A Preferred-Habitat Model of the Term Structure of Interest Rates

Dimitri Vayanos*
London School of Economics
CEPR and NBER

Jean-Luc Vila†
Merrill Lynch

March 15, 2007‡

Abstract

This paper develops a term-structure model in which investors with preferences for specific maturities trade with risk-averse arbitrageurs. Arbitrageurs integrate the markets for different maturities, incorporating information about expected short rates into bond prices. We show that bond risk premia are negatively related to short rates and positively to term-structure slope. Moreover, forward rates under-react to expected short rates, especially for long maturities, while investor demand impacts mainly long maturities. Thus, the short end of the term structure is mainly driven by short-rate expectations, while the long end by demand. Despite the presence of two distinct economic factors, the first principal component "explains" about 90% of movement. Our results are consistent with empirical evidence and generate novel testable implications.

* d.vayanos@lse.ac.uk

† Jean-Luc_Vila@ml.com

‡ We thank Markus Brunnermeier, Andrea Buraschi, Peter DeMarzo, Xavier Gabaix, Robin Greenwood, Arvind Krishnamurthy, Anna Pavlova, Jeremy Stein, seminar participants at LSE, and conference participants at Imperial 2007 and SITE 2006 for helpful comments. We are grateful to Giorgio Fossi, Jason Kastner and David Webb for explaining to us aspects of the UK pension reform and its impact on the term structure.

1 Introduction

The government-bond market involves many distinct investor clienteles. For example, pension funds invest typically in maturities longer than fifteen years as a way to hedge their long-term liabilities. Life-insurance companies also have preferences for long maturities, typically around the fifteen-year mark. On the other hand, asset managers and banks' treasury departments are the typical investors for maturities shorter than ten years. Clienteles' demands vary over time in response to regulatory and other changes. This time-variation can be an important source of term-structure movements.

The UK pension reform provides a stark illustration of how clientele demands can impact the term structure. Starting in 2005, pension funds were required to mark their liabilities to market, discounting them at the rates of long-maturity bonds. As a result, hedging demands increased significantly, especially for inflation-indexed bonds because pensions are linked to inflation. The impact on the term structure was dramatic. For example, in January 2006 the inflation-indexed bond maturing in 2011 was yielding 1.5%, while the 2055 bond was yielding only 0.6%. The 0.6% yield is very low relative to the 3% historical average of long real rates in the UK. Moreover, the downward-sloping term structure is hard to attribute to expectations of future spot rates since real rates would have to drop below 0.6% after 2011. Section 2 presents more extensive evidence, and argues that the reform had strong and long-lived effects both on the real and the nominal term structure.

The role of clienteles is emphasized in some early work on the term structure. The market-segmentation hypothesis of Culbertson (1957) and others posits that investors have strong preferences for specific maturities. Each maturity then constitutes a separate market and the yield is determined by local demand and supply independently of other maturities. Modigliani and Sutch (1966) appeal to market segmentation in formulating their preferred-habitat hypothesis. On the other hand, modern theories such as Cox, Ingersoll and Ross (CIR 1985) do not build on these ideas because they assume a representative agent rather than heterogeneous clienteles. In fact, CIR criticize the market-segmentation hypothesis since it implies that bonds with similar maturities can trade at very different prices.

In this paper we take the view that the time-varying demands of heterogenous clienteles are an important source of term-structure movements. And while we acknowledge the criticism of the market-segmentation hypothesis, we formalize a less extreme version in which investors with strong maturity preferences trade with arbitrageurs. Arbitrageurs in our model integrate the markets for different maturities, ensuring that bonds with similar maturities trade at similar prices. But because arbitrageurs are risk averse, clientele demands impact the term structure. We characterize the impact of demand shocks, and compare with shocks to short-rate expectations. We show that consistent with common views held by practitioners, short-rate expectations are the main driver of term-structure movements for short maturities, while demand dominates for longer maturities. Our model also provides a theory for the risk premia inherent in the term structure that is both consistent with a number of puzzling empirical facts and generates novel testable implications.

Our model, described in Section 3, is set in continuous time. The short rate follows an exogenous Ornstein-Uhlenbeck process, and the prices of zero-coupon bonds are determined endogenously through trading between investors and arbitrageurs. Following the spirit of the market-segmentation hypothesis, we assume that for each maturity there exists an investor clientele consuming at that maturity and demanding only the corresponding zero-coupon bond. Thus, in the absence of arbitrageurs, each maturity would form a separate market. The (hypothetical) spot rates clearing these markets can, in general, depend both on maturity and time, but we focus on the case where they only depend on time. Thus, shocks to clientele demands are common across maturities, and the term structure would be flat in the absence of arbitrageurs. Arbitrageurs can invest in all maturities, and maximize a mean-variance objective over instantaneous changes in wealth.

Section 4 considers the simple case where there are no demand shocks, meaning that the term structure would be constant over time without arbitrageurs. This yields an one-factor model with the short rate as the only risk factor. In the one-factor model, arbitrageurs bridge the disconnect between the constant term structure and the time-varying short rate, incorporating information about expected short rates into bond prices. Suppose, for example, that the short rate increases, thus becoming attractive relative to investing in bonds. Investors do not take advantage of this opportunity because they prefer the safety of the bond that matures at the time when they need to consume. But arbitrageurs do take advantage by shorting bonds and investing at the short rate. Through this reverse-carry trade, bond prices decrease, thus responding to the high short rate. Conversely, following a negative shock to the short rate, arbitrageurs engage in a carry trade, borrowing short and buying bonds.

Bond risk premia can be deduced from arbitrageurs' trading strategies. Consider, for example, the case where arbitrageurs are short bonds because the short rate is high. Since arbitrageurs are the marginal agents and are risk-averse, bonds earn negative premia. Conversely, premia are positive when the short rate is low. Premia are, therefore, negatively related to the short rate. Moreover, since an increase in the short rate translates to a decrease in the slope of the term structure, premia are positively related to slope. In this sense, our model can provide a natural explanation for the puzzling empirical finding in Fama and Bliss (1987) that premia are positively related to term-structure slope and are highly variable relative to their unconditional average.

The behavior of bond risk premia is reflected into that of forward rates. Because information about expected short rates is incorporated into bond prices by risk-averse arbitrageurs, forward rates under-react to changes in expected short rates. This relates to the behavior of bond risk premia: for example, an increase in short rates raises forward rates, but the effect is tempered by premia becoming negative. We show that under-reaction becomes stronger for longer maturities because premia for those maturities are more important: speculating on a difference between a long-maturity forward rate and the corresponding short rate involves more risk. Forward rates are also influenced by investor demand (a constant parameter in the one-factor model) and the effect is stronger for longer maturities. This suggests that short-rate expectations are the main driver of term-structure movements for short maturities, while demand–manifested through risk premia–is the main driver for long maturities. We confirm this result in Section 5, where we consider stochastic demand shocks.

Our results are linked to properties of the risk-neutral measure (which exists in our model because arbitrageurs render the term structure arbitrage-free). The risk-neutral measure solves a fixed-point problem: it depends on the risk-averse arbitrageurs' bond holdings, which in turn depend on bond prices and thus on the risk-neutral measure. We show that the risk-neutral dynamics of the short rate are Ornstein-Uhlenbeck, as are the true dynamics, but with faster mean-reversion and a long-run mean influenced by demand. The under-reaction of forward rates to expected short rates is linked to the fast mean-reversion, while the effects of demand are larger for long maturities because demand enters through the long-run mean. Moreover, the market price of risk in our model changes sign, in line with the reduced-form specifications proposed by Dai and Singleton (2002) and Duffee (2002).[1]

Section 5 considers the case where there are demand shocks, meaning that in the absence of arbitrageurs the term structure moves up (negative demand shock) or down (positive demand shock) while remaining flat and independent of the short rate. Demand shocks reinforce the positive relationship between bond risk premia and term-structure slope, derived in the one-factor model. Indeed, an upward-sloping term structure can arise because the short rate is low or because demand is low. In both cases arbitrageurs hold long positions in bonds and premia are positive. Demand

[1] Dai and Singleton (2002) and Duffee (2002) argue that CIR-type specifications in which premia are proportional to volatility do not match important properties of the data, and they propose new specifications in which premia can change sign. Such specifications arise naturally in our structural model.

shocks also induce a negative relationship between term-structure slope and changes to long rates, consistent with Campbell and Shiller's (1991) empirical finding. Indeed, if an upward-sloping term structure arises because the short rate is low, then long rates are likely to increase as the short rate mean-reverts. But if it arises because demand is low, then long rates are likely to decrease as demand mean-reverts.

Because demand shocks increase arbitrageurs' risk, they accentuate the under-reaction of forward rates to expected short rates. In fact, long-maturity forward rates can even move in the opposite direction to short rates. The intuition is that arbitrageurs incorporate an increase in short rates into prices by shorting bonds. Because, however, demand risk makes long-maturity bonds particularly risky, arbitrageurs do not extend their shorting to long maturities, buying instead those maturities to hedge the demand risk of their positions in shorter maturities. On the other hand, demand risk accentuates the effect of any given demand shock, especially for long maturities as arbitrageurs are reluctant to take large positions in those maturities.[2] Therefore, demand risk both makes precise and strengthens the intuition from the one-factor model that short-rate expectations are the main driver of term-structure movements for short maturities, while demand is the main driver for long maturities.

That independent shocks to demand and short-rate expectations drive different segments of the term structure might seem at odds with Litterman and Scheinkman (1991), who find that one principal component (PC) explains 90% of bond-return variation. We show, however, that PC analysis generates the same result in our model even when shocks to the two factors have the same variance. Moreover, the first PC is an amalgam of short-rate and demand shocks, and its effect on yields can be very different than of each separate shock. Our results caution against interpreting PCs as economic factors, and suggest what the effect of true factors should be.

Our explanation for the relationship between bond risk premia and term-structure slope differs from a number of recent papers. Wachter (2006) and Buraschi and Jiltsov (2007) consider representative-agent models with habit formation, in which periods of low consumption are associated with high short rates and bond risk premia. The time-variation in the premia generates a positive relationship with slope, but in contrast to our model premia are positively related to the short rate. Xiong and Yan (2006) consider a CIR-type model with two agents holding heterogeneous beliefs about the time-varying mean of the short rate. When agents are overly optimistic about the mean, they undervalue the bonds, and this leads an econometrician who infers the mean

[2] This result is reminiscent of De Long, Summers, Shleifer and Waldman (1990), who show in an one-asset model that the risk of noise trading in the future amplifies the effect of current noise trading.

correctly to observe positive premia and positive slope. While we do not dispute the relevance of habit formation or heterogeneous beliefs for asset pricing, we believe that in episodes such as the UK pension reform, changes in risk premia were generated by an entirely different mechanism. Considering this mechanism leads to a new set of intuitions and predictions.

A number of recent papers study the pricing of multiple assets within a class when arbitrage is limited. In Barberis and Shleifer (2003), arbitrageurs absorb demand shocks of investors with preferences for specific asset styles. These shocks generate comovement of assets within a style.[3] In Pavlova and Rigobon (2005), style arises because of portfolio constraints, and can be the source of international financial contagion.[4] Gabaix, Krishnamurthy and Vigneron (2007) assume that the marginal holder of mortgage-backed securities is a risk-averse arbitrageur, whose wealth is tied to a mortgage portfolio rather than to economy-wide wealth. They find empirical support for their theory because pre-payment risk is priced according to the covariance with the mortgage portfolio. Garleanu, Pedersen and Poteshman (2006) assume that the marginal options trader is a risk-averse market maker who absorbs demand shocks of other investors. They characterize how demand for a given option affects the prices of all options, and find empirical support for their theory by using measures of demand pressure.[5] Our work relates to the above papers because we study the role of arbitrageurs in enforcing pricing relationships across a large number of bonds. In contrast to most of these papers, we carry our analysis in continuous time, bringing ideas of limited arbitrage more firmly into the continuous-time contingent-claims framework.

2 UK Pension Reform and the Term Structure

SECTION TO BE WRITTEN

3 Model

Time is continuous and goes from zero to infinity. The term structure at time t consists of a continuum of zero-coupon bonds in zero net supply. The maturities of the bonds are in the interval $(0, T]$, and the bond with maturity τ pays \$1 at time $t + \tau$. We denote by $P_{t,\tau}$ the time-t price of

[3] See also Spiegel (1998) for a general correlation structure of demand shocks.

[4] Kyle and Xiong (2001) derive financial contagion from the wealth effects of arbitrageurs with logarithmic preferences. In Gromb and Vayanos (2002), wealth effects arise because of arbitrageurs' margin constraints. These constraints hinder the arbitrageurs' ability to provide liquidity and to integrate segmented markets.

[5] See also Bollen and Whaley (2004) for an empirical analysis of demand effects for options, and Bates (2006) for a model in which option demand is generated by crash-averse agents.

the bond with maturity τ and by $R_{t,\tau}$ the spot rate for that maturity. The spot rate is related to the price through

$$R_{t,\tau} = -\frac{\log(P_{t,\tau})}{\tau}. \tag{1}$$

The short rate r_t is the limit of $R_{t,\tau}$ when τ goes to zero. We take r_t as exogenous and assume that it follows the Ornstein-Uhlenbeck process

$$dr_t = \kappa_r(\overline{r} - r_t)dt + \sigma_r dB_{r,t}, \tag{2}$$

where $(\overline{r}, \kappa_r, \sigma_r)$ are constants and $B_{r,t}$ is a Brownian motion. The short rate r_t could be determined by the Central Bank and the macro-economic environment, but we do not model these mechanisms. Our focus instead is on how exogenous movements in r_t influence the bond prices $P_{\tau,t}$ that are endogenously determined in equilibrium.

Agents are of two types: investors and arbitrageurs. Investors have preferences for bonds of specific maturities. Examples are pension funds, whose typical preferences are for maturities longer than fifteen years, life-insurance companies, with preferences for maturities around fifteen years, and asset managers and banks' treasury departments, with preferences for maturities shorter than ten years. We assume that preferences take an extreme form whereby each investor demands only a specific maturity. This assumption is very convenient analytically, and we argue below that it should not affect the main intuitions and results. The set of investors demanding maturity τ constitutes the clientele for the bond with the same maturity, and the maturity τ is the clientele's habitat. We assume that the demand of the clientele for the bond with maturity τ is an increasing function of the bond's yield $R_{t,\tau}$, and we adopt a simple linear specification where the demand is for

$$y_{t,\tau} = \alpha(\tau)\tau(R_{t,\tau} - \beta_{t,\tau}) \tag{3}$$

time-t dollars worth of the bond. We impose no restrictions on the function $\alpha(\tau)$ except that it is positive, and consider specifications for the demand intercept $\beta_{t,\tau}$ later in this section.

Eq. (3) implies an extreme form of market segmentation. If investors were the only market participants, then each maturity would constitute a separate market with the yield being determined by local demand and supply. Given the demand (3) and the bonds' zero net supply, the equilibrium yield for maturity τ would be $R_{t,\tau} = \beta_{t,\tau}$. Of course, such segmentation does not occur in equilibrium because of the arbitrageurs. Arbitrageurs integrate markets, ensuring that bonds with similar maturities trade at similar prices.

In Appendix A we provide a utility-based foundation for the demand (3) based on the notion that investors have infinite risk aversion for consumption at specific times. In particular, we assume that the clientele for maturity τ consumes at time $t+\tau$ and is infinitely risk-averse over consumption. Infinite risk aversion ensures that the clientele considers only the bond with maturity τ and not other bonds.[6] Of course, infinite risk aversion is an extreme assumption, but we believe that it should not affect the main intuitions and results. Indeed, under finite risk aversion, the clientele for maturity τ would substitute between bonds with maturities close to τ, trading off risk and return. In a sense, however, arbitrageurs are doing exactly this type of substitution. So endowing clienteles with finite risk aversion would be qualitatively similar to reducing the risk aversion of arbitrageurs. The advantage of allowing only arbitrageurs to substitute between bonds is analytical convenience: if substitution can also be done by investors, the model would be complicated because we would need to keep track of the portfolios of diverse clienteles.

We assume that the demand intercept $\beta_{t,\tau}$ is independent of maturity τ, and denote it by β_t. Under this assumption, demand shocks are common across maturities and the term structure would be flat in the absence of arbitrageurs. Allowing demand shocks to differ across maturities is an important extension of our analysis and we leave it for future research. Section 4 considers the case where β_t is constant over time and equal to a constant β. This yields an one-factor model because the only risk factor is the short rate r_t. Section 5 assumes that β_t is time-varying and follows the Ornstein-Uhlenbeck process

$$d\beta_t = \kappa_\beta(\overline{\beta} - \beta_t) + \sigma_\beta dB_{\beta,t}, \tag{4}$$

where $(\overline{\beta}, \kappa_\beta, \sigma_\beta)$ are constants and $B_{\beta,t}$ is a Brownian motion independent of $B_{r,t}$. This yields a two-factor model in (r_t, β_t).

Arbitrageurs choose a bond portfolio to trade off instantaneous mean and variance. Denoting the arbitrageurs' time-t wealth by W_t and their dollar investment in the bond with maturity τ by $x_{t,\tau}$, the arbitrageurs' budget constraint is

$$dW_t = \left(W_t - \int_0^T x_{t,\tau}\right) r_t dt + \int_0^T x_{t,\tau} \frac{dP_{t,\tau}}{P_{t,\tau}}. \tag{5}$$

[6]This is as long as prices involve no arbitrage, which is the case in equilibrium because of the arbitrageurs. No arbitrage ensures that (i) any strategy that generates a riskless payoff at time $t+\tau$ is equivalent to buying the bond with maturity τ, and (ii) any strategy that generates a risky payoff at time $t+\tau$ does not dominate uniformly the strategy of buying the bond with maturity τ.

Arbitrageurs solve a standard mean-variance problem

$$\max_{\{x_{t,\tau}\}_{\tau\in(0,T]}} \left[E_t(dW_t) - \frac{a}{2} Var_t(dW_t)\right], \tag{6}$$

over instantaneous mean and variance, where a is a risk-aversion coefficient. Arbitrageurs can be interpreted as hedge funds or proprietary-trading desks, and their preferences over instantaneous mean and variance could be arising from short-term compensation. Intertemporal optimization under logarithmic utility would also give rise to such preferences, but the risk-aversion coefficient a would then depend on wealth. In taking a to be constant, we suppress wealth effects. We appeal informally to wealth effects, however, when drawing some of the model's empirical implications.

4 One-Factor Model

This section considers the case where r_t is the only risk factor. The one-factor model yields many of the general intuitions and results, while being very simple and tractable. In the one-factor model the term structure without arbitrageurs is constant and flat at β, while the short rate r_t is time-varying. Arbitrageurs bridge this disconnect, bringing information about the short-rate process into the term structure. We determine the extent to which arbitrageurs are able to perform this role by solving for equilibrium below.

4.1 Equilibrium

We conjecture that equilibrium bond yields are affine in r_t, i.e.,

$$P_{t,\tau} = e^{-[A_r(\tau)r_t + C(\tau)]} \tag{7}$$

for two functions $A_r(\tau), C(\tau)$ that depend on maturity τ. Applying Ito's Lemma to (7) and using the dynamics (2) of the short rate, we find that instantaneous bond returns are

$$\frac{dP_{t,\tau}}{P_{t,\tau}} = \mu_{t,\tau} dt - A_r(\tau)\sigma_r dB_{r,t}, \tag{8}$$

where

$$\mu_{t,\tau} \equiv A_r'(\tau)r_t + C'(\tau) - A_r(\tau)\kappa_r(\overline{r} - r_t) + \frac{1}{2}A_r(\tau)^2\sigma_r^2. \tag{9}$$

Using (8), we write the arbitrageurs' budget constraint (5) as

$$dW_t = \left[W_t r_t + \int_0^T x_{t,\tau}(\mu_{t,\tau} - r_t)d\tau\right] dt - \left[\int_0^T x_{t,\tau} A_r(\tau) d\tau\right] \sigma_r dB_{r,t},$$

and the arbitrageurs' optimization problem (6) as

$$\max_{\{x_{t,\tau}\}_{\tau\in(0,T]}} \left[\int_0^T x_{t,\tau}(\mu_{t,\tau} - r_t)d\tau - \frac{a\sigma_r^2}{2}\left[\int_0^T x_{t,\tau} A_r(\tau) d\tau\right]^2\right].$$

The first-order condition is

$$\mu_{t,\tau} - r_t = A_r(\tau)\lambda_r, \tag{10}$$

where

$$\lambda_r \equiv a\sigma_r^2 \int_0^T x_{t,\tau} A_r(\tau) d\tau. \tag{11}$$

Eq. (10) is the key pricing equation of the one-factor model. The left-hand side is a bond's expected return $\mu_{t,\tau}$ in excess of the short rate r_t, and the right-hand side corresponds to the bond's risk. The term $A_r(\tau)$ is the bond's loading on the short rate r_t, the only risk factor. The term λ_r is the factor risk premium, evaluated from the viewpoint of arbitrageurs who are the marginal agents. The risk premium λ_r is the product of the arbitrageurs' risk aversion coefficient a, the short rate's instantaneous variance σ_r^2, and the loading $\int_0^T x_{t,\tau} A_r(\tau) d\tau$ of the arbitrageurs' aggregate position on the short rate.

To solve for equilibrium, we combine the first-order condition (10) with market clearing. Since bonds are in zero net supply, arbitrageurs' and investors' positions are opposites. Using (3) we find

$$x_{t,\tau} = -y_{t,\tau} = \alpha(\tau)\tau(\beta - R_{t,\tau}) = \alpha(\tau)\left[\beta\tau - [A_r(\tau)r_t + C(\tau)]\right]. \tag{12}$$

Substituting (9), (11) and (12) into (10), we find an affine equation in r_t. Setting constant and linear terms to zero yields two linear ODEs in $A_r(\tau), C(\tau)$. Proposition 1 computes the solutions to the ODEs.

Proposition 1. *The functions $A_r(\tau), C(\tau)$ are given by*

$$A_r(\tau) = \frac{1 - e^{-\kappa_r^*\tau}}{\kappa_r^*}, \tag{13}$$

$$C(\tau) = \kappa_r^*\bar{r}^* \int_0^\tau A_r(u)du - \frac{\sigma_r^2}{2}\int_0^\tau A_r(u)^2 du, \tag{14}$$

where κ_r^* is the unique solution to

$$\kappa_r^* = \kappa_r + a\sigma_r^2 \int_0^T \alpha(\tau) A_r(\tau)^2 d\tau, \tag{15}$$

and

$$\overline{r}^* \equiv \overline{r} + \frac{(\beta - \overline{r}) z_\beta + z_c}{\kappa_r^*}, \tag{16}$$

$$z_\beta \equiv \frac{a\sigma_r^2 \int_0^T \alpha(\tau)\tau A_r(\tau) d\tau}{1 + a\sigma_r^2 \int_0^T \alpha(\tau) \left[\int_0^\tau A_r(u) du\right] A_r(\tau) d\tau}, \tag{17}$$

$$z_c \equiv \frac{\frac{a\sigma_r^4}{2} \int_0^T \alpha(\tau) \left[\int_0^\tau A_r(u)^2 du\right] A_r(\tau) d\tau}{1 + a\sigma_r^2 \int_0^T \alpha(\tau) \left[\int_0^\tau A_r(u) du\right] A_r(\tau) d\tau}. \tag{18}$$

To explain the intuition for Proposition 1, we consider the dynamics of the short rate under the risk-neutral measure. Since the factor risk premium λ_r is affine in r_t (from (11) and (12)), the risk-neutral dynamics are Ornstein-Uhlenbeck, as the true dynamics. In the proof of Proposition 1 we show that the risk-neutral dynamics are

$$dr_t = \kappa_r^* (\overline{r}^* - r_t) dt + \sigma_r d\hat{B}_{r,t}, \tag{19}$$

where $(\kappa_r^*, \overline{r}^*)$ are defined by (15) and (16), and $\hat{B}_{r,t}$ is the Brownian motion under the risk-neutral measure. Eqs. (13) and (14) are the standard term-structure equations when risk-neutral dynamics are Ornstein-Uhlenbeck (Vasicek (1977)). The novel aspect of our analysis consists in identifying how the risk-neutral parameters $(\kappa_r^*, \overline{r}^*)$ differ from their true counterparts $(\kappa_r, \overline{r})$. Any such differences are driven by arbitrageurs' risk aversion: with risk-neutral arbitrageurs ($a = 0$), (15)-(18) confirm that the two sets of parameters coincide.

Eq. (15) implies that $\kappa_r^* > \kappa_r$, i.e., the short rate mean-reverts faster under the risk-neutral than under the true measure. This result plays a key role in our analysis, and the intuition is as follows. Consider a positive shock to the short rate. This makes investing in the short rate attractive relative to investing in bonds. Investors do not take advantage of this opportunity because they are locked in their habitats: bonds give them a guaranteed return at the time when they need to consume, while investing in the short rate gives a random return. On the other hand, arbitrageurs take advantage of the opportunity by shorting bonds and investing at the short rate. This trade, known as reverse-carry, leaves arbitrageurs exposed to the risk that the short rate goes down because bond prices then increase. As a result, the risk-neutral probability of the short rate going down exceeds the true probability, meaning that mean-reversion is faster. Conversely, a negative shock to the short rate leads arbitrageurs to engage in a carry trade, borrowing at the short rate and buying bonds. Arbitrageurs' risk is that the short rate goes up because bond prices then decrease. As a result, the risk-neutral probability of the short rate going up exceeds the true probability, implying again faster mean-reversion.

Eq. (16) shows that the short rate's long-run mean $\overline{r}^*$ under the risk-neutral measure depends not only on the true mean $\overline{r}$, but also on the demand parameter β. Thus, investor demand influences bond prices, even when current and future expected short rates are held constant. Since $z_\beta > 0$, $\overline{r}^*$ is increasing in β, meaning that bond prices decrease when demand decreases. Moreover, $\overline{r}^*$ increases less than one-to-one with $\overline{r}$, meaning that bond prices do not respond fully to changes in future expected short rates. Section 4.2 examines in more detail how the term structure responds to changes in short-rate expectations and investor demand.

Lastly, it is worth noting that the solution in Proposition 1 involves a fixed-point problem. The pricing functions $A_r(\tau), C(\tau)$ are determined from the risk-neutral measure through standard techniques. The novel aspect of our analysis, and the source of the fixed-point problem, is that the risk-neutral measure is derived endogenously from the arbitrageurs' risk aversion and bond holdings. Arbitrageurs' holdings depend on yields since these influence investors' demand. Therefore, holdings depend on the pricing functions $A_r(\tau), C(\tau)$, giving rise to the fixed-point problem. In terms of Proposition 1, the function $A_r(\tau)$ depends on the risk-neutral mean-reversion rate κ_r^* through (13), and κ_r^* depends on $A_r(\tau)$ through (15).

4.2 Term-Structure Movements

Using our characterization of the equilibrium term structure, we can examine the impact of shocks to short-rate expectations and investor demand. We determine how such shocks affect forward rates and denote by $f_{t,\tau-\Delta\tau,\tau}$ the time-t forward rate between maturities $\tau - \Delta\tau$ and τ. The forward rate is related to the price through

$$f_{t,\tau-\Delta\tau,\tau} = -\frac{\log\left(\frac{P_{t,\tau}}{P_{t,\tau-\Delta\tau}}\right)}{\Delta\tau}. \tag{20}$$

We denote by $f_{t,\tau}$ the instantaneous forward rate for maturity τ, defined as the limit of $f_{t,\tau-\Delta\tau,\tau}$ when $\Delta\tau$ goes to zero. Eq. (20) implies that

$$f_{t,\tau} = -\frac{\partial \log(P_{t,\tau})}{\partial \tau}. \tag{21}$$

Studying the term structure in terms of forward rather than spot rates has some conceptual advantages. In particular, the instantaneous forward rate for maturity τ can be compared to the expected short rate at τ, while the spot rate for maturity τ combines expectations of all short rates up to τ.[7] Of course, mapping instantaneous forward rates into spot rates is straightforward: (1) and (21) imply that

$$R_{t,\tau} = \frac{\int_0^\tau f_{t,u} du}{\tau},$$

i.e., the spot rate for maturity τ is the average of all instantaneous forward rates up to τ.

Proposition 2 determines how forward rates respond to changes in short-rate expectations. Because the short rate follows a univariate process in our model, changes in short-rate expectations are generated by changes in the short rate. Therefore, we consider a shock to the short rate at time t and compare the responses of the instantaneous forward rate for maturity τ and the expected short rate at τ.

Proposition 2 (Effect of Short-Rate Expectations).

- *Forward rates under-react to changes in expected short rates:* $0 < \frac{\partial f_{t,\tau}}{\partial r_t} < \frac{\partial E_t(r_{t+\tau})}{\partial r_t}$.
- *Under-reaction is more severe for longer maturities:* $\frac{\frac{\partial f_{t,\tau}}{\partial r_t}}{\frac{\partial E_t(r_{t+\tau})}{\partial r_t}}$ *is decreasing in* τ.

Proposition 2 shows that the expectations hypothesis (EH) fails to hold in our model: according to the EH forward rates should move one-to-one with expected short rates, but we find under-reaction. The intuition for the under-reaction is that in the absence of arbitrageurs, forward rates are determined by investor demand and are not influenced by the short-rate process. This is because investors hold the bonds that mature at the time when they need to consume, and do not take the risk of investing in the short rate. Arbitrageurs bring information about the short-rate process into the term structure, but the extent of their activity is limited by risk aversion. As a result, short-rate

[7] Suppose, for example, that the 29- and 30-year spot rates are 4% and 3.88%, respectively. The forward rate between maturities 29 and 30 is 0.46%, and might appear very low in comparison to the expected one-year spot rate in 29 years. This comparison is not as transparent if one looks at the 29- and 30-year spot rates.

information is not fully incorporated, implying under-reaction. The intuition why under-reaction is more severe for longer maturities is subtler and we elaborate on it below. But crudely put, longer maturities are harder to arbitrage: speculating on a difference between a forward rate and the corresponding expected short rate involves more risk for longer maturities.

To gain more intuition on the relation between forward rates and expected short rates, consider the strategy that speculates on the difference between the two rates, e.g., lend at $f_{t,\tau}$ and borrow the same amount at $r_{t+\tau}$. Since this strategy requires no investment at time t, its expected payoff under the risk-neutral measure is zero, i.e.,

$$E_t^*\left[e^{-\int_0^\tau r_{t+u}du}(f_{t,\tau} - r_{t+\tau})\right] = 0, \tag{22}$$

where E_t^* denotes risk-neutral expectation. In the proof of Proposition 2 we show that under Ornstein-Uhlenbeck dynamics, (22) becomes

$$\begin{aligned} f_{t,\tau} &= E_t^*(r_{t+\tau}) - \frac{\sigma_r^2}{2}A_r(\tau)^2 \\ &= E_t(r_{t+\tau}) + [E_t^*(r_{t+\tau}) - E_t(r_{t+\tau})] - \frac{\sigma_r^2}{2}A_r(\tau)^2. \end{aligned} \tag{23}$$

The difference between the forward rate and the expected short rate is the sum of two terms: a risk premium equal to the difference between the short rate's risk-neutral and true expectation, and a convexity adjustment that does not matter for our discussion.[8] Because the short rate mean-reverts faster under the risk-neutral than under the true measure, the risk-neutral expectation of $r_{t+\tau}$ is less sensitive to shocks to r_t than the true expectation. Therefore, a positive shock to r_t generates a negative risk premium, while a negative shock generates a positive risk premium. Forward rates under-react to the shocks because the risk premia are of opposite signs to the shocks. Moreover, risk premia are small for short maturities because the expectations under the two measures are similar. They cumulate, however, to larger values for longer maturities as the expectations diverge, which is why the under-reaction becomes more severe.[9]

Proposition 3 determines how forward rates respond to changes in investor demand, measured by the parameter β. This parameter is constant in the one-factor model so the effects in Proposition 3 concern an unanticipated, one-off change.

[8] The adjustment is because losses from the speculative strategy occur when short rates are high, and are thus discounted more heavily than profits. This makes the strategy attractive and lowers the no-arbitrage forward rate. The adjustment is zero for a futures rather than a forward rate because profits and losses are settled continuously.

[9] The risk premia converge to zero when τ goes to infinity. This does not contradict Proposition 2 because the under-reaction is expressed as a ratio rather than a difference.

Proposition 3 (Effect of Investor Demand).

- *Holding expected short rates constant, forward rates increase when investor demand decreases:* $0 < \frac{\partial f_{t,\tau}}{\partial \beta} < 1$.
- *The effect is stronger for longer maturities:* $\frac{\partial f_{t,\tau}}{\partial \beta}$ *is increasing in* τ.

The intuition for Proposition 3 mirrors that of Proposition 2. In the absence of arbitrageurs, the term structure is flat at β, and an increase in β (decrease in demand) shifts it up uniformly. Arbitrageurs dampen this effect because they bring the term structure more in line with expected short rates. But because they are risk-averse, the change in β has an effect. Moreover, the effect is stronger for longer maturities because these are harder to arbitrage. Eq. (23) confirms these intuitions. Recall that the short-rate's long-run mean $\overline{r}^*$ under the risk-neutral measure is increasing in β. Therefore, an increase in β raises the short rate's risk-neutral expectation in (23), while leaving the true expectation unchanged. This raises the risk premium and the forward rate. The effect is small for short maturities because the risk-neutral and true expectation are close. It cumulates, however, to larger values for longer maturities as the expectations diverge.

Taken together, Propositions 2 and 3 suggest that short-rate expectations are the main driver of forward rates for short maturities, while demand–manifested through risk premia–is the main driver for longer maturities. Of course, such a result cannot be a formal consequence of the one-factor model because the demand parameter β is constant. But we show this result in the two-factor model of Section 5. We should emphasize that this result is very much in line with practitioners' view that the term structure is driven mainly by Central-Bank policy at the short end, inflation and growth expectations at the medium end, and demand at the long end.

4.3 Risk Premia and Predictability

We next examine the implications of our model for bond risk premia and predictability. We argue, in particular, that the combination of maturity clienteles and limited arbitrage can provide a natural explanation for a number of puzzling empirical facts.

Fama and Bliss (FB 1987) find that expected returns on bonds vary only weakly with maturity.[10] This means that long-maturity bonds carry small risk premia relative to bonds of shorter maturities. But while premia are small on average, they seem to vary dramatically over time. In

[10] See also Cochrane (1999) who updates the FB findings with more recent data.

particular, they are positive when the term structure is upward sloping, and negative when it is downward sloping. FB perform the regression

$$\frac{1}{\Delta\tau}\log\left(\frac{P_{t+\Delta\tau,\tau-\Delta\tau}}{P_{t,\tau}}\right) - R_{t,\Delta\tau} = \alpha_p + \gamma_p(f_{t,\tau-\Delta\tau,\tau} - R_{t,\Delta\tau}) + \epsilon_{t+\Delta\tau}. \tag{24}$$

The dependent variable is the return on a zero-coupon bond with maturity τ held over a period $\Delta\tau$, in excess of the spot rate for maturity $\Delta\tau$. The independent variable is the forward rate between maturities $\tau - \Delta\tau$ and τ, minus the the spot rate for maturity $\Delta\tau$. FB perform this regression for $\Delta\tau = 1$ year and $\tau = 2, 3, 4, 5$ years. They find that in all cases γ_p is positive and statistically significant. This means that risk premia of long-maturity bonds are positive when forward rates exceed spot rates, i.e., the term structure is upward sloping, and negative when the term structure is downward sloping. The time-variation is significant: the standard deviation of predicted risk premia is about 1-1.5%, while average premia are about 0.5%. FB's finding is in violation of the expectations hypothesis (EH) which predicts that premia are zero and $\gamma_p = 0$.

A positive relationship between risk premia and term-structure slope arises naturally in our model. The term structure is upward sloping at times when the short rate r_t is low. Expected short rates are then low, and arbitrageurs incorporate this information into the term structure by borrowing short and buying bonds. Since arbitrageurs are the marginal agents, bonds carry positive premia. Conversely, the term structure is downward sloping at times when r_t is high. Arbitrageurs then short bonds and premia are negative. To show the positive relationship between premia and slope in our model, we compute the FB regression coefficient γ_p in the analytically convenient case where returns are over a short period $\Delta\tau$.

Proposition 4. *For $\Delta\tau \to 0$ and for all τ, the FB regression coefficient in (24) is $\gamma_p = \frac{\kappa_r^* - \kappa_r}{\kappa_r^*} > 0$.*

Additionally, our model generates a negative relationship between risk premia and the short rate. To show this relationship, we compute the regression coefficient of expected excess returns on the spot rate

$$\frac{1}{\Delta\tau}\log\left(\frac{P_{t+\Delta\tau,\tau-\Delta\tau}}{P_{t,\tau}}\right) - R_{t,\Delta\tau} = \alpha_s + \gamma_s R_{t,\Delta\tau} + \epsilon_{t+\Delta\tau}, \tag{25}$$

for small $\Delta\tau$.

Proposition 5. *For $\Delta\tau \to 0$, the regression coefficient in (25) is $\gamma_s = -(\kappa_r^* - \kappa_r)A_r(\tau) < 0$.*

The negative relationship between premia and the short rate is reflected in the short rate's factor risk premium λ_r. In the proof of Proposition 5 we show that

$$\lambda_r = \kappa_r^* \overline{r}^* - \kappa_r \overline{r} - r_t(\kappa_r^* - \kappa_r). \tag{26}$$

Therefore, λ_r is an affine function of r_t, positive for small values and negative for large values. The changing sign of λ_r ensures that bond risk premia change sign with the short rate (and with term-structure slope). It is worth noting that in many equilibrium term-structure models, λ_r does not change sign and neither do bond risk premia. For example, in Cox, Ingersoll and Ross (CIR 1985), this is because λ_r is proportional to the squared volatility of the short rate. CIR-type specifications are adopted in many subsequent reduced-form models. Dai and Singleton (DS 2002) and Duffee (2002) argue that such specifications do not match important properties of the data. They propose a new class of specifications that generate a better match while retaining the tractability of affine models. In their specifications risk premia can change sign. Our relative contribution is to show that such specifications can arise naturally in an equilibrium model with maturity clienteles and limited arbitrage. Moreover, because our model is structural rather than reduced-form, it can suggest which specifications within the large class considered in DS and Duffee are more economically plausible. Within our model, we can also give economic interpretations to term-structure factors and risk premia, and relate them to exogenous variables.

We next explore the implications of our model for the predictability of long-rate changes. Campbell and Shiller (CS 1991) find that the slope of the term structure can predict changes to long rates, but to a weaker and typically opposite extent than implied by the EH. Their regression is

$$R_{t+\Delta\tau,\tau-\Delta\tau} - R_{t,\tau} = \alpha_r + \gamma_r \frac{\Delta\tau}{\tau - \Delta\tau}(R_{t,\tau} - R_{t,\Delta\tau}) + \epsilon_{t+\Delta\tau}. \tag{27}$$

The dependent variable is the change, between times t and $t + \Delta\tau$, in the yield of a zero-coupon bond that has maturity τ at time t. The independent variable is an appropriately normalized difference between the spot rates for maturities τ and $\Delta\tau$. According to the EH, the coefficient γ_r should be one. CS find, however, that γ_r is smaller than one, typically negative, and decreasing with τ. This finding is related to the behavior of risk premia documented in FB. Indeed, suppose that risk premia are positive when the term structure is upward sloping. Then, because bonds earn positive expected returns, their yields increase on average by less than in the EH benchmark. Therefore, the CS regression coefficient γ_r is smaller than one. That γ_r is negative especially for long maturities can be viewed as evidence that risk premia of long-maturity bonds are strongly time-varying. Proposition 6 computes the CS regression coefficient γ_r in our model, in the analytically convenient case where $\Delta\tau$ is small.

Proposition 6. *For $\Delta\tau \to 0$, the CS regression coefficient in (27) is*

$$\gamma_r = 1 - \frac{(\kappa_r^* - \kappa_r)A_r(\tau)}{1 - \frac{A_r(\tau)}{\tau}} < 1. \tag{28}$$

It is increasing in τ, negative for small τ if $\kappa_r^* > 2\kappa_r$, and positive for large τ.

Since our model generates a positive relationship between premia and slope, it also generates a CS regression coefficient γ_r smaller than one. Contrary to CS, however, γ_r is increasing in τ and becomes positive for large τ. This is because with the short rate as the only factor, risk premia for long maturities do not exhibit sufficiently strong time-variation. The variation in these premia becomes stronger, however, when the demand parameter β is stochastic. Indeed, the effects of β are precisely through the premia, and are especially important for long maturities from Proposition 3. Section 5 shows that the coefficient γ_r in the two-factor model can be negative and decreasing in τ.

An important message of Propositions 2-6 is that the combination of maturity clienteles and limited arbitrage can underlie a wide range of phenomena: under-reaction of forward rates to short-rate expectations (Proposition 2), demand effects on the term structure (Proposition 3), positive relationship between premia and slope (Proposition 4), negative relationship between premia and the short rate (Proposition 5), and CS coefficients smaller than one (Proposition 6). Some of these phenomena have been documented empirically, while others constitute novel predictions of our model. Moreover, because our model traces all these phenomena to risk-averse arbitrageurs integrating markets for different maturities, it has novel implications on how the strength of the phenomena should vary with arbitrageurs' risk aversion.

Corollary 1. *When arbitrageurs are more risk averse (larger a), or the short rate is more volatile (larger σ_r):*

- *Forward rates under-react more strongly to changes in expected short rates.*
- *Demand has larger effects on forward rates.*
- *The FB regression coefficient γ_p of risk premia on term-structure slope, computed in Proposition 4, is larger.*
- *The regression coefficient γ_s of risk premia on the short rate, computed in Proposition 5, is smaller.*

- *The CS regression coefficient γ_r of long-rate changes on term-structure slope, computed in Proposition 6, is smaller if $\kappa_r^* < 5.27\kappa_r$.*

Corollary 1 is a comparative-statics result in our model because the parameters a and σ_r are assumed constant over time. Stepping outside of the model, however, we can interpret the corollary as concerning the effects of time-variation in a and σ_r. If, for example, a is decreasing in arbitrageurs' wealth, then time-variation in a could be measured by arbitrageurs' returns. Empirical proxies for the latter are the returns of hedge funds or the profit-loss positions of proprietary-trading desks. But our model suggests an even more direct proxy (in the sense of requiring only term-structure data), derived from arbitrageurs' trading strategies. For example, at times when the term structure is upward sloping, our model predicts that arbitrageurs are engaged in the carry trade. Therefore, arbitrageurs' wealth decreases when the carry trade loses money, i.e., when r_t increases. Conversely, when the term structure is downward sloping, arbitrageurs are engaged in the reverse-carry trade. Therefore, their wealth decreases when the reverse-carry trade loses money, i.e., when r_t decreases.

5 Two-Factor Model

This section considers the case where investor demand, represented by the parameter β_t, is an additional risk factor to the short rate r_t. The two-factor model confirms intuitions derived from the one-factor model and generates several new results.

5.1 Equilibrium

We conjecture that equilibrium bond yields are affine in (r_t, β_t), i.e.,

$$P_{t,\tau} = e^{-\left[A_r(\tau) r_t + A_\beta(\tau)\beta_t + C(\tau)\right]} \tag{29}$$

for functions $A_r(\tau), A_\beta(\tau), C(\tau)$ that depend on maturity τ. Proceeding as in the one-factor model, we find that instantaneous bond returns are

$$\frac{dP_{t,\tau}}{P_{t,\tau}} = \mu_{t,\tau} dt - \sum_{z=r,\beta} A_z(\tau)\sigma_z dB_{z,t}, \tag{30}$$

where

$$\mu_{t,\tau} \equiv \sum_{z=r,\beta} \left[A_z'(\tau) z_t - A_z(\tau)\kappa_r(\overline{z} - z_t) + \frac{1}{2} A_z(\tau)^2 \sigma_z^2 \right] + C'(\tau). \tag{31}$$

The arbitrageurs' first-order condition is

$$\mu_{t,\tau} - r_t = \sum_{z=r,\beta} A_z(\tau)\lambda_z, \tag{32}$$

where

$$\lambda_z \equiv a\sigma_z^2 \int_0^T x_{t,\tau} A_z(\tau) d\tau. \tag{33}$$

Eq. (32) is the counterpart of (10) in the context of the two-factor model. It relates a bond's expected excess return $\mu_{t,\tau} - r_t$ to the bond's risk. Risk is measured by the bond's loadings on the factors r_t and β_t. The loading on factor $z_t \in \{r_t, \beta_t\}$ is $A_z(\tau)$ from (30), and it is multiplied by the factor risk premium λ_z. The premium λ_z depends on arbitrageurs' bond holdings, and these can be computed by generalizing the market-clearing condition (12) to

$$x_{t,\tau} = -y_{t,\tau} = \alpha(\tau)\tau(\beta - R_{t,\tau}) = \alpha(\tau) \left[\beta\tau - [A_r(\tau) r_t + A_\beta(\tau)\beta_t + C(\tau)] \right]. \tag{34}$$

Substituting (31), (33) and (34) into (32), we find an affine equation in r_t and β_t. Setting constant and linear terms to zero, yields three linear ODEs in $A_r(\tau), A_\beta(\tau), C(\tau)$. Two ODEs constitute a system in $A_r(\tau), A_\beta(\tau)$, and third determines $C(\tau)$ given $A_r(\tau), A_\beta(\tau)$.

Solving for $A_r(\tau), A_\beta(\tau)$ involves a fixed-point problem. Namely, the risk-neutral measure, implicit in the coefficients of the linear ODEs, depends on arbitrageurs' bond holdings, but the latter depend on $A_r(\tau), A_\beta(\tau)$ because they depend on bond prices. More precisely, the 2×2 matrix of coefficients of $A_r(\tau), A_\beta(\tau)$ in the two ODEs involves integrals of $A_r(\tau), A_\beta(\tau)$. This makes it difficult to compute eigenvalues and eigenvectors of the matrix. In the one-factor model, the matrix has only one element and the fixed-point problem reduces to one non-linear (scalar) equation. But with two factors, the problem involves four equations. Solving the equations numerically is simple, but proving analytical results is more difficult. So we leave proofs to a future revision of this paper, and focus below on an extended numerical example. This example is not meant to be a full-fledged calibration; it conveys, however, many intuitions on how demand risk affects the term structure.

To construct the example, we need to pick values for the exogenous parameters $(T, \kappa_r, \sigma_r, \kappa_\beta, \sigma_\beta, a)$ and the function $\alpha(\tau)$. We set $T = 30$, meaning that bonds' maturities go up to 30 years. We set

$\kappa_r = 0.15$ and $\sigma_r = 0.02$, meaning that shocks to the short rate have a half-life of $\log(2)/0.15 = 4.62$ years and a yearly standard deviation 2%. These numbers are consistent with the estimates $\kappa_r = 0.178$ and $\sigma_r = 0.02$ in Chan, Karolyi, Longstaff and Sanders (1992). We set $\kappa_\beta = 0.15$ so that shocks to β_t have the same half-life as shocks to r_t. Moreover, since β_t is an interest rate (the level of the term structure in the absence of arbitrageurs), we assume that σ_β is comparable in size to σ_r and set $\sigma_\beta = 0.02$. We assume that the function $\alpha(\tau)$ measuring the demand elasticity for the bond with maturity τ decays exponentially with τ, so that bonds with longer maturities have smaller demand in present-value terms. We set the decay parameter to 0.1, meaning that $\alpha(\tau) = \alpha e^{-0.1\tau}$ for a constant α. Eqs. (33) and (34) imply that $\alpha(\tau)$ matters only through its product with the arbitrageurs' risk-aversion parameter a. We set $a\alpha = 1.5$ and discuss below the sensitivity of our results to this parameter.

5.2 Term-Structure Movements

In the context of the numerical example, we can examine how shocks to short-rate expectations and investor demand affect the term structure. Figure 1 shows the effect of short-rate expectations.

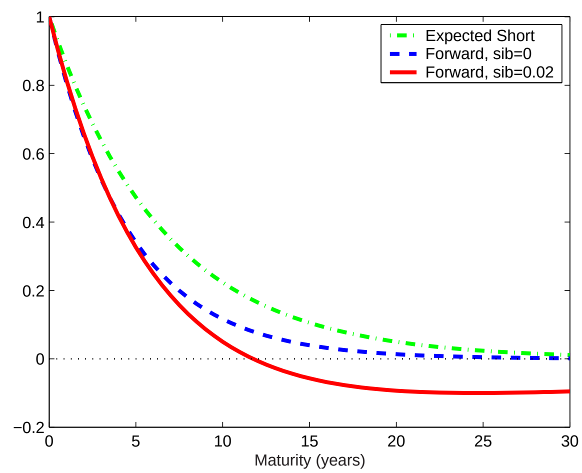

Figure 1: Effect of a unit increase in the short rate r_t on the term structure of expected short rates and instantaneous forward rates. The dashed-dotted line represents the effect on expected short rates. The dashed line represents the effect on instantaneous forward rates when $\sigma_\beta = 0$, and the solid line represents the effect when $\sigma_\beta = 0.02$.

The dashed-dotted line shows how a unit increase in the short rate r_t affects the term structure of expected short rates. The effect decays exponentially with maturity, at a rate equal to the short rate's mean-reversion parameter κ_r. The dashed line shows how the increase in r_t affects instantaneous forward rates in the absence of demand risk ($\sigma_\beta = 0$). Consistent with Proposition

2, forward rates under-react relative to expected short rates since the dashed line is below the dashed-dotted line. Moreover, under-reaction is more severe for longer maturities since the ratio of the dashed to the dashed-dotted line decreases with maturity.

The solid line shows how the increase in r_t affects instantaneous forward rates in the presence of demand risk ($\sigma_\beta = 0.02$). This line coincides approximately with the dashed line for short maturities, but is significantly below for longer maturities. Moreover, it becomes negative, meaning that an increase in expected short rates can lower long-maturity forward rates. This surprising effect is driven by arbitrageurs' hedging activity. Indeed, arbitrageurs incorporate the increase in short rates into bond prices by shorting bonds. Their shorting activity is larger for bonds of shorter maturities because the change in expected short rates is more pronounced for those maturities. At the same time, shorting activity exposes arbitrageurs to the risk that bond prices increase, either because the short rate decreases (r-risk) or because investor demand increases (β-risk). To hedge demand risk, arbitrageurs buy long-maturity bonds, which are particularly sensitive to that risk. The hedging activity of arbitrageurs on long-maturity bonds dominates their shorting activity, leading to the increase in long-maturity forward rates. In other words, arbitrageurs buy long-maturity bonds to hedge the demand risk of their short positions in bonds of shorter maturities.

Consider next the effect of investor demand. Figure 2 shows how a unit increase in β_t, i.e., a decrease in demand, impacts the term structure of instantaneous forward rates. The dashed line corresponds to the case of no demand risk (where the demand shock should be interpreted as an one-off) and the solid line to the case of demand risk. In both cases, the increase in β_t has a hump-shaped effect, impacting intermediate maturities the most. That the effect is not increasing with maturity does not contradict Proposition 3 because the increase in β_t is mean-reverting while it is permanent in Proposition 3. Mean-reversion mitigates the effect of demand for long maturities: arbitrageurs expect demand to increase back to its normal level before those maturities and, therefore, they buy long-maturity bonds more aggressively. The effect of mean-reversion is weaker in the presence of demand risk because arbitrageurs are deterred from taking aggressive positions in long-maturity bonds. Thus, the hump in the solid line is to the right relative to the dashed line.

Figures 1 and 2 make precise the intuition derived from the one-factor model that short-rate expectations are the main driver of the term structure for short maturities while demand is for long maturities. Indeed, the two-factor model allows for shocks both to expectations (r-shocks) and demand (β-shocks), and both types of shocks have the same standard deviation and mean-reversion rate in the numerical example. Figures 1 and 2 confirm that the relative impact of r- to β-shocks (ratio of the solid lines) is larger for short maturities. These figures also illustrate

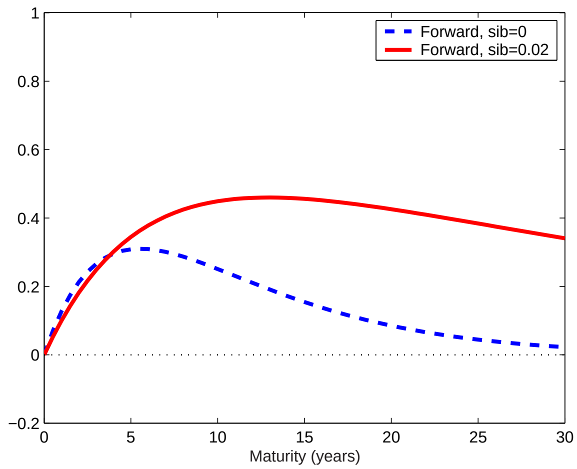

Figure 2: Effect of a unit increase in β_t (decrease in demand) on the term structure of instantaneous forward rates. The dashed line corresponds to the case $\sigma_\beta = 0$, and the solid line to the case $\sigma_\beta = 0.02$.

the effect of demand risk. In particular, Figure 2 shows that β-shocks have larger impact in the presence of demand risk, i.e., when arbitrageurs expect additional β-shocks in the future.

We next draw the implications of our model for principal-component analyses of the term structure. Litterman and Scheinkman (LS 1991) decompose bond returns into principal components (PC) and show that the first PC explains 89.5% of return variation. Thus, returns are highly correlated across maturities, and a common interpretation of this finding is that one main economic factor drives returns. This seems to contradict our two-factor model, at least in the context of the numerical example. Indeed, returns in our model are driven by independent r- and β-shocks. Moreover, Figures 1 and 2 show that the relative impact of the two types of shocks varies significantly with maturity.[11]

To map our model to the empirical findings, we perform the LS analysis with instantaneous returns, considering bonds with maturities from six months to 30 years in six-month increments. Eq. (30) shows that the loadings of a bond with maturity τ on r- and β-shocks are $A_r(\tau)$ and $A_\beta(\tau)$, respectively. Since r- and β-shocks are independent, the instantaneous covariance matrix associated to the vector of 60 returns is

$$\sigma_r^2 A_r A_r' + \sigma_\beta^2 A_\beta A_\beta', \tag{35}$$

where A_z is the column vector with elements $\{A_z(\tau)\}_{\tau=0.5,..,30}$ for $z = r, \beta$. The principal compo-

[11] Figures 1 and 2 concern forward rates, but the effects are similar for bond returns. For example, r-shocks explain 99.92% of the variance of the one-year bond's instantaneous return, 92.20% for the five-year bond, 65.22% for the ten-year bond, 18.50% for the twenty-year bond, and 4.79% for the thirty-year bond.

nents are the eigenvectors of this matrix and the fraction of variation they explain is the ratio of the corresponding eigenvalue to the sum of all eigenvalues. Since in our model returns are driven by two factors, the matrix (35) has two non-zero eigenvalues. The largest one is 94.61% of the sum, meaning that the first PC explains 94.61% of return variation. Thus, the mere fact that one PC explains a large fraction of term-structure movements does not mean that movements are caused by one main economic factor. Our result also cautions against giving any economic interpretation to PCs. For example, the first PC in the numerical example is $0.44A_r + A_\beta$ (up to normalization), meaning that the corresponding factor is $0.44r_t + \beta_t$. This is an amalgam of r- and β-shocks, and can neither be interpreted as shocks to short-rate expectations nor as shocks to demand.

Figure 3 draws the first PC in yield space, together with the effects of r- and β-shocks (which are equal to $A_r(\tau)/\tau$ and $A_\beta(\tau)/\tau$, respectively, from (1) and (7)). The first PC is almost flat with maturity but this obscures the stark difference between r- and β-shocks: the effect of r-shocks is decreasing with maturity, while that of β-shocks is increasing. But because the first PC is an amalgam of the two types of shocks, it turns out to be approximately flat.

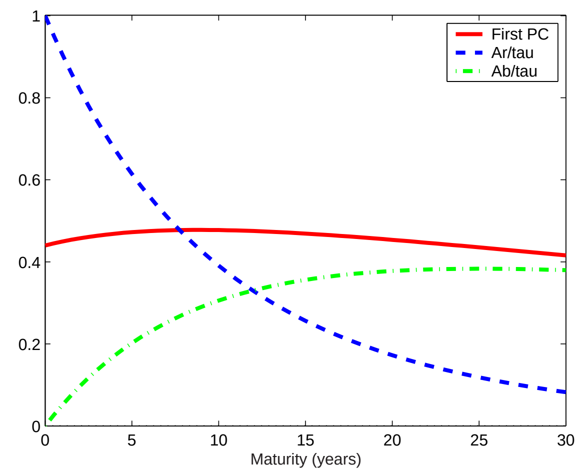

Figure 3: The solid line represents the first principal component of bond returns, plotted in yield space and normalized. The dashed line represents the effect of an r-shock on yields and the dashed-dotted line represents the effect of a β-shock.

The flat shape of the first PC in the numerical example is consistent with LS. Our model, however, does not always generate a flat shape. For example, the first PC is decreasing with maturity when $\sigma_\beta = 0$ because there are only r-shocks, and increasing when $\sigma_\beta = 0.04$ because β-shocks dominate. Similar comparative statics hold when αa varies because demand shocks have larger impact when arbitrageurs are more risk-averse. Therefore, the robust implication of our analysis is not that the first PC is flat, but that it explains a large fraction of variation. For example, the fraction explained for $\sigma_\beta = 0$ or $\sigma_\beta = 0.04$ is larger than for $\sigma_\beta = 0.02$ because in the first two cases one type of shock is dominant.

5.3 Risk Premia and Predictability

We next examine how demand risk affects the relationship between term-structure slope, bond risk premia and spot-rate changes. Consider the Fama-Bliss (FB) regression (24) with dependent variable the excess return of a τ-year zero-coupon bond over an one-year interval ($\Delta\tau = 1$), and independent variable the difference between the one-year forward rate starting in year $\tau - 1$ and the one-year spot rate. Figure 4 plots the regression coefficient γ_p for all maturities $\tau > 1$. The figure shows that demand risk reinforces the positive relationship between premia and slope, especially for long maturities. The intuition is that the term structure is upward sloping at times when the short rate r_t is low or bond demand is low (high β_t). In both cases arbitrageurs hold long positions in bonds and premia are positive. Moreover, demand risk renders long-maturity bonds particularly risky, thus raising the premia that arbitrageurs require for holding them. This strengthens the relationship between premia and slope for long maturities.

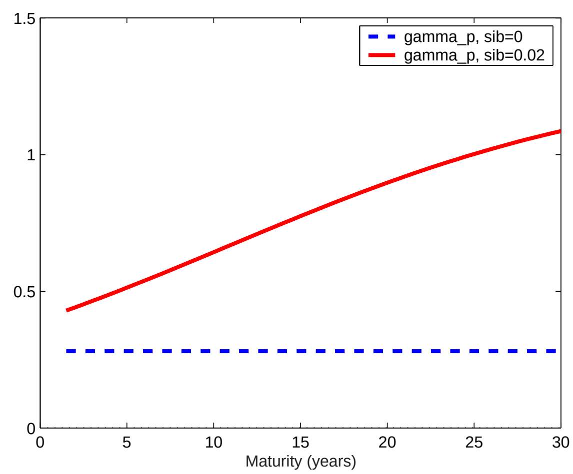

Figure 4: The coefficient γ_p of the Fama-Bliss regression (24). The dependent variable is the excess return of a τ-year zero-coupon bond over an one-year interval ($\Delta\tau = 1$). The independent variable is the difference between the one-year forward rate starting in year $\tau - 1$ and the one-year spot rate. The coefficient γ_p is plotted for all maturities $\tau > 1$. The dashed line corresponds to the case $\sigma_\beta = 0$, and the solid line to the case $\sigma_\beta = 0.02$.

Consider next the Campbell-Shiller (CS) regression (27) with dependent variable the change in yield of a τ-year zero-coupon bond over an one-year interval ($\Delta\tau = 1$), and independent variable a normalized difference between the τ- and the one-year spot rate. The normalization is so that the regression coefficient γ_r is equal to one under the expectations hypothesis (EH). Figure 5 plots γ_r for all maturities $\tau > 1$. The figure shows that demand risk renders γ_r negative and decreasing with maturity. The intuition is that demand risk generates strongly time-varying risk premia for long-maturity bonds. Suppose, for example, that the term structure is upward sloping. If the positive slope is due to the short rate r_t being low, then long yields are likely to increase as the short rate mean-reverts. But if the positive slope is due to bond demand being low (high β_t), then long yields are likely to decrease as demand mean-reverts. In other words, the low demand generates large positive premia, which are expected to decrease as demand decreases.

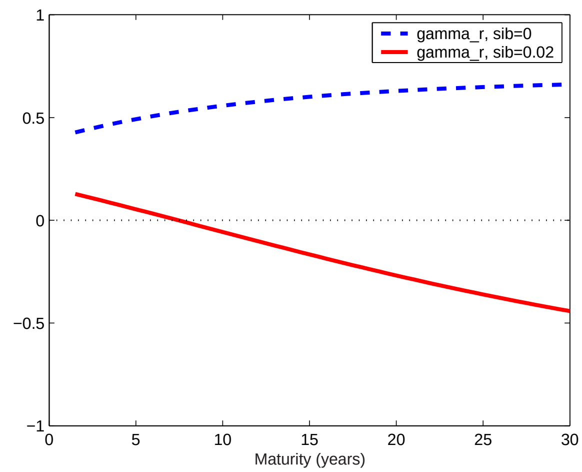

Figure 5: The coefficient γ_r of the Campbell-Shiller regression (27). The dependent variable is the change in yield of a τ-year zero-coupon bond over an one-year interval ($\Delta\tau = 1$). The independent variable is the difference between the τ- and the one-year spot rate, normalized so that the regression coefficient γ_r is equal to one under the expectations hypothesis. The coefficient γ_r is plotted for all maturities $\tau > 1$. The dashed line corresponds to the case $\sigma_\beta = 0$, and the solid line to the case $\sigma_\beta = 0.02$.

How do the coefficients in Figures 4 and 5 compare to their empirical counterparts? FB compute γ_p for maturities $\tau = 2, 3, 4, 5$ and find $\gamma_p \approx 1$. CS compute γ_r for the same maturities and find that it decreases from about -1 to -2. For those maturities, Figures 4 and 5 imply $\gamma_p \approx 0.5$ and $\gamma_r \approx 0$. In both cases, our model implies weaker violations of the EH than in the data, but the numbers are within the FB and CS confidence intervals. Note that demand risk reduces the discrepancy between our results and the empirical estimates: it raises the FB coefficients, and renders the CS coefficients negative and decreasing with maturity.

The coefficients γ_p and γ_r can be made closer to their empirical counterparts by choosing larger values of demand risk σ_β or arbitrageur risk aversion a. For such values, however, the first PC tends to become non-flat because long yields become more variable than short yields. We leave a more complete analysis of this issue, as well as a more detailed calibration of the two-factor model, to a future revision of this paper.

6 Conclusion

SECTION TO BE WRITTEN

A Foundation for Investors' Demand Function

SECTION TO BE WRITTEN

B Proofs of Propositions 1-6 and Corollary 1

Proof of Proposition 1: Substituting (9), (11) and (12) into (10), we find

$$A_r'(\tau)r_t + C'(\tau) - A_r(\tau)\kappa_r(\overline{r} - r_t) + \frac{1}{2}A_r(\tau)^2\sigma_r^2 - r_t$$
$$= A_r(\tau)a\sigma_r^2 \int_0^T \alpha(\tau)\left[\beta\tau - [A_r(\tau)r_t + C(\tau)]\right] A_r(\tau)d\tau. \tag{B.1}$$

This equation is affine in r_t. Setting the linear terms to zero, we find the ODE

$$A_r'(\tau) + \kappa_r A_r(\tau) - 1 = -a\sigma_r^2 A_r(\tau) \int_0^T \alpha(\tau)A_r(\tau)^2 d\tau, \tag{B.2}$$

and setting the constant terms to zero, we find the ODE

$$C'(\tau) - \kappa_r\overline{r}A_r(\tau) + \frac{1}{2}\sigma_r^2 A_r(\tau)^2 = a\sigma_r^2 A_r(\tau) \int_0^T \alpha(\tau)\left[\beta\tau - C(\tau)\right] A_r(\tau)d\tau. \tag{B.3}$$

These ODEs must be solved with the initial conditions $A_r(0) = C(0) = 0$. The solution to (B.2) is (13), provided that κ_r^* is a solution to (15). Eq. (15) has a unique solution because the right-hand side is decreasing in κ_r^* and is equal to zero for $\kappa_r^* = \infty$. The solution to (B.3) is

$$C(\tau) = z\int_0^\tau A_r(u)du - \frac{\sigma_r^2}{2}\int_0^\tau A_r(u)^2 du, \tag{B.4}$$

where

$$z \equiv \kappa_r\overline{r} + a\sigma_r^2 \int_0^T \alpha(\tau)\left[\beta\tau - C(\tau)\right] A_r(\tau)d\tau. \tag{B.5}$$

Substituting $C(\tau)$ from (B.4) into (B.5), we find

$$z = \kappa_r\overline{r} + a\sigma_r^2\beta\int_0^T \alpha(\tau)\tau A_r(\tau)d\tau - a\sigma_r^2 z\int_0^T \alpha(\tau)\left[\int_0^\tau A_r(u)du\right]A_r(\tau)d\tau$$
$$+\frac{a\sigma_r^4}{2}\int_0^T \alpha(\tau)\left[\int_0^\tau A_r(u)^2du\right]A_r(\tau)d\tau$$
$$\Rightarrow z = \frac{\kappa_r\overline{r} + a\sigma_r^2\beta\int_0^T \alpha(\tau)\tau A_r(\tau)d\tau + \frac{a\sigma_r^4}{2}\int_0^T \alpha(\tau)\left[\int_0^\tau A_r(u)^2du\right]A_r(\tau)d\tau}{1 + a\sigma_r^2\int_0^T \alpha(\tau)\left[\int_0^\tau A_r(u)du\right]A_r(\tau)d\tau}.$$

The function $C(\tau)$ coincides with (14) if $z = \kappa_r^*\overline{r}^*$. Eqs. (16)-(18) imply that $z = \kappa_r^*\overline{r}^*$ if

$$\frac{\kappa_r + a\sigma_r^2\int_0^T \alpha(\tau)\tau A_r(\tau)d\tau}{1 + a\sigma_r^2\int_0^T \alpha(\tau)\left[\int_0^\tau A_r(u)du\right]A_r(\tau)d\tau} = \kappa_r^*. \tag{B.6}$$

Eq. (B.6) follows from

$$\kappa_r + a\sigma_r^2\int_0^T \alpha(\tau)\tau A_r(\tau)d\tau$$
$$= \kappa_r^* + a\sigma_r^2\int_0^T \alpha(\tau)\left[\tau - A_r(\tau)\right]A_r(\tau)d\tau$$
$$= \kappa_r^* + \kappa_r^* a\sigma_r^2\int_0^T \alpha(\tau)\left[\int_0^\tau A_r(u)du\right]A_r(\tau)d\tau,$$

where the first step follows from (15) and the second from (13).

To show that the risk-neutral dynamics are given by (19), we rewrite (B.1) as

$$A_r'(\tau)r_t + C'(\tau) - A_r(\tau)\kappa_r(\overline{r} - r_t) + \frac{1}{2}A_r(\tau)^2\sigma_r^2 - r_t$$
$$= A_r(\tau)a\sigma_r^2\left[\int_0^T \alpha(\tau)\left[\beta\tau - C(\tau)\right]A_r(\tau)d\tau - r_t\int_0^T \alpha(\tau)A_r(\tau)^2d\tau\right]$$
$$\Leftrightarrow A_r'(\tau)r_t + C'(\tau) - A_r(\tau)\kappa_r(\overline{r} - r_t) + \frac{1}{2}A_r(\tau)^2\sigma_r^2 - r_t = A_r(\tau)\left[z - \kappa_r\overline{r} - r_t(\kappa_r^* - \kappa_r)\right]$$
$$\Leftrightarrow A_r'(\tau)r_t + C'(\tau) - A_r(\tau)\kappa_r^*(\overline{r}^* - r_t) + \frac{1}{2}A_r(\tau)^2\sigma_r^2 - r_t = 0, \tag{B.7}$$

where the second step follows from (15) and (B.5), and the third from $z = \kappa_r^*\overline{r}^*$. The risk-neutral dynamics are given by (19) because of (B.7). ∎

Proof of Proposition 2: Eqs. (7) and (21) imply that

$$f_{t,\tau} = A_r'(\tau) r_t + C'(\tau). \tag{B.8}$$

Therefore,

$$\frac{\partial f_{t,\tau}}{\partial r_t} = A_r'(\tau) = e^{-\kappa_r^* \tau}, \tag{B.9}$$

where the second step follows from (13). On the other hand, (2) implies that

$$E_t(r_{t+\tau}) = (1 - e^{-\kappa_r \tau})\overline{r} + e^{-\kappa_r \tau} r_t. \tag{B.10}$$

Therefore,

$$\frac{\partial E_t(r_{t+\tau})}{\partial r_t} = e^{-\kappa_r \tau}. \tag{B.11}$$

The claims in the proposition follow from (B.9), (B.11), and $\kappa_r^* > \kappa_r$.

We next prove the claim made in the text that (22) implies (23). We set $T \equiv t + \tau$ and treat T as fixed when t varies. We first show that

$$E_t^*(r_{t+\tau}) = A_r'(\tau) r_t + \hat{C}'(\tau), \tag{B.12}$$

where $\hat{C}(\tau)$ solves the ODE

$$\hat{C}'(\tau) - \kappa_r \overline{r} A_r(\tau) = a\sigma_r^2 A_r(\tau) \int_0^T \alpha(\tau) \left[\beta\tau - C(\tau)\right] A_r(\tau) d\tau \tag{B.13}$$

with the initial condition $\hat{C}(0) = 0$. (Eq. (B.13) differs from (B.3) because of the missing term $\sigma_r^2 A_r(\tau)^2/2$.) Since $A_r'(0) = 1$ and $\hat{C}'(0) = 0$, (B.12) holds if the process $A_r'(\tau) r_t + \hat{C}'(\tau)$ is a martingale under the risk-neutral measure. Given the risk-neutral dynamics (19), the martingale property holds if

$$-A_r''(\tau) r_t - \hat{C}''(\tau) + A_r'(\tau)\kappa_r^*(\overline{r}^* - r_t) = 0. \tag{B.14}$$

Multiplying (B.2)by r_t and adding to (B.13), we find

$$A_r'(\tau) r_t + \hat{C}'(\tau) - A_r(\tau)\kappa_r(\overline{r} - r_t) - r_t$$
$$= A_r(\tau) a\sigma_r^2 \left[\int_0^T \alpha(\tau) \left[\beta\tau - C(\tau)\right] A_r(\tau) d\tau - r_t \int_0^T \alpha(\tau) A_r(\tau)^2 d\tau\right].$$

Following the same steps as when deriving (B.7), we find

$$A_r'(\tau)r_t + C'(\tau) - A_r(\tau)\kappa_r^*(\overline{r}^* - r_t) - r_t = 0. \tag{B.15}$$

Eq. (B.14) follows by differentiating (B.15) with respect to τ.

We next show that

$$E_t^*\left[e^{-\int_t^{t+\tau} r_u du} r_{t+\tau}\right] = e^{-[A_r(\tau)r_t + C(\tau)]}\left[E_t^*(r_{t+\tau}) - \frac{1}{2}A_r(\tau)^2\sigma_r^2\right]. \tag{B.16}$$

Since $A_r(0) = C(0) = 0$, (B.16) holds if the process

$$e^{-\int_0^t r_u du - [A_r(\tau)r_t + C(\tau)]}\left[E_t^*(r_{t+\tau}) - \frac{1}{2}A_r(\tau)^2\sigma_r^2\right]$$

is a martingale under the risk-neutral measure. Substituting the diffusion of $E_t^*(r_{t+\tau})$ from (B.12) and using (B.7), we find that the martingale property holds. Combining (B.16) with (22) and

$$E_t^*\left[e^{-\int_t^{t+\tau} r_u du}\right] = e^{-[A_r(\tau)r_t + C(\tau)]}$$

we find (23). ∎

Proof of Proposition 3: Eqs. (13)-(16) and (B.8) imply that

$$\frac{\partial f_{t,\tau}}{\partial \beta} = \frac{\partial C'(\tau)}{\partial \beta} = \frac{\partial \overline{r}^*}{\partial \beta}(1 - e^{-\kappa_r^*\tau}) = \frac{z_\beta}{\kappa_r^*}(1 - e^{-\kappa_r^*\tau}). \tag{B.17}$$

The claims in the proposition follow from (B.17), $z_\beta > 0$, and $z_\beta/\kappa_r^* < 1$. (The latter inequality follows from (17) and (B.6)). ∎

Proof of Proposition 4: Eqs. (1), (7) and (20) imply that the dependent variable in (24) is

$$\frac{1}{\Delta\tau}\left[A_r(\tau)r_t + C(\tau) - [A_r(\tau - \Delta\tau)r_{t+\Delta\tau} + C(\tau - \Delta\tau)] - [A_r(\Delta\tau)r_t + C(\Delta\tau)]\right]$$

and the independent variable is

$$\frac{1}{\Delta\tau}\left[A_r(\tau)r_t + C(\tau) - [A_r(\tau - \Delta\tau)r_t + C(\tau - \Delta\tau)] - [A_r(\Delta\tau)r_t + C(\Delta\tau)]\right].$$

Therefore, the FB regression coefficient is

$$\gamma_p = \frac{\text{Cov}\left[[A_r(\tau) - A_r(\Delta\tau)]r_t - A_r(\tau-\Delta\tau)r_{t+\Delta\tau}, [A_r(\tau) - A_r(\tau-\Delta\tau) - A_r(\Delta\tau)]r_t\right]}{\text{Var}\left[[A_r(\tau) - A_r(\tau-\Delta\tau) - A_r(\Delta\tau)]r_t\right]}. \quad \text{(B.18)}$$

Eqs. (2) and (B.10) imply that

$$\text{Cov}(r_{t+\Delta\tau}, r_t) = \text{Var}(r_t)e^{-\kappa_r\Delta\tau}.$$

Substituting into (B.18), we find

$$\gamma_p = \frac{A_r(\tau) - A_r(\tau-\Delta\tau)e^{-\kappa_r\Delta\tau} - A_r(\Delta\tau)}{A_r(\tau) - A_r(\tau-\Delta\tau) - A_r(\Delta\tau)}.$$

Taking the limit $\Delta\tau \to 0$ and noting from (13) that $A_r(\Delta\tau)/\Delta\tau \to 1$, we find

$$\gamma_p \to \frac{A_r'(\tau) + \kappa_r A_r(\tau) - 1}{A_r'(\tau) - 1} = \frac{(\kappa_r^* - \kappa_r)A_r(\tau)}{\kappa_r^* A_r(\tau)} = \frac{\kappa_r^* - \kappa_r}{\kappa_r^*},$$

where the second step follows from (B.2). ∎

Proof of Proposition 5: Proceeding as in the proof of Proposition 4, we find

$$\gamma_s = \frac{A_r(\tau) - A_r(\tau-\Delta\tau)e^{-\kappa_r\Delta\tau} - A_r(\Delta\tau)}{A_r(\Delta\tau)}.$$

Taking the limit $\Delta\tau \to 0$, we find

$$\gamma_p \to A_r'(\tau) + \kappa_r A_r(\tau) - 1 = -(\kappa_r^* - \kappa_r)A_r(\tau).$$

To show that the factor risk premium λ_r is given by (26), we note from (11) and (12) that

$$\begin{aligned}\lambda_r &= a\sigma_r^2 \int_0^T \alpha(\tau)\left[\beta\tau - [A_r(\tau)r_t + C(\tau)]\right] A_r(\tau)d\tau \\ &= z - \kappa_r\overline{r} - (\kappa_r^* - \kappa_r)r_t,\end{aligned}$$

where the second step follows from (15) and (B.5). Eq. (26) follows from $z = \kappa_r^*\overline{r}^*$. ∎

Proof of Proposition 6: Eqs. (1) and (7) imply that the dependent variable in (27) is

$$\frac{A_r(\tau-\Delta\tau)r_{t+\Delta\tau} + C(\tau-\Delta\tau)}{\tau-\Delta\tau} - \frac{A_r(\tau)r_t + C(\tau)}{\tau}$$

and the independent variable is

$$\frac{\Delta\tau}{\tau-\Delta\tau}\left[\frac{A_r(\tau)r_t+C(\tau)}{\tau}-\frac{A_r(\Delta\tau)r_t+C(\Delta\tau)}{\Delta\tau}\right].$$

Proceeding as in the proof of Proposition 4, we find

$$\gamma_r=\frac{\tau-\Delta\tau}{\Delta\tau}\frac{\frac{A_r(\tau-\Delta\tau)e^{-\kappa_r\Delta\tau}}{\tau-\Delta\tau}-\frac{A_r(\tau)}{\tau}}{\frac{A_r(\tau)}{\tau}-\frac{A_r(\Delta\tau)}{\Delta\tau}}.$$

Taking the limit $\Delta\tau\to 0$, we find

$$\gamma_r=\frac{A_r'(\tau)-\frac{A_r(\tau)}{\tau}+\kappa_r A_r(\tau)}{1-\frac{A_r(\tau)}{\tau}},$$

which coincides with (28) because of (B.2).

Eqs. (13) and (28) imply that γ_r is increasing in τ if the function

$$f(x)\equiv\frac{1-e^{-x}}{1-\frac{1-e^{-x}}{x}}$$

is decreasing in x for $x>0$. The derivative $f'(x)$ has the same sign as

$$f_1(x)\equiv -e^{\frac{x}{2}}+e^{-\frac{x}{2}}+x.$$

The function $f_1(x)$ is negative because $f_1(0)=0$ and $f_1'(x)<0$. Therefore, $f(x)$ is decreasing.

Eqs. (13) implies that for small τ, $A_r(\tau)=\tau-\frac{\kappa_r^*\tau^2}{2}+o(\tau^2)$. Substituting into (28), we find

$$\gamma_r=1-\frac{\kappa_r^*-\kappa_r}{\frac{\kappa_r^*}{2}}+o(1)=2\frac{\kappa_r}{\kappa_r^*}-1+o(1).$$

Therefore, γ_r is negative for small τ if $\kappa_r^*>2\kappa_r$. Eq. (13) implies that when $\tau\to\infty$, $A_r(\tau)\to 1/\kappa_r^*$. Substituting into (28), we find

$$\gamma_r\to 1-\frac{\kappa_r^*-\kappa_r}{\kappa_r^*}=\frac{\kappa_r}{\kappa_r^*}>0.$$

∎

Proof of Corollary 1: The key observation is that the mean-reversion parameter κ_r^* under the risk-neutral measure is an increasing function of $a\sigma_r^2$. Indeed, κ_r^* is defined implicitly by (15), whose right-hand side is decreasing in κ_r^* (because of (13)) and increasing in $a\sigma_r^2$.

To prove the first claim in the corollary, we recall from (B.9) and (B.11) that the response of forward rates to expected short rates is

$$\frac{\frac{\partial f_{t,\tau}}{\partial r_t}}{\frac{\partial E_t(r_{t+\tau})}{\partial r_t}} = e^{-(\kappa_r^* - \kappa_r)\tau}.$$

The under-reaction of forward rates is stronger for larger a and σ_r because κ_r^* is increasing in these variables.

To prove the second claim, we recall from (B.17) that the effect of demand on forward rates is

$$\frac{\partial f_{t,\tau}}{\partial \beta} = \frac{z_\beta}{\kappa_r^*}(1 - e^{-\kappa_r^* \tau}).$$

To show that the effect is stronger for larger a and σ_r, it suffices to show that z_β/κ_r^* is increasing in $a\sigma_r^2$. Eqs. (17) and (B.6) imply that

$$\begin{aligned}
\frac{z_\beta}{\kappa_r^*} &= 1 - \frac{\kappa_r}{\kappa_r^*} \frac{1}{1 + a\sigma_r^2 \int_0^T \alpha(\tau) \left[\int_0^\tau A_r(u)du\right] A_r(\tau)d\tau} \\
&= 1 - \frac{\kappa_r}{\kappa_r^* + a\sigma_r^2 \int_0^T \alpha(\tau)[\tau - A_r(\tau)]A_r(\tau)d\tau} \\
&= 1 - \frac{\kappa_r}{\kappa_r + a\sigma_r^2 \int_0^T \alpha(\tau)\tau A_r(\tau)d\tau},
\end{aligned}$$

where the second step follows from (B.2) and the third from (15). Therefore, z_β/κ_r^* is increasing in $a\sigma_r^2$ if $a\sigma_r^2 \int_0^T \alpha(\tau)\tau A_r(\tau)d\tau$ is increasing in $a\sigma_r^2$. To show the latter, it suffices from (13) to show that $a\sigma_r^2/\kappa_r^*$ is increasing in $a\sigma_r^2$. Eq. (15) implies that

$$\frac{a\sigma_r^2}{\kappa_r^*} = \frac{1 - \frac{\kappa_r}{\kappa_r^*}}{\int_0^T \alpha(\tau)A_r(\tau)^2 d\tau}.$$

The numerator in the right-hand side is increasing in $a\sigma_r^2$, and the denominator is decreasing in $a\sigma_r^2$ from (13). Therefore, $a\sigma_r^2/\kappa_r^*$ is increasing in $a\sigma_r^2$.

The third and fourth claim follow because the functions $(\kappa_r^* - \kappa_r)/\kappa_r^*$ and $(\kappa_r^* - \kappa_r)A_r(\tau)$ are increasing in κ_r^*. To prove the fifth claim, we note from (13) and (28) that γ_r is smaller for larger a and σ_r if the function

$$g(\kappa_r^*) \equiv \frac{\kappa_r^* - \kappa_r}{\frac{\kappa_r^*}{1 - e^{-\kappa_r^*\tau}} - \frac{1}{\tau}}$$

is increasing in κ_r^*. The derivative $g'(\kappa_r^*)$ is positive if $\kappa_r^*/\kappa_r < g_1(\kappa_r^*\tau)$, where

$$g_1(x) \equiv x \frac{1 - e^{-x} - xe^{-x}}{(1 - e^{-x})^2 - x^2 e^{-x}}.$$

The function $g_1(x)$ is minimized for $x = 3.1676$ and the minimum value is 5.2752. Therefore, the function $g(\kappa_r^*)$ is increasing in κ_r^* if $\kappa_r^* < 5.27\kappa_r$. $\blacksquare$

References

Barberis, N. and A. Shleifer, 2003, "Style Investing," *Journal of Financial Economics,* 68, 161-199.

Bates, D., 2006, "The Market for Crash Risk," working paper, University of Iowa.

Bollen, N., and R. Whaley, 2004, "Does Net Buying Pressure Affect the Shape of Implied Volatility Functions?," *Journal of Finance,* 59, 711753.

Buraschi, A. and A. Jiltsov, 2007, "Term Structure of Interest Rates Implications of Habit Persistence," *Journal of Finance*, forthcoming.

Chan, K., G. Karolyi, F. Longstaff, and A. Sanders, 1992, "An Empirical Comparison of Alternative Models of the Short-Term Interest Rate," *Journal of Finance,* 47, 1209-1227.

Cochrane, J., 1999, "New Facts in Finance," *Economic Perspectives* (Federal Reserve Bank of Chicago), 23, 36-58.

Cox, J., J. Ingersoll, and S. Ross, 1985, "A Theory of the Term Structure of Interest Rates," *Econometrica,* 53, 385-408.

Culbertson, J., 1957, "The Term Structure of Interest Rates," *Quarterly Journal of Economics,* 71, 485-517.

De Long, B., A. Shleifer, L. Summers, and R. Waldmann, 1990, "Noise Trader Risk in Financial Markets," *Journal of Political Economy,* 98, 703-738.

Dai, Q. and K. Singleton, 2002, " Expectations Puzzles, Time-Varying Risk Premia, and Affine Models of the Term Structure," *Journal of Financial Economics,* 63, 415-441.

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

Fama, E. and R. Bliss, 1987, "The Information in Long-Maturity Forward Rates," *American Economic Review,* 77, 680-692.

Gabaix, X., A. Krishnamurthy, and O. Vigneron, 2007, "Limits of Arbitrage: Theory and Evidence from the Mortgage Backed Securities Market," *Journal of Finance,* forthcoming.

Garleanu, N., L. Pedersen, and A. Poteshman, 2006, "Demand-Based Option Pricing," working paper, University of Pennsylvania.

Gromb, D. and D. Vayanos, 2002, "Equilibrium and Welfare in Markets with Financially Constrained Arbitrageurs," *Journal of Financial Economics,* 66, 361407.

Kyle, A., and W. Xiong, 2001, "Contagion as a Wealth Effect of Financial Intermediaries," *Journal of Finance* 56, 1401-1440.

Litterman, R. and J. Scheinkman, 1991, "Common Factors Affecting Bond Returns," *Journal of Fixed Income,* 1, 54-61.

Modigliani, F. and R. Sutch, 1966, "Innovations in Interest-Rate Policy," *American Economic Review,* 56, 178-197.

Pavlova, A. and R. Rigobon, 2005, "Wealth Transfers, Contagion, and Portfolio Constraints," NBER working paper 11440.

Spiegel, M., 1998, "Stock Price Volatility in a Multiple Security Overlapping Generations Model," *Review of Financial Studies,* 11, 419-447.

Vasicek, O., 1977, "An Equilibrium Characterization of the Term Structure," *Journal of Financial Economics,* 5, 177-188.

Wachter, J., 2006, "A Consumption-Based Model of the Term Structure of Interest Rates," *Journal of Financial Economics,* 79, 365-399.

Xiong, W. and H. Yan, 2006, "Heterogeneous Expectations and Bond Markets," working paper, Princeton University.