of Proposition 1. The proof follows Xiong and Yan (2010). For any random variable X_τ such that $E^i[X_\tau] < \infty$ define $Y_\tau = \frac{w_\tau^i}{w_t^i}X_\tau$. Since markets are complete, each dealer is a marginal agent. Then a claim to the cash-flow Y_τ is priced as:

$$E_t^i\left[e^{-\rho(\tau-t)}\frac{u'(c_\tau^i)}{u'(c_t^i)}Y_\tau\right]=e^{-\rho(\tau-t)}E_t^i\left[\frac{c_t^i}{c_\tau^i}Y_\tau\right]=e^{-\rho(\tau-t)}E_t^i\left[\frac{w_t^i}{w_\tau^i}Y_\tau\right]=e^{-\rho(\tau-t)}E_t^i\left[X_\tau\right] \quad (79)$$

Define

$$\omega_t^{i,j}=\frac{w_t^i}{w_t^j} \quad (80)$$

Then the same security is priced by dealer j as:

$$E_t^j\left[e^{-\rho(\tau-t)}\frac{u'(c_\tau^j)}{u'(c_t^j)}Y_\tau\right]=e^{-\rho(\tau-t)}E_t^j\left[\frac{c_t^j}{c_\tau^j}Y_\tau\right]=e^{-\rho(\tau-t)}E_t^j\left[\frac{w_t^j}{w_\tau^j}Y_\tau\right]=e^{-\rho(\tau-t)}E_t^j\left[\frac{\omega_\tau^{i,j}}{\omega_t^{i,j}}X_\tau\right] \quad (81)$$

Then the wealth ratio acts as a Radon-Nikodym derivative. In equilibrium it must be the case that:

$$E_t^i\left[X_\tau\right]=E_t^j\left[\frac{\omega_\tau^{i,j}}{\omega_t^{i,j}}X_\tau\right] \quad (82)$$

The price of an asset paying off a single dividend at time τ is:

$$\begin{aligned}
P_{X,t}^{(\tau)}&=E_t^j\left[e^{-\rho(\tau-t)}\frac{w_t^j}{w_\tau^j}X_\tau\right]=e^{-\rho(\tau-t)}E_t^j\left[\frac{w_t^j}{w_\tau^j}\frac{W_\tau}{W_t}\frac{W_t}{W_\tau}X_\tau\right]=\frac{w_t^j}{W_t}e^{-\rho(\tau-t)}E_t^j\left[\frac{W_\tau}{w_\tau^j}\frac{W_t}{W_\tau}X_\tau\right]\\
&=\frac{w_t^j}{W_t}e^{-\rho(\tau-t)}E_t^j\left[\frac{\sum_{i=1}^N w_\tau^i}{w_\tau^j}\frac{W_t}{W_\tau}X_\tau\right]=\frac{w_t^j}{W_t}e^{-\rho(\tau-t)}E_t^j\left[\sum_{i=1}^N\frac{w_\tau^i}{w_\tau^j}\frac{W_t}{W_\tau}X_\tau\right]=\frac{w_t^j}{W_t}e^{-\rho(\tau-t)}E_t^j\left[\sum_{i=1}^N\omega_\tau^{i,j}\frac{W_t}{W_\tau}X_\tau\right]\\
&=\frac{w_t^j}{W_t}e^{-\rho(\tau-t)}E_t^j\left[\frac{W_t}{W_\tau}X_\tau\right]+\frac{w_t^j}{W_t}e^{-\rho(\tau-t)}\left(\sum_{i\neq j}E_t^j\left[\omega_\tau^{i,j}\frac{W_t}{W_\tau}X_\tau\right]\right)\\
&=\frac{w_t^j}{W_t}e^{-\rho(\tau-t)}E_t^j\left[\frac{W_t}{W_\tau}X_\tau\right]+\frac{w_t^j}{W_t}e^{-\rho(\tau-t)}\left(\sum_{i\neq j}\omega_t^{i,j}E_t^j\left[\frac{\omega_\tau^{i,j}}{\omega_t^{i,j}}\frac{W_t}{W_\tau}X_\tau\right]\right)\\
&=\frac{w_t^j}{W_t}e^{-\rho(\tau-t)}E_t^j\left[\frac{W_t}{W_\tau}X_\tau\right]+\sum_{i\neq j}\frac{w_t^i}{W_t}e^{-\rho(\tau-t)}E_t^i\left[\frac{W_t}{W_\tau}X_\tau\right]=\sum_{i=1}^N\frac{w_t^i}{W_t}E_t^i\left[e^{-\rho(\tau-t)}\frac{W_t}{W_\tau}X_\tau\right]\\
&=\sum_{i=1}^N\omega_t^iE_t^i\left[e^{-\rho(\tau-t)}\frac{W_t}{W_\tau}X_\tau\right]
\end{aligned} \quad (83)$$

Now, notice that the term

$$E_t^i\left[e^{-\rho(\tau-t)}\frac{W_t}{W_\tau}X_\tau\right] \quad (84)$$

would correspond to the price of the security if a dealer of type i owned all wealth in this economy. This would be the case if the economy was populated by a single dealer type. Thus, the price of any security with a single dividend at time τ in the heterogenous dealer economy is given by:

$$P_{X,t}^{(\tau)} = \sum_{i=1}^{N} \frac{w_t^i}{W_t} P_{X,t}^{(\tau),i} \tag{85}$$

For zero-coupon bonds, $X_t = 1$.

□

***Proof of Proposition* 2.** Suppose $w_t^i = W_t$, that is, only type i dealers exist. Then I have $P_t^{(\tau),i} = P_t^{(\tau)}$. The optimality condition must still be satisfied. Conjecture:

$$P_t^{(\tau),i} = \exp\Big(a(\tau) + b_r(\tau)r_t + C_\mu(\tau)\mu_t^i + C_\epsilon^i(\tau)\epsilon_t^i\Big) \tag{86}$$

By Itô's lemma:

$$\begin{aligned}\frac{dP_t^{(\tau),i}}{P_t^{(\tau),i}} &= \underbrace{\frac{1}{P_t^{(\tau),i}}\Big[P_{r,t}^{(\tau),i}\kappa_r(\mu_t - r_t) + P_{\mu,t}^{(\tau),i}\kappa_\mu(\bar{\mu} - \mu_t) - P_{\epsilon^i,t}^{(\tau),i}\kappa_\epsilon\epsilon_t^i - P_{\tau,t}^{(\tau),i} + \frac{1}{2}P_{rr,t}^{(\tau),i}\sigma_r^2 + \frac{1}{2}P_{\mu\mu,t}^{(\tau),i}\sigma_\mu^2 + \frac{1}{2}P_{\epsilon^i\epsilon^i,t}^{(\tau),i}\sigma_\epsilon^{i2}\frac{\kappa_r^2}{\sigma_r^2} + P_{r\mu,t}^{(\tau),i}\sigma_r\sigma_\mu + P_{r\epsilon^i,t}^{(\tau),i}\sigma_r\sigma_\epsilon^i\Big]}_{\ell^i(\tau)} dt \\ &+ \underbrace{\frac{1}{P_t^{(\tau),i}}\Big[P_{r,t}^{(\tau),i}\sigma_r + P_{\mu,t}^{(\tau),i}\sigma_\mu + P_{\epsilon^i,t}^{(\tau),i}\sigma_\epsilon^i\Big]}_{\zeta_{r,t}^{\tau,i}} dB_t\end{aligned} \tag{87}$$

The optimality condition implies:

$$\begin{aligned}&C_r(\tau)\kappa_r(\mu_t + \epsilon_t^i - r_t) + C_\mu(\tau)\kappa_\mu(\bar{\mu} - \mu_t - \epsilon_t^i) - C_\epsilon^i(\tau)\kappa_\epsilon\epsilon_t^i \\ &- a'(\tau) - C_r'(\tau)r_t - C_\mu'(\tau)(\mu_t + \epsilon_t^i) - C_\epsilon^{i\prime}(\tau)\epsilon_t^i + \frac{1}{2}C_r(\tau)^2\sigma_r^2 + \frac{1}{2}C_\mu(\tau)^2\sigma_\mu^2 + \frac{1}{2}C_\epsilon^i(\tau)^2\sigma_\epsilon^{i2} \\ &+ C_r(\tau)C_\mu(\tau)\sigma_r\sigma_\mu + C_r(\tau)C_\epsilon^i(\tau)\sigma_r\sigma_\epsilon^i - r_t = x_t^i C_\epsilon^i(\tau)\sigma_\epsilon^i\int_0^\infty \theta e^{-\theta u}C_\epsilon^i(u)\sigma_\epsilon^i du \\ &+ x_t^i\big[C_r(\tau)\sigma_r + C_\mu(\tau)\sigma_\mu + C_\epsilon^i(\tau)\sigma_\epsilon^i\big]\int_0^\infty \theta e^{-\theta u}\big[C_r(u)\sigma_r + C_\mu(u)\sigma_\mu + C_\epsilon^i(u)\sigma_\epsilon^i\big] \\ &- \big[C_r(\tau)\sigma_r + C_\mu(\tau)\sigma_\mu + C_\epsilon^i(\tau)\sigma_\epsilon^i\big]\frac{\kappa_r}{\sigma_r}\epsilon_t^i\end{aligned} \tag{88}$$

Coefficients on the two sides of the equality must be identical. Collecting the coefficients on r_t yields:

$$C_r'(\tau) + \kappa_r C_r(\tau) + 1 = 0 \tag{89}$$

This is a first-order linear ODE with boundary condition $b_r(0) = 0$, so the solution is:

$$C_r(\tau) = -\frac{(1 - e^{-\kappa_r \tau})}{\kappa_r} \tag{90}$$

Similarly, collecting the coefficients of μ_t:

$$C_\mu'(\tau) + \kappa_\mu C_\mu(\tau) - \kappa_r C_r(\tau) = 0 \implies C_\mu'(\tau) + \kappa_\mu C_\mu(\tau) + 1 - e^{-\kappa_r \tau} = 0 \tag{91}$$

Imposing the boundary condition $C_\mu(0) = 0$ once again, I find:

$$C_\mu(\tau) = -\frac{1}{\kappa_\mu}\left(1 - \frac{\kappa_r e^{-\kappa_\mu \tau} - \kappa_\mu e^{-\kappa_r \tau}}{\kappa_r - \kappa_\mu}\right) \tag{92}$$

Now, considering the terms multiplying ϵ_t^i I find:

$$2\kappa_r C_r(\tau) - \left(\kappa_\mu - \frac{\kappa_r}{\sigma_r}\sigma_\mu\right) C_\mu(\tau) - \left(\kappa_\epsilon - \frac{\kappa_r}{\sigma_r}\sigma_\epsilon^i\right) C_\epsilon^i(\tau) - C_\mu'(\tau) - C_\epsilon^{i\prime}(\tau) = 0 \tag{93}$$

Plugging in $C_r(\tau)$ and $C_\mu(\tau)$:

$$\begin{aligned} C_\epsilon^{i\prime}(\tau) &= -2(1 - e^{-\kappa_r \tau}) + \left(1 - \frac{\kappa_r \sigma_\mu}{\kappa_\mu \sigma_r}\right)\left(1 - \frac{\kappa_r e^{-\kappa_\mu \tau} - \kappa_\mu e^{-\kappa_r \tau}}{\kappa_r - \kappa_\mu}\right) \\ &+ \frac{\kappa_r}{\kappa_r - \kappa_\mu}\left(e^{-\kappa_\mu \tau} - e^{-\kappa_r \tau}\right) - \left(\kappa_\epsilon - \frac{\kappa_r}{\sigma_r}\sigma_\epsilon^i\right) C_\epsilon^i(\tau) \end{aligned} \tag{94}$$

The solution is:

$$C_\epsilon^i(\tau) = \frac{\kappa_\mu \sigma_r + \kappa_r \sigma_\mu}{\kappa_\mu(\kappa_r \sigma_\epsilon^i - \kappa_\epsilon \sigma_r)}\left(1 - e^{-\left(\kappa_\epsilon - \kappa_r \frac{\sigma_\epsilon^i}{\sigma_r}\right)\tau}\right) - \frac{\kappa_\mu \sigma_r + \kappa_r \sigma_\mu}{\kappa_\mu(\kappa_r \sigma_\epsilon^i - \kappa_\epsilon \sigma_r + \kappa_\mu \sigma_r)}\left(e^{-\kappa_r \tau} - e^{-\left(\kappa_\epsilon - \kappa_r \frac{\sigma_\epsilon^i}{\sigma_r}\right)\tau}\right) \tag{95}$$

Now, note the following:

$$B_{\theta,r} = \int_0^\infty \theta e^{-\theta \tau} C_r(\tau) d\tau = -\frac{1}{\kappa_r + \theta} \tag{96}$$

$$B_{\theta,\mu} = \int_0^\infty \theta e^{-\theta\tau} C_\mu(\tau) d\tau = -\frac{\kappa_r}{(\kappa_r+\theta)(\kappa_\mu+\theta)} < 0 \tag{97}$$

$$B^i_{\theta,\epsilon} = \int_0^\infty \theta e^{-\theta\tau} C^i_\epsilon(\tau) d\tau = \frac{\kappa_r(\kappa_r\,\sigma_\mu + \kappa_\mu\,\sigma_r + \sigma_r\,\theta)}{(\kappa_r+\theta)(\kappa_\mu+\theta)(\kappa_r\,\sigma^i_\epsilon - \kappa_\epsilon\,\sigma_r - \sigma_r\,\theta)} < 0 \tag{98}$$

Recall that:

$$\zeta^{(\tau),i}_{r,t} = C_r(\tau)\sigma_r + C_\mu(\tau)\sigma_\mu + C^i_\epsilon(\tau)\sigma^i_\epsilon \tag{99}$$

Hence by the linearity of integral:

$$Z^i_{\theta,r} = \int_0^\infty \theta e^{-\theta\tau} \zeta^{(\tau),i}_{r,t} d\tau = \frac{\sigma_r(\kappa_\epsilon+\theta)(\kappa_r\,\sigma_\mu + \kappa_\mu\,\sigma_r + \sigma_r\,\theta)}{(\kappa_r+\theta)(\kappa_\mu+\theta)(\kappa_r\,\sigma^i_\epsilon - \kappa_\epsilon\,\sigma_r - \sigma_r\,\theta)} < 0 \tag{100}$$

Finally, collecting the remaining terms together I obtain the following ODE:

$$a'(\tau) = C_\mu(\tau)\kappa_\mu\bar{\mu} + \frac{1}{2}C_r(\tau)^2\sigma_r^2 + \frac{1}{2}C_\mu(\tau)^2\sigma_\mu^2 + \frac{1}{2}C^i_\epsilon(\tau)^2\sigma^{i\,2}_\epsilon + C_r(\tau)C_\mu(\tau)\sigma_r\sigma_\mu + C_r(\tau)C^i_\epsilon(\tau)\sigma_r\sigma^i_\epsilon$$
$$- x^i_t\zeta^i_{r,t}(\tau)Z^i_{\theta,r} \tag{101}$$

Note that $a'(\tau)$ enters the drift of bond returns with a negative sign. Then I can back out the relation between long-term asset exposure x^i_t and expected bond returns as:

$$x^i_t\zeta^i_{r,t}(\tau)Z^i_{\theta,r} \tag{102}$$

The sign of this expression depends on the sign of x^i_t. If x^i_t is positive, indicating positive exposure to long-term assets, then high exposure is associated with higher expected bond returns.

Since $C_r(0) = C_\mu(0) = C^i_\epsilon(0) = 0$, I have $a'(0) = 0$. The boundary condition implies $a(0) = 0$. I further have:

$$\int_0^\infty \theta e^{-\theta\tau} C_r(\tau)^2 d\tau = \frac{2}{(\kappa_r+\theta)(2\,\kappa_r+\theta)} \tag{103}$$

$$\int_0^\infty \theta e^{-\theta\tau} C_\mu(\tau)^2 d\tau = \frac{2\,{\kappa_r}^2(2\,\kappa_r + 2\,\kappa_\mu + 3\,\theta)}{(\kappa_r+\theta)(\kappa_\mu+\theta)(2\,\kappa_r+\theta)(2\,\kappa_\mu+\theta)(\kappa_r+\kappa_\mu+\theta)} \tag{104}$$

$$\int_0^\infty \theta e^{-\theta\tau} C_\mu(\tau)C_r(\tau) d\tau = \frac{\kappa_r(2\,\kappa_r + 2\,\kappa_\mu + 3\,\theta)}{(\kappa_\mu+\theta)(2\,{\kappa_r}^2 + 3\,\kappa_r\,\theta + \theta^2)(\kappa_r+\kappa_\mu+\theta)} \tag{105}$$

$$\int_0^\infty \theta e^{-\theta\tau} C_\epsilon^i(\tau) C_r(\tau) d\tau = \Bigg[\kappa_r \Bigg(3\,{\sigma_r}^2\,\theta^3 + 2\,{\kappa_r}^3\,\sigma_r\,\sigma_\mu - 2\,{\kappa_r}^3\,\sigma_\mu\,\sigma_\epsilon^i + 2\,\kappa_r\,{\kappa_\mu}^2\,{\sigma_r}^2 + 2\,{\kappa_r}^2\,\kappa_\mu\,{\sigma_r}^2 + 2\,{\kappa_\mu}^2\,\kappa_\epsilon\,{\sigma_r}^2$$

$$+ 5\,\kappa_r\,{\sigma_r}^2\,\theta^2 + 2\,{\kappa_r}^2\,{\sigma_r}^2\,\theta + 6\,\kappa_\mu\,{\sigma_r}^2\,\theta^2 + 3\,{\kappa_\mu}^2\,{\sigma_r}^2\,\theta + 2\,\kappa_\epsilon\,{\sigma_r}^2\,\theta^2 + 2\,\kappa_r\,\kappa_\mu\,\kappa_\epsilon\,{\sigma_r}^2 + 2\,{\kappa_r}^2\,\kappa_\mu\,\sigma_r\,\sigma_\mu - 2\,\kappa_r\,{\kappa_\mu}^2\,\sigma_r\,\sigma_\epsilon^i$$

$$- 2\,{\kappa_r}^2\,\kappa_\mu\,\sigma_r\,\sigma_\epsilon^i + 2\,{\kappa_r}^2\,\kappa_\epsilon\,\sigma_r\,\sigma_\mu - 2\,{\kappa_r}^2\,\kappa_\mu\,\sigma_\mu\,\sigma_\epsilon^i + 7\,\kappa_r\,\kappa_\mu\,{\sigma_r}^2\,\theta + 2\,\kappa_r\,\kappa_\epsilon\,{\sigma_r}^2\,\theta + 4\,\kappa_\mu\,\kappa_\epsilon\,{\sigma_r}^2\,\theta + 4\,\kappa_r\,\sigma_r\,\sigma_\mu\,\theta^2$$

$$+ 5\,{\kappa_r}^2\,\sigma_r\,\sigma_\mu\,\theta - 2\,\kappa_r\,\sigma_r\,\sigma_\epsilon^i\,\theta^2 - 2\,{\kappa_r}^2\,\sigma_r\,\sigma_{\epsilon,r}^i\,\theta - 3\,{\kappa_r}^2\,\sigma_\mu\,\sigma_\epsilon^i\,\theta + 2\,\kappa_r\,\kappa_\mu\,\kappa_\epsilon\,\sigma_r\,\sigma_\mu$$

$$+ 3\,\kappa_r\,\kappa_\mu\,\sigma_r\,\sigma_\mu\,\theta - 4\,\kappa_r\,\kappa_\mu\,\sigma_r\,\sigma_\epsilon^i\,\theta + 3\,\kappa_r\,\kappa_\epsilon\,\sigma_r\,\sigma_\mu\,\theta\Bigg)\Bigg] \times \Bigg[(\kappa_\mu + \theta)\left(2\,{\kappa_r}^2 + 3\,\kappa_r\,\theta + \theta^2\right)(\kappa_r + \kappa_\mu + \theta)$$

$$\left(-{\kappa_r}^2\,\sigma_r\,\sigma_\epsilon^i + {\kappa_r}^2\,{\sigma_{\epsilon,r}}^{i2} + \kappa_r\,\kappa_\epsilon\,{\sigma_r}^2 - 2\,\kappa_r\,\kappa_\epsilon\,\sigma_r\,\sigma_\epsilon^i + \kappa_r\,{\sigma_r}^2\,\theta - 2\,\kappa_r\,\sigma_r\,\sigma_\epsilon^i\,\theta + {\kappa_\epsilon}^2\,{\sigma_r}^2 + 2\,\kappa_\epsilon\,{\sigma_r}^2\,\theta + {\sigma_r}^2\,\theta^2\right)\Bigg]^{-1} \tag{106}$$

The expression for the integral involving $C_\epsilon^i(\tau)^2$ is highly involved, yet a closed form expression is available. These ensure that the drift of the bond holdings portfolio is well-defined. Moreover:

$$\lim_{\tau\to\infty} C_r(\tau) = -\frac{1}{\kappa_r} \tag{107}$$

$$\lim_{\tau\to\infty} C_\mu(\tau) = -\frac{1}{\kappa_\mu} \tag{108}$$

$$\lim_{\tau\to\infty} C_\epsilon^i(\tau) = \frac{\kappa_\mu\sigma_r + \kappa_r\sigma_\mu}{\kappa_\mu(\kappa_r\sigma_\epsilon^i - \kappa_\epsilon\sigma_r)} \tag{109}$$

Then by the properties of the exponential function:

$$\lim_{\tau\to\infty} a'(\tau) = \underbrace{-\bar{\mu}}_{<0} + \underbrace{\frac{1}{2}\frac{\sigma_r^2}{\kappa_r^2} + \frac{1}{2}\frac{\sigma_\mu^2}{\kappa_\mu^2} + \frac{1}{2}\left(\frac{\kappa_\mu\sigma_r + \kappa_r\sigma_\mu}{\kappa_\mu(\kappa_r\sigma_\epsilon^i - \kappa_\epsilon\sigma_r)}\right)^2 \sigma_\epsilon^{i2} + \frac{\sigma_r\sigma_\mu}{\kappa_r\kappa_\mu} - \frac{\kappa_\mu\sigma_r + \kappa_r\sigma_\mu}{\kappa_r\kappa_\mu(\kappa_r\sigma_\epsilon^i - \kappa_\epsilon\sigma_r)}\sigma_r\sigma_\epsilon^i}_{>0}$$
$$\underbrace{-x_t^i\Big[-\frac{\sigma_r}{\kappa_r} - \frac{\sigma_\mu}{\kappa_\mu} + \sigma_\epsilon^i\frac{\kappa_\mu\sigma_r + \kappa_r\sigma_\mu}{\kappa_\mu(\kappa_r\sigma_\epsilon^i - \kappa_\epsilon\sigma_r)}\Big]Z_{\theta,r}^i}_{-sign(x_t^i)} \tag{110}$$

In any reasonable calibration of the model parameters, I have that $\lim_{\tau\to\infty} a'(\tau) > 0$. Further, consider the following:

$$a(\tau) = \int \Bigg(C_\mu(\tau)\kappa_\mu\bar{\mu} + \frac{1}{2}C_r(\tau)^2\sigma_r^2 + \frac{1}{2}C_\mu(\tau)^2\sigma_\mu^2 + \frac{1}{2}C_\epsilon^i(\tau)^2\sigma_\epsilon^{i2} + C_r(\tau)C_\mu(\tau)\sigma_r\sigma_\mu + C_r(\tau)C_\epsilon^i(\tau)\sigma_r\sigma_\epsilon^i$$
$$- x_t^i\zeta_{r,t}^i(\tau)Z_{\theta,r}^i\Bigg)d\tau \tag{111}$$

I established the finiteness of the limits for the functions $b_r(\tau), C_\mu(\tau), C_\epsilon^(\tau)$. By the properties of the exponential function, it is straightforward to establish the existence and finiteness of limits for any positive power of these functions. Then, integrating these functions and the constant terms one can express $a(\tau)$ as:

$$a(\tau) = C\tau + E(\tau) \tag{112}$$

where $E(\tau)$ is the collection of the exponential terms, and $C \in \mathbb{R}$ is some constant. Then by the properties of the exponential function, I have:

$$\lim_{\tau\to\infty} \frac{-a(\tau)}{\tau} = -C \tag{113}$$

ensuring that bond yields are well-defined. □

B.5 Representative Agent Risk Premium

Proof of Proposition 5. Recall that the risk premium is:

$$\begin{aligned}\ell_t^i(\tau) &= C_r(\tau)\kappa_r(\mu_t + \epsilon_t^i - r_t) + C_\mu(\tau)\kappa_\mu(\bar{\mu} - \mu_t - \epsilon_t^i) - C_\epsilon^i(\tau)\kappa_\epsilon\epsilon_t^i \\ &- a'(\tau) - C_r'(\tau)r_t - C_\mu'(\tau)(\mu_t + \epsilon_t^i) - C_\epsilon^{i\prime}(\tau)\epsilon_t^i \\ &+ \frac{1}{2}C_r(\tau)^2\sigma_r^2 + \frac{1}{2}C_\mu(\tau)^2\sigma_\mu^2 + \frac{1}{2}C_\epsilon^i(\tau)^2\sigma_\epsilon^{i2}\end{aligned} \tag{114}$$

In the representative agent case, I have:

$$\ell_t^i(\tau) - r_t = -C_r(\tau)C_\mu(\tau)\sigma_r\sigma_\mu - x_t^i\Big(C_r(\tau)\sigma_r + C_\mu(\tau)\sigma_\mu\Big)\left(\frac{(\kappa_r\,\sigma_\mu + \kappa_\mu\,\sigma_r + \sigma_r\,\theta)}{(\kappa_r + \theta)(\kappa_\mu + \theta)}\right) \tag{115}$$

In the heterogeneous agent case:

$$\begin{aligned}\ell_t^i(\tau) - r_t = &-\Big(\sigma_\mu C_\mu(\tau) + \sigma_\epsilon^i C_\epsilon^i(\tau)\Big)\frac{\kappa_r}{\sigma_r}\epsilon_t^i - C_r(\tau)C_\mu(\tau)\sigma_r\sigma_\mu - C_r(\tau)C_\epsilon^i(\tau)\sigma_r\sigma_\epsilon^i \\ &- x_t^i\Big(C_r(\tau)\sigma_r + C_\mu(\tau)\sigma_\mu + C_\epsilon^i(\tau)\sigma_\epsilon^i\Big)\left(\frac{(\kappa_r\,\sigma_\mu + \kappa_\mu\,\sigma_r + \sigma_r\,\theta)}{(\kappa_r + \theta)\,(\kappa_\mu + \theta)}\right)\left(\frac{\sigma_r\,(\kappa_\epsilon + \theta)}{(\kappa_r\,\sigma_\epsilon^i - \kappa_\epsilon\,\sigma_r - \sigma_r\,\theta)}\right)\end{aligned} \tag{116}$$

I can rewrite this as:

$$\begin{aligned} rp_t^i(\tau) = rp_t^{rep} \quad & \underbrace{-\left(\sigma_\mu C_\mu(\tau) + \sigma_\epsilon^i C_\epsilon^i(\tau)\right)\frac{\kappa_r}{\sigma_r}\epsilon_t^i}_{\geqslant 0} \quad \underbrace{- \quad C_r(\tau) C_\epsilon^i(\tau)\sigma_r\sigma_\epsilon^i}_{sign(\sigma_r\sigma_\epsilon^i)} \\ & \underbrace{+x_t^i C_\epsilon^i(\tau)\sigma_\epsilon^i\left(\frac{(\kappa_r\,\sigma_\mu + \kappa_\mu\,\sigma_r + \sigma_r\,\theta)}{(\kappa_r+\theta)\,(\kappa_\mu+\theta)}\right)\left(\frac{\sigma_r\,(\kappa_\epsilon+\theta)}{(\kappa_r\,\sigma_\epsilon^i - \kappa_\epsilon\,\sigma_r - \sigma_r\,\theta)}\right)}_{\geqslant 0} \end{aligned} \tag{117}$$

□

B.6 Wealth Evolution

In this section, I work with aggregate wealth W_t. Since $Y_t = \rho W_t$ and $c_t^i = \rho w_t^i$, the results of this section extend immediately to aggregate and individual consumption processes.

Denote the process for aggregate wealth as:

$$\frac{dW_t}{W_t} = (r_t + \mu_W)dt + \sigma_W dB_t \tag{118}$$

Following the martingale approach of Cox and Huang (1989), the problem of dealer i can be rewritten as the static optimization problem:

$$\begin{gathered} \max_{c_i} \quad E_0^i\left[\int_0^\infty e^{-\rho t} u\left(c_t^i\right) dt\right] \\ s.t. \\ E_0^i\left[\int_0^\infty \pi_t^i c_t^i dt\right] \leqslant s_0^i E_0^i\left[\int_0^\infty \pi_t^i \rho W_t dt\right] \end{gathered} \tag{119}$$

where π_t^i denotes the (unique) state-price density of dealer i, and s_0^i denotes the initial wealth share. Under i's beliefs, the problem of dealer j becomes:

$$\begin{gathered} \max_{c_i} \quad E_0^i\left[\int_0^\infty \xi_t^{j,i} e^{-\rho t} u\left(c_t^j\right) dt\right] \\ s.t. \\ E_0^i\left[\int_0^\infty \pi_t^i c_t^j dt\right] \leqslant s_0^j E_0^i\left[\int_0^\infty \pi_t^i \rho W_t dt\right] \end{gathered} \tag{120}$$

where $\xi_t^{j,i}$ denotes the Radon-Nikodym derivative of dealer j and i's beliefs. The optimality conditions for the two dealers are respectively:

$$w_t^i = e^{-\rho t}\left(\rho\lambda^i\pi_t^i\right)^{-1} \tag{121}$$

and

$$w_t^j = e^{-\rho t}\xi_t^{j,i}\left(\rho\lambda^j\pi_t^i\right)^{-1} \tag{122}$$

where λ^i is the Lagrange multiplier on the static budget constraint. Plugging in the optimality condition into the static budget constraint yields a condition for λ^i:

$$\rho\lambda^i s_0^i E_0^i\left[\int_0^\infty \pi_t^i\rho W_t dt\right] = 1 \tag{123}$$

Now, using the optimality condition and the aggregate resource constraint I obtain a single equation in one unknown π_t^i:

$$\underbrace{e^{-\rho t}\left(\rho\lambda^i\pi_t^i\right)^{-1}}_{w_t^i} + \sum_{j\neq i}\underbrace{e^{-\rho t}\xi_t^{j,i}\left(\rho\lambda^j\pi_t^i\right)^{-1}}_{w_t^j} = W_t \tag{124}$$

Solving for π_t^i and plugging in the optimality condition then yields:

$$\omega_t^i = \frac{w_t^i}{W_t} = \left(1 + \sum_{j\neq i}\xi_t^{j,i}\frac{\lambda^i}{\lambda^j}\right)^{-1} \tag{125}$$

I further have:

$$\pi_t^i = e^{-\rho t}(\rho\lambda^i w_t^i)^{-1} \tag{126}$$

and thus

$$\rho\pi_t^i W_t = e^{-\rho t}W_t\frac{1}{\lambda^i}\frac{1}{w_t^i} = e^{-\rho t}\frac{1}{\lambda^i} + e^{-\rho t}\sum_{j\neq i}\xi_t^{j,i}\frac{1}{\lambda^j} \tag{127}$$

Then:

$$\begin{aligned}
E_0^i\left[\int_0^\infty \pi_t^i\rho W_t dt\right] &= \frac{1}{\lambda^i}\int_0^\infty e^{-\rho t}dt + E_0^i\left[\int_0^\infty e^{-\rho t}\sum_{j\neq i}\xi_t^{j,i}\frac{1}{\lambda^j}dt\right] \\
&= \frac{1}{\rho\lambda_i} + \int_0^\infty e^{-\rho t}\sum_{j\neq i}E_0^i\left[\xi_t^{j,i}\right]\frac{1}{\lambda^j}dt \\
&= \frac{1}{\rho\lambda_i} + \int_0^\infty e^{-\rho t}\sum_{j\neq i}\frac{\lambda_i}{\lambda_j^2}dt \\
&= \frac{1}{\rho\lambda_i} + \frac{1}{\rho}\sum_{j\neq i}\frac{\lambda_i}{\lambda_j^2}
\end{aligned} \tag{128}$$

Then:

$$\begin{aligned}\rho\lambda^i s_0^i E_0^i\left[\int_0^\infty \pi_t^i \rho W_t dt\right] &= s_0^i\left(1+\sum_{j\neq i}\frac{\lambda_i^2}{\lambda_j^2}\right) = 1 \\ \implies \lambda_i &= \sqrt{\left(\frac{1}{s_0^i}-1\right)\left(\sum_{j\neq i}\frac{1}{\lambda_j^2}\right)}\end{aligned} \tag{129}$$

This equation has to hold for each dealer, giving us I non-linear equations in I unknowns. I solve this system numerically. Importantly, this equation highlights that the Lagrange multipliers on the static budget constraint only depend on each other and the initial wealth share.

Now, recall that the relative wealth ratios $\omega_t^{j,i} = \frac{w_t^j}{w_t^i}$ serve as Radon-Nikodym derivatives with logarithmic preferences. The Radon-Nikodym derivative between two dealers' beliefs must be unique, which yields:

$$\frac{d\xi_t^{j,i}}{\xi_t^{j,i}} = \frac{d\omega_t^{j,i}}{\omega_t^{j,i}} = (\epsilon_t^j - \epsilon_t^i)\frac{\kappa_r}{\sigma_r}dB_t^i \tag{130}$$

Thus, knowing w_t^i and $\xi_t^{j,i}$ fully characterizes the wealth distribution at any time t. Under the econometrician's measure:

$$\frac{d\xi_t^{j,i}}{\xi_t^{j,i}} = \frac{d\omega_t^{j,i}}{\omega_t^{j,i}} = (\epsilon_t^j - \epsilon_t^i)\frac{\kappa_r}{\sigma_r}dB_t - (\epsilon_t^j - \epsilon_t^i)\epsilon_t^i\left(\frac{\kappa_r}{\sigma_r}\right)^2 dt \tag{131}$$

I further have the following relation between relative wealth ratios and likelihood ratios:

$$\omega_t^{j,i} = \xi_t^{j,i}\frac{\lambda^i}{\lambda^j} \tag{132}$$

Applying Itô's lemma to w_t^i, I get the loadings on the Brownian risk factors as:

$$\left(-W_t\left(1+\sum_{j\neq i}\xi_t^{j,i}\frac{\lambda^i}{\lambda^j}\right)^{-2}\sum_{j\neq i}\frac{\lambda^i}{\lambda^j}\xi_t^{j,i}(\epsilon_t^j-\epsilon_t^i)\frac{\kappa_r}{\sigma_r} + \left(1+\sum_{j\neq i}\xi_t^{j,i}\frac{\lambda^i}{\lambda^j}\right)^{-1}\sigma_W W_t\right)dB_t^i \tag{133}$$

Using the definition of w_t^i and algebraic manipulation yields:

$$\left(\sigma_W - \frac{w_t^i}{W_t}\sum_{j\neq i}\frac{\lambda^i}{\lambda^j}\xi_t^{j,i}(\epsilon_t^j-\epsilon_t^i)\frac{\kappa_r}{\sigma_r}\right)w_t^i dB_t^i \tag{134}$$

Recall that wealth evolves as:

$$\frac{dw_t^i}{w_t^i} = (r_t - \rho^{-1})dt + \int_0^\infty \alpha_t^i(\tau)(\ell_t^\tau - r_t)dtd\tau + \int_0^\infty \alpha_t^i(\tau)\zeta_{r,t}^\tau d\tau dB_t \tag{135}$$

Thus, I obtain the following equalities for portfolio shares:

$$\int_0^\infty \alpha_t^i(\tau)\zeta_{r,t}^\tau d\tau = \left(\sigma_W - \frac{w_t^i}{W_t}\sum_{j\neq i}\frac{\lambda^i}{\lambda^j}\xi_t^{j,i}(\epsilon_t^j - \epsilon_t^i)\frac{\kappa_r}{\sigma_r}\right) \tag{136}$$

or equivalently

$$\int_0^\infty \alpha_t^i(\tau)\zeta_{r,t}^\tau d\tau = \left(\sigma_W - \frac{w_t^i}{W_t}\sum_{j\neq i}\omega_t^{j,i}(\epsilon_t^j - \epsilon_t^i)\frac{\kappa_r}{\sigma_r}\right) \tag{137}$$

Now, consider the pricing function

$$P_t^{(\tau)} = \sum_{i=1}^N \frac{w_t^i}{W_t}P_t^{(\tau),i} \tag{138}$$

Start by considering the dynamics of $\frac{w_t^i}{W_t}P_t^{(\tau),i} = \omega_t^i P_t^{(\tau),i}$. By Itô's lemma,

$$d\omega_t^i P_t^{(\tau),i} = \omega_t^i dP_t^{(\tau),i} + P_t^{(\tau),i}d\omega_t^i + d\omega_t^i dP_t^{(\tau),i} \tag{139}$$

The second term yields the diffusion term:

$$-P_t^{(\tau),i}\omega_t^{i2}\sum_{j\neq i}\xi_t^{j,i}\frac{\lambda^i}{\lambda^j}(\epsilon_t^j - \epsilon_t^i)dB_t \tag{140}$$

The first term yields the diffusion term:

$$\omega_t^i P_t^{(\tau),i}\zeta_{r,t}^{\tau,i}dB_t \tag{141}$$

Then the diffusion of $d\omega_t^i P_t^{(\tau),i}$ is simply:

$$\omega_t^i P_t^{(\tau),i}\left(\zeta_{r,t}^{\tau,i} - \omega_t^i\sum_{j\neq i}\xi_t^{j,i}\frac{\lambda^i}{\lambda^j}(\epsilon_t^j - \epsilon_t^i)\right)dB_t \tag{142}$$

Hence I can recover the total bond return volatility as:

$$\underbrace{\frac{\sum_{i=1}^{N}\left(\omega_t^i P_t^{(\tau),i}\left(\zeta_{r,t}^{\tau,i} - \omega_t^i \sum_{j\neq i} \frac{\lambda^i}{\lambda^j}\xi_t^{j,i}(\epsilon_t^j - \epsilon_t^i)\right)\right)}{\sum_{i=1}^{N}\omega_t^i P_t^{(\tau),i}}}_{\zeta_{r,t}^{\tau}} dB_t \tag{143}$$

Hence exposures solve:

$$x_t^i \int_0^\infty \theta e^{-\theta\tau} \frac{\sum_{i=1}^{N}\left(\omega_t^i P_t^{(\tau),i}\left(\zeta_{r,t}^{\tau,i} - \sum_{j\neq i}\omega_t^j(\epsilon_t^j - \epsilon_t^i)\right)\right)}{\sum_{i=1}^{N}\omega_t^i P_t^{(\tau),i}} d\tau = \left(\sigma_W - \sum_{j\neq i}\omega_t^j(\epsilon_t^j - \epsilon_t^i)\frac{\kappa_r}{\sigma_r}\right) \tag{144}$$

Rearranging the left-hand side I get:

$$\begin{aligned}
& x_t^i \int_0^\infty \theta e^{-\theta\tau} \frac{\sum_{i=1}^{N}\left(\omega_t^i P_t^{(\tau),i}\left(\zeta_{r,t}^{\tau,i} - \sum_{j\neq i}\omega_t^j(\epsilon_t^j - \epsilon_t^i)\right)\right)}{\sum_{i=1}^{N}\omega_t^i P_t^{(\tau),i}} d\tau \\
&= x_t^i \int_0^\infty \theta e^{-\theta\tau} \frac{\sum_{i=1}^{N}\left(\omega_t^i P_t^{(\tau),i}\left(\zeta_{r,t}^{\tau,i} - \sum_{i=1}^{N}\omega_t^i\epsilon_t^i + \epsilon_t^i\right)\right)}{\sum_{i=1}^{N}\omega_t^i P_t^{(\tau),i}} d\tau \\
&= x_t^i \int_0^\infty \theta e^{-\theta\tau} \frac{\sum_{i=1}^{N}\left(\omega_t^i P_t^{(\tau),i}\left(\zeta_{r,t}^{\tau,i} + \epsilon_t^i\right)\right)}{\sum_{i=1}^{N}\omega_t^i P_t^{(\tau),i}} d\tau - x_t^i \sum_{i=1}^{N}\omega_t^i\epsilon_t^i \int_0^\infty \theta e^{-\theta\tau} d\tau \qquad (145) \\
&= x_t^i \int_0^\infty \theta e^{-\theta\tau} \frac{\sum_{i=1}^{N}\left(\omega_t^i P_t^{(\tau),i}\left(\zeta_{r,t}^{\tau,i} + \epsilon_t^i\right)\right)}{\sum_{i=1}^{N}\omega_t^i P_t^{(\tau),i}} d\tau - x_t^i \sum_{i=1}^{N}\omega_t^i\epsilon_t^i \\
&= x_t^i \int_0^\infty \theta e^{-\theta\tau} \left[\frac{\sum_{i=1}^{N}\omega_t^i P_t^{(\tau),i}\zeta_{r,t}^{\tau,i}}{\sum_{i=1}^{N}\omega_t^i P_t^{(\tau),i}} + \frac{\sum_{i=1}^{N}\omega_t^i P_t^{(\tau),i}\epsilon_t^i}{\sum_{i=1}^{N}\omega_t^i P_t^{(\tau),i}}\right] d\tau - x_t^i \sum_{i=1}^{N}\omega_t^i\epsilon_t^i
\end{aligned}$$

The right-hand side becomes:

$$\sigma_W - \sum_{j\neq i} \omega_t^j (\epsilon_t^j - \epsilon_t^i) \frac{\kappa_r}{\sigma_r} = \sigma_W - \frac{\kappa_r}{\sigma_r} \sum_{k=1}^{N} \omega_t^i \epsilon_t^i + \frac{\kappa_r}{\sigma_r} \epsilon_t^i \tag{146}$$

Multiplying each side of this equation, I obtain:

$$\sigma_W^2 - \sum_{k=1}^{N} \omega_t^i \sigma_W \epsilon_t^i \frac{\kappa_r}{\sigma_r} + \sigma_W \frac{\kappa_r}{\sigma_r} \epsilon_t^i = \mu_W + \sigma_W \frac{\kappa_r}{\sigma_r} \epsilon_t^i = \mu_W^i \tag{147}$$

and

$$\sigma_W x_t^i \int_0^\infty \theta e^{-\theta\tau} \left[\frac{\sum_{i=1}^{N} \omega_t^i P_t^{(\tau),i} \zeta_{r,t}^{\tau,i}}{\sum_{i=1}^{N} \omega_t^i P_t^{(\tau),i}} + \frac{\sum_{i=1}^{N} \omega_t^i P_t^{(\tau),i} \epsilon_t^i}{\sum_{i=1}^{N} \omega_t^i P_t^{(\tau),i}} \right] d\tau - x_t^i \sum_{i=1}^{N} \omega_t^i \epsilon_t^i \tag{148}$$

Hence I get the familiar Merton portfolio choice solution under the subjective beliefs of the dealer:

$$x_t^i = \left\{ \int_0^\infty \theta e^{-\theta\tau} \left[\frac{\sum_{i=1}^{N} \omega_t^i P_t^{(\tau),i} \zeta_{r,t}^{\tau,i}}{\sum_{i=1}^{N} \omega_t^i P_t^{(\tau),i}} + \frac{\sum_{i=1}^{N} \omega_t^i P_t^{(\tau),i} \epsilon_t^i}{\sum_{i=1}^{N} \omega_t^i P_t^{(\tau),i}} \right] d\tau - \sum_{i=1}^{N} \omega_t^i \epsilon_t^i \right\}^{-1} \times \frac{\mu_W^i}{\sigma_W} \tag{149}$$

Next, I characterize the drift of aggregate wealth.

Proof of Lemma 1. Note the following:

$$\pi_t^i = e^{-\rho t} (\rho \lambda^i w_t^i)^{-1} = e^{-\rho t} (\rho \lambda^i)^{-1} W_t^{-1} \left(1 + \sum_{j\neq i} \eta_t^{j,i} \frac{\lambda^i}{\lambda^j} \right) \tag{150}$$

Itô's lemma implies that the drift of $\frac{d\pi_t^i}{\pi_t^i}$ under dealer i's measure is :

$$-r_t - \mu_W - \sigma_W \epsilon_t^i \frac{\kappa_r}{\sigma_r} - \rho + \sigma_W^2 - \sigma_W \frac{\kappa_r}{\sigma_r} \frac{\sum_{j\neq i} \eta_t^{j,i} \frac{\lambda^i}{\lambda^j} (\epsilon_t^j - \epsilon_t^i)}{\left(1 + \sum_{j\neq i} \eta_t^{j,i} \frac{\lambda^i}{\lambda^j} \right)} \tag{151}$$

By no-arbitrage the subjective drift should equal $-r_t$, implying:

$$\begin{aligned}
\mu_W &= -\rho - \sigma_W \epsilon_t^i \frac{\kappa_r}{\sigma_r} + \sigma_W^2 - \sigma_W \frac{\kappa_r}{\sigma_r} \frac{\sum_{j\neq i} \omega_t^{j,i}(\epsilon_t^j - \epsilon_t^i)}{\left(1 + \sum_{j\neq i} \omega_t^{j,i}\right)} \\
&= -\rho - \sigma_W \epsilon_t^i \frac{\kappa_r}{\sigma_r} + \sigma_W^2 - \sigma_W \frac{\kappa_r}{\sigma_r} \sum_{j\neq i} \omega_t^j (\epsilon_t^j - \epsilon_t^i) \\
&= -\rho + \sigma_W^2 - \sigma_W \frac{\kappa_r}{\sigma_r} \sum_{i=1}^N \omega_t^i \epsilon_t^i \\
&= \sum_{i=1}^N \omega_t^i \left(-\rho + \sigma_W^2 - \sigma_W \frac{\kappa_r}{\sigma_r} \epsilon_t^i \right)
\end{aligned} \tag{152}$$

I have the following equivalence:

$$\sum_{i=1}^N \omega_t^i \sum_{j\neq i} \omega_t^j \epsilon_t^j = \sum_{i=1}^N (1 - \omega_t^i) \omega_t^i \epsilon_t^i \tag{153}$$

and

$$\sum_{i=1}^N \omega_t^i \sum_{j\neq i} \epsilon_t^j = \sum_{i=1}^N (1 - \omega_t^i) \epsilon_t^i \tag{154}$$

Then:

$$\begin{aligned}
&\mu_W + \sigma_W \epsilon_t^i \frac{\kappa_r}{\sigma_r} = -\rho + \sigma_W^2 - \sigma_W \frac{\kappa_r}{\sigma_r} \sum_{j\neq i} \omega_t^j \epsilon_t^j + \sigma_W \frac{\kappa_r}{\sigma_r} (1 - \omega_t^i) \epsilon_t^i \\
&\implies \sum_{i=1}^N \omega_t^i \left(\mu_W + \sigma_W \epsilon_t^i \frac{\kappa_r}{\sigma_r} \right) = -\rho + \sigma_W^2 - \sigma_W \frac{\kappa_r}{\sigma_r} \sum_{i=1}^N \omega_t^i \sum_{j\neq i} \omega_t^j \epsilon_t^j + \sigma_W \frac{\kappa_r}{\sigma_r} \sum_{i=1}^N \omega_t^i (1 - \omega_t^i) \epsilon_t^i \\
&= -\rho + \sigma_W^2 - \sigma_W \frac{\kappa_r}{\sigma_r} \sum_{i=1}^N (1 - \omega_t^i) \omega_t^i \epsilon_t^i + \sigma_W \frac{\kappa_r}{\sigma_r} \sum_{i=1}^N (1 - \omega_t^i) \omega_t^i \epsilon_t^i = -\rho + \sigma_W^2
\end{aligned} \tag{155}$$

That is, the wealth-share weighted average of subjective expected wealth growth is constant and equals $-\rho + \sigma_W^2$. Using the equivalence $Y_t = \rho W_t$ completes the proof. □

Now, consider:

$$\omega_t^i = \left(1 + \sum_{j\neq i} \omega_t^{j,i}\right)^{-1} \tag{156}$$

$$d\omega_t^i \epsilon_t^i = \epsilon_t^i d\omega_t^i + \omega_t^i d\epsilon_t^i + d\omega_t^i d\epsilon_t^i \tag{157}$$

$$\frac{d\eta_t^{j,i}}{\eta_t^{j,i}} = \frac{d\omega_t^{j,i}}{\omega_t^{j,i}} = (\epsilon_t^j - \epsilon_t^i)\frac{\kappa_r}{\sigma_r}dB_t - (\epsilon_t^j - \epsilon_t^i)\epsilon_t^i\left(\frac{\kappa_r}{\sigma_r}\right)^2 dt \tag{158}$$

Then under the objective measure, ω_t^i evolves as:

$$\begin{aligned}\frac{d\omega_t^i}{\omega_t^i} &= \left(\frac{\kappa_r}{\sigma_r}\right)^2 \sum_{j\neq i}\left[\omega_t^j(\epsilon_t^j - \epsilon_t^i)\epsilon_t^i + {\omega_t^j}^2(\epsilon_t^j - \epsilon_t^i)^2\right] \\ &\sum_{j\neq i}\omega_t^j(\epsilon_t^j - \epsilon_t^i)\frac{\kappa_r}{\sigma_r}dB_t\end{aligned} \tag{159}$$

Now recall that under i's beliefs, the relative wealth ratios evolve as:

$$\frac{d\omega_t^{j,i}}{\omega_t^{j,i}} = (\epsilon_t^j - \epsilon_t^i)\frac{\kappa_r}{\sigma_r}dB_t^i \tag{160}$$

Then:

$$\begin{aligned} d\left(1+\sum_{j\neq i}\omega_t^{j,i}\right) &= \sum_{j\neq i} d\omega_t^{j,i} = \sum_{j\neq i}\omega_t^{j,i}(\epsilon_t^j - \epsilon_t^i)\frac{\kappa_r}{\sigma_r}dB_t^i \\ &= \frac{1}{\omega_t^i}\sum_{j\neq i}\omega_t^j(\epsilon_t^j - \epsilon_t^i)\frac{\kappa_r}{\sigma_r}dB_t^i = \frac{1}{\omega_t^i}\sum_{j=1}^{N}\omega_t^j\epsilon_t^j\frac{\kappa_r}{\sigma_r}dB_t^i - \frac{1}{\omega_t^i}\epsilon_t^i\frac{\kappa_r}{\sigma_r}dB_t^i\end{aligned} \tag{161}$$

Note that:

$$\left(1+\sum_{j\neq i}\omega_t^{j,i}\right) = \left(1+\frac{1}{\omega_t^i}\sum_{j\neq i}\omega_t^j\right) = 1 + \frac{1-\omega_t^i}{\omega_t^i} = \frac{1}{\omega_t^i} \tag{162}$$

Define: $\omega_t^i = \frac{w_t^i}{w_t^i + X_t}$:

$$\begin{aligned}\frac{d\omega_t^i}{\omega_t^i(1-\omega_t^i)} &= \frac{dw_t^i}{w_t^i} - \frac{dX_t}{X_t} - \omega_t^i\left(\frac{dw_t^i}{w_t^i}\right)^2 + (1-\omega_t^i)\left(\frac{dX_t}{X_t}\right)^2 + \omega_t^i\frac{dw_t^i}{w_t^i}\frac{dX_t}{X_t} - (1-\omega_t^i)\frac{dw_t^i}{w_t^i}\frac{dX_t}{X_t} \\ &= \frac{dw_t^i}{w_t^i} - \frac{dX_t}{X_t} - \omega_t^i\frac{dw_t^i}{w_t^i}\left(\frac{dw_t^i}{w_t^i} - \frac{dX_t}{X_t}\right) - (1-\omega_t^i)\frac{dX_t}{X_t}\left(\frac{dw_t^i}{w_t^i} - \frac{dX_t}{X_t}\right) \\ &= \left(1 - \frac{dW_t}{W_t}\right)\left(\frac{dw_t^i}{w_t^i} - \frac{dX_t}{X_t}\right) \\ \implies d\omega_t^i &= \omega_t^i(1-\omega_t^i)\left[\frac{dw_t^i}{w_t^i} - \frac{d\sum_{j\neq i}w_t^j}{\sum_{j\neq i}w_t^j} - \frac{dW_t}{W_t}\left(\frac{dw_t^i}{w_t^i} - \frac{d\sum_{j\neq i}w_t^j}{\sum_{j\neq i}w_t^j}\right)\right]\end{aligned} \tag{163}$$

This equation characterizes the evolution of i's wealth share. Recall that:

$$w_t^i = \left(1 + \sum_{j \neq i} \omega_t^{j,i}\right)^{-1} W_t \tag{164}$$

Then by Itô's lemma, I have:

$$\frac{dw_t^i}{w_t^i} = r_t + \mu_W - \sigma_W^2 + \sigma_W \frac{\kappa_r}{\sigma_r} \sum_{j=1}^N \omega_t^j \epsilon_t^j + \left(-\sigma_W + \frac{\kappa_r}{\sigma_r} \sum_{j=1}^N \omega_t^j \epsilon_t^j - \frac{\kappa_r}{\sigma_r} \epsilon_t^i\right)^2 + \left(\sigma_W - \sum_{j \neq i} \omega_t^j (\epsilon_t^j - \epsilon_t^i) \frac{\kappa_r}{\sigma_r}\right) dB_t^i \tag{165}$$

Given the definition of μ_W, this reduces to:

$$\frac{dw_t^i}{w_t^i} = \left[r_t - \rho + \left(-\sigma_W + \frac{\kappa_r}{\sigma_r} \sum_{j=1}^N \omega_t^j \epsilon_t^j - \frac{\kappa_r}{\sigma_r} \epsilon_t^i\right)^2\right] dt - \left(-\sigma_W + \frac{\kappa_r}{\sigma_r} \sum_{j=1}^N \omega_t^j \epsilon_t^j - \frac{\kappa_r}{\sigma_r} \epsilon_t^i\right) dB_t^i \tag{166}$$

Note that combined with the optimality condition, this equation imposes the exact same restriction on exposures.

Now note that under i's beliefs, the drift of $\sum_{j \neq i} \omega_t^{j,i}$ is zero, meaning that the drift of $\frac{X_t}{w_t^i}$ is zero. I also have:

$$d\omega_t^i = d\left(\frac{w_t^i}{w_t^i + X_t}\right) = -d\left(\frac{X_t}{w_t^i + X_t}\right) = -d\left(\sum_{j \neq i} \omega_t^j\right) = -d(1 - \omega_t^i) \tag{167}$$

Then:

$$d\left(\frac{X_t}{w_t^i}\right) = d\left(\frac{X_t/W_t}{w_t^i/W_t}\right) = \frac{X_t}{w_t^i}\left[\frac{d(X_t/W_t)}{X_t/W_t} - \frac{d(w_t^i/W_t)}{w_t^i/W_t} + \frac{d(w_t^i/W_t)}{w_t^i/W_t}\left(\frac{d(w_t^i/W_t)}{w_t^i/W_t} - \frac{d(X_t/W_t)}{X_t/W_t}\right)\right] \tag{168}$$

Now replace each $d(X_t/W_t)$ with $-d(w_t^i/W_t)$:

$$\begin{aligned} d\left(\frac{X_t}{w_t^i}\right) &= d\left(\frac{X_t/W_t}{w_t^i/W_t}\right) = \frac{X_t}{w_t^i}\left[-\frac{d(w_t^i/W_t)}{X_t/W_t} - \frac{d(w_t^i/W_t)}{w_t^i/W_t} + \frac{d(w_t^i/W_t)}{w_t^i/W_t}\left(\frac{d(w_t^i/W_t)}{w_t^i/W_t} + \frac{d(w_t^i/W_t)}{X_t/W_t}\right)\right] \\ &= \frac{W_t}{w_t^i}\left[-\frac{d(w_t^i/W_t)}{w_t^i/W_t} + \left(\frac{d(w_t^i/W_t)}{w_t^i/W_t}\right)^2\right] \end{aligned} \tag{169}$$

B.7 Estimation

Recall the processes:

$$dr_t = -\kappa_r(r_t - \mu_t)dt + \sigma_r dB_t \tag{170}$$

where the drift μ_t follows the law of motion:

$$d\mu_t = -\kappa_\mu(\mu_t - \bar{\mu})dt + \sigma_\mu dB_t \tag{171}$$

$$\mu_t^i = \mu_t + \epsilon_t^i \tag{172}$$

where the law of motion of ϵ_t^i is:

$$d\epsilon_t^i = -\kappa_\epsilon \epsilon_t^i dt + \sigma_\epsilon^i dB_t \tag{173}$$

$$\frac{d\xi_t^i}{\xi_t^i} = \epsilon_t^i \frac{\kappa_r}{\sigma_r} dB_t \tag{174}$$

Recall

$$d\tilde{B}_t^i = dB_t - \epsilon_t^i \frac{\kappa_r}{\sigma_r} dt \tag{175}$$

Under i's beliefs, ϵ_t^i follows:

$$d\epsilon_t^i = -\left(\kappa_\epsilon - \sigma_\epsilon^i \frac{\kappa_r}{\sigma_r}\right)\epsilon_t^i dt + \sigma_\epsilon^i d\tilde{B}_t^i \tag{176}$$

Under the econometrician's measure, the dynamics of μ_t^i follow:

$$d\mu_t^i = d\mu_t + d\epsilon_t^k = -\kappa_\mu(\mu_t^i - \bar{\mu})dt + \big(\kappa_\mu - \kappa_\epsilon\big)\epsilon_t^i dt + \big(\sigma_\mu + \sigma_\epsilon^i\big)dB_t \tag{177}$$

Under i's beliefs, I get the following instead:

$$d\mu_t^i = d\mu_t + d\epsilon_t^k = -\kappa_\mu(\mu_t - \bar{\mu})dt - \left(\kappa_\epsilon - \frac{\kappa_r}{\sigma_r}\left(\sigma_\mu + \sigma_\epsilon^i\right)\right)\epsilon_t^i dt + \left(\sigma_\mu + \sigma_\epsilon^i\right)d\tilde{B}_t^i \tag{178}$$

The corresponding (subjective) state-space in discrete time is as follows:

$$\epsilon_{t+1}^i = \left(1 - \left(\kappa_\epsilon - \sigma_\epsilon^i \frac{\kappa_r}{\sigma_r}\right)\Delta t\right)\epsilon_t^i + \sigma_\epsilon^i \sqrt{\Delta t}\ \varepsilon_{t+1}^i \tag{179}$$

$$\mu_{r,t+1}^i = \kappa_\mu \bar{\mu}\Delta t + (1 - \kappa_\mu \Delta t)\ \mu_t + \left(\kappa_\epsilon - \frac{\kappa_r}{\sigma_r}\left(\sigma_\mu + \sigma_\epsilon^i\right)\right)\Delta t\ \epsilon_t^i + \left(\sigma_\mu + \sigma_\epsilon^i\right)\sqrt{\Delta t}\epsilon_{t+1}^i \tag{180}$$

$$r_{t+1} = (1-\kappa_r\Delta t)\ r_t + \kappa_r\Delta t\ \mu_t^i + \sigma_r\sqrt{\Delta t}\ \varepsilon_{t+1}^i \tag{181}$$

where Δt corresponds to the time-interval dt, and ϵ_{t+1}^i corresponds to the Brownian increment $B_{t+\Delta t}^i - B_t^i$. More concisely, I have:

$$\begin{aligned} \epsilon_{t+1}^i &= \theta_\epsilon^i \epsilon_t^i + \nu_\epsilon^i \varepsilon_{t+1}^i \\ \mu_{r,t+1}^i &= (1-\theta_\mu)\bar{\mu} + \theta_\mu \mu_t + \theta_{\mu,\epsilon}^i\ \epsilon_t^i + \nu_\mu \varepsilon_{t+1}^i \\ r_{t+1} &= \theta_r r_t + \theta_{r,\mu}\mu_t^i + \nu_r \varepsilon_{t+1}^i \end{aligned} \tag{182}$$

True state-space system for short-rate dynamics is:

$$\begin{aligned} r_{t+1} &= \theta_r r_t + (1-\theta_r)\mu_t + \nu_r \varepsilon_{t+1} \\ \mu_{r,t+1} &= (1-\theta_\mu)\bar{\mu} + \theta_\mu \mu_t + \nu_\mu \varepsilon_{t+1} \end{aligned} \tag{183}$$

Rewriting I get the relation:

$$\begin{aligned} \mu_t &= \frac{1}{(1-\theta_r)} r_{t+1} - \frac{\theta_r}{(1-\theta_r)} r_t - \frac{\nu_r}{(1-\theta_r)} \varepsilon_{t+1} \\ r_{t+1} - \theta_r r_t &= \frac{(1-\theta_r)}{\theta_\mu}\mu_{r,t+1} - \frac{(1-\theta_r)}{\theta_\mu}(1-\theta_\mu)\bar{\mu} + \left(\nu_r + \frac{(1-\theta_r)}{\theta_\mu}\nu_\mu\right)\varepsilon_{t+1} \end{aligned} \tag{184}$$

I further have:

$$E_t^i[r_{t+1}] = \theta_r r_t + (1-\theta_r)\mu_t + (1-\theta_r)\epsilon_t^i \tag{185}$$

$$\begin{aligned} E_{t+1}^i[r_{t+2}] &= \theta_r r_{t+1} + \theta_{r,\mu}\mu_{r,t+1}^i \\ &= \theta_r^2 r_t + \theta_r(1-\theta_r)\mu_t + \theta_r \nu_r \varepsilon_{t+1} \\ &+ (1-\theta_r)(1-\theta_\mu)\bar{\mu} + (1-\theta_r)\theta_\mu\mu_t + (1-\theta_r)\nu_\mu\varepsilon_{t+1} \\ &+ (1-\theta_r)\theta_\epsilon \epsilon_t^i + (1-\theta_r)\nu_\epsilon^i \varepsilon_{t+1} \end{aligned} \tag{186}$$

Then the forecast errors are:

$$r_{t+1} - E_t^i[r_{t+1}] = -(1-\theta_r)\epsilon_t^i + \nu_r \varepsilon_{t+1} \tag{187}$$

$$\begin{aligned} r_{t+2} - E_{t+1}^i[r_{t+2}] &= -(1-\theta_r)\epsilon_{t+1}^i + \nu_r\varepsilon_{t+2} \\ &= -\theta_\epsilon(1-\theta_r)\epsilon_t^i - \nu_\epsilon^i(1-\theta_r)\varepsilon_{t+1} + \nu_r\varepsilon_{t+2} \\ &= \theta_\epsilon\Big(r_{t+1} - E_t^i[r_{t+1}]\Big) - \Big(\nu_\epsilon^i(1-\theta_r) + \theta_\epsilon\nu_r\Big)\varepsilon_{t+1} + \nu_r\varepsilon_{t+2} \end{aligned} \tag{188}$$

Hence I obtain the ARMA(1,1) representation of dealer forecast errors.