Mortgage Risk and the Yield Curve*

Aytek Malkhozov
McGill[†]

Philippe Mueller
LSE[‡]

Andrea Vedolin
LSE[§]

Gyuri Venter
CBS[¶]

Abstract

We study the feedback from the risk of outstanding mortgage-backed securities (MBS) on the level and volatility of interest rates. We incorporate the supply shocks resulting from changes in MBS duration into a stylized equilibrium dynamic term structure model and derive two predictions that are strongly supported in the data: (i) MBS duration positively predicts excess bond returns, especially for longer maturities; (ii) MBS convexity increases interest rate volatility, and this effect has a hump-shaped term structure. Empirically, the predictive power of MBS duration and convexity is not subsumed by information in the cross section of Treasury yields.

First Version: August 2012
This Version: 2 June, 2014

*We would like to thank Caio Almeida, Michael Bauer, Ruslan Bikbov, Mike Chernov, Pierre Collin-Dufresne, Jefferson Duarte, Greg Duffee, Mathieu Fournier, András Fülöp, Christian Julliard, Anh Le, Neil Pearson, Douglas McManus, Marcel Rindisbacher, Dimitri Vayanos, Paul Whelan, and seminar and conference participants at the Bank of England, Banque de France, CBS, Deutsche Bank, Duke University, Erasmus University, Exeter University, Imperial Business School, LSE, Manchester Business School, the Oxford-MAN Institute, Toulouse School of Economics, Luxembourg School of Finance, University of Nottingham, SAC Capital Advisors, University of Bern, the University of Bonn, the 2012 Arne Ryde Workshop for Financial Economics, the 2012 IFSID Conference on Structured Products and Derivatives, the 2012 Young Scholars Nordic Finance Workshop, the 2012 Junior Faculty Research Roundtable at UNC, the 2012 Conference on Advances in the Analysis of Hedge Fund Strategies, the 2013 Mathematical Finance Days, the 2013 Bank of Canada Fixed Income conference, the 2013 Financial Econometrics conference in Toulouse, the 2013 SFS Cavalcade, the 2013 Western Finance Meeting, and the UNC/Atlanta Fed Conference on Housing for thoughtful comments. Malkhozov gratefully acknowledges financial support from the Institut de Finance Mathématique de Montréal. Mueller and Vedolin gratefully acknowledge financial support from the LSE Research Committee Seed Fund. Venter gratefully acknowledges financial support from the Center for Financial Frictions (FRIC), grant no. DNRF102. All authors thank the Paul Woolley Centre at the LSE and the Dauphine-Amundi Chair in Asset Management for financial support.

[†]Desautels Faculty of Management, Email: aytek.malkhozov@mcgill.ca
[‡]Department of Finance, Email: p.mueller@lse.ac.uk
[§]Department of Finance, Email: a.vedolin@lse.ac.uk
[¶]Department of Finance, Email: gv.fi@cbs.dk

Mortgage-backed securities (MBS) and, more generally, mortgage loans expose investors to interest rate risk. Unlike with regular bonds, this exposure can change with interest rate conditions.[1] This is the case because mortgages typically feature an embedded prepayment option that makes their convexity negative: Lower interest rates increase the probability that outstanding mortgages will be prepaid in the future and thereby considerably decrease their duration. This leaves financial institutions who invest in MBS short of duration exposure, until either interest rates revert back to higher levels, or the prepayment option becomes sufficiently in the money for a large number of households to refinance and take on new mortgages. Because households do not play an active role in bond markets and do not hedge their time-varying interest rate risk exposure, it is the position of financial institutions that determines the pricing of interest rate risk (see Gabaix, Krishnamurthy, and Vigneron (2007)). In other words, a fall in mortgage duration is equivalent to a negative transitory shock to the supply of long-term bonds and therefore can have an effect on their prices.[2] Moreover, mortgage investors who want to keep the duration of their portfolios constant after a drop in MBS duration (for hedging or portfolio rebalancing reasons) induce additional buying pressure on Treasuries and push interest rates further down. Thus, negative convexity due to the prepayment option creates an amplification channel for interest rate shocks.

The negative convexity channel described above has attracted the attention of practitioners, policy makers, and empirical researchers alike.[3] In this paper we build a parsimonious model that formalizes the intuition behind this channel and allows us to derive novel predictions about the effect of mortgage risk on the yield curve. The model implies a link between mortgage convexity and the volatility of interest rates and between mortgage duration and bond excess returns. We empirically test the predictions of the model and find strong support for them in the data.

[1] In contrast with Treasuries, MBS duration can drop by more than 50%, e.g., from 5 to 2.5 years in a short period (see also Figure 1). Taking into account the value of outstanding mortgage debt, we calculate that one monthly standard deviation shock to MBS duration is a dollar duration equivalent of a USD 368bn shock to the supply of 10-year Treasuries.

[2] See Lou, Yan, and Zhang (2013) on how financial intermediaries' supply shocks can affect Treasury bond prices.

[3] See Perli and Sack (2003), Chang, McManus, and Ramagopal (2005), Duarte (2008), Hanson (2013), and Li and Wei (2013) among others.

In our model the term structure of interest rates is determined by the interaction between (i) exogenous shocks to the short rate and (ii) changes in the net supply of long-term bonds that are endogenously driven by the interest rate risk exposure of mortgages. The term structure model we derive takes the form of a standard Vasicek (1977) short rate model augmented by an additional affine factor that captures the duration of outstanding MBS. This factor drives the market price of interest rate risk and affects the risk premia of long-term bonds but not the dynamics of the risk-free rate itself. Its contribution tends to zero if the risk-bearing capacity of financial institutions is high.

The model has two sets of testable predictions. First, the duration of outstanding MBS predicts bond excess returns. Moreover, this effect is stronger for longer maturity bonds. This happens because in the model the market price of interest rate risk is proportional to the quantity of duration risk that investors have to hold. Longer maturity bonds with a higher exposure to interest rate risk are more strongly affected through this channel.

Second, the average volatility of all yields is increasing in the convexity of outstanding MBS through the negative convexity amplification channel discussed above. This effect has a hump-shaped pattern across maturities with intermediate maturities being most strongly affected. In the model, supply shocks create transitory variations in the market price of interest rate risk. Short-maturity bonds are a close substitute to the short rate and are not very sensitive to variations in the market price of risk, while for long maturities, the market price of risk is expected to mean revert. As a result, the effect of negative convexity on yield volatility is strongest for intermediate maturities. The same intuition applies to swaption implied volatilities.

We test the aforementioned theoretical predictions in the data. We first regress bond excess returns onto measures of MBS duration and find a highly significant positive link between duration and excess returns. Moreover, the link becomes stronger as bond maturity increases. The relationship is also economically significant: A one standard deviation change in MBS duration implies a change of 273 basis points in the expected one-year excess return on a 10-year bond.[4] Both the statistical significance and the magnitude of the coefficients are robust to other documented predictors of bond risk premia.

For second moments of bond yields, we find that more negative MBS convexity significantly increases bond yield and swaption implied volatilities. As implied by the model, the effect is most pronounced for intermediate maturities between two and three years. For example, any one standard deviation change in MBS convexity increases bond yield volatility for these maturities by approximately 120 basis points. Moreover, this strong link remains if we add other determinants of interest rate volatility.

We also test how MBS convexity affects bond return volatility—both implied and realized—and find that the estimated slope coefficients are highly significant. To test how MBS convexity affects compensation for volatility risk in fixed income markets, we use measures of bond variance risk premia. In line with the intuition of the model we find that MBS convexity loads positively and highly significantly on these proxies of volatility risk.

Since both MBS duration and convexity depend on interest rate conditions, it is natural to ask whether duration and convexity contain any information above the one encoded in yields. Empirically, we find that neither duration nor convexity is fully spanned by the cross section of Treasury yields. Regressing duration and convexity on the first three principal components of yields results in R^2 of a mere 25% and 40%, respectively. Moreover, running multivariate regressions from bond excess returns (or bond volatility) both on duration (or convexity) as well as the first three yield factors does not reduce the statistical and economic significance of the estimated coefficients. While our stylized model is not designed to address the possibility that MBS duration is unspanned, two features of the model nevertheless speak to the empirical facts. First, aggregate MBS duration depends on the refinancing incentive, i.e., the difference between the average fixed rate paid on outstanding mortgages and the current mortgage rate. This means that duration is not a function of the current level of interest rates alone, but also depends on past levels. Thus, the short rate can only explain a fraction of

[4] We find similar results when we use swap rather than Treasury data.

the variation in duration even though they share the same shocks. Second, mortgage duration decreases after a negative shock to interest rates. This results in a lower term premium and a flatter yield curve. At the same time, a drop in the short rate has a smaller effect on the long end of the curve, leading to a steeper yield curve. Because the two effects partially offset each other, the correlation between duration and the slope of the yield curve is low. In our two factor model, the short rate factor explains most of the variation in the cross section of yields while duration accounts for all the predictability in excess returns.[5]

Our work is related to a series of empirical papers on the negative convexity channel. In our paper, the empirical analysis is guided by an equilibrium term structure model that, to the best of our knowledge, is the first to formalize the intuition behind the negative convexity channel. Moreover, our model has simultaneous implications for bond risk premia and interest rate volatility. Perli and Sack (2003), Chang, McManus, and Ramagopal (2005), and Duarte (2008) test the presence of a linkage between various proxies for MBS hedging activity and interest rate volatility, without considering the simultaneous effect that hedging activity can have on bond pricing. Li and Wei (2013) study a no arbitrage model of the term structure that includes an unspanned MBS supply factor.

In contemporaneous work, Hanson (2013) reports results similar to ours regarding the predictability of bond excess returns by MBS duration. The author's focus is mainly on documenting how MBS duration affects term premia and forward rates. Our paper is different along several dimensions. First, in contrast to his theoretical framework, ours allows us to derive formal predictions regarding the effects of mortgage risk on bond risk premia for different maturities. Second, our model also allows for an analysis of the negative convexity effect on interest rate and bond return volatilities across different maturities, and we provide novel predictions and empirical evidence in this regard.

[5] A simple calibration shows that the short rate explains more than 97% of the variation in the cross section of yields. Empirically, the first three principal components explain roughly the same amount of the cross sectional variation. At the same time, model implied R^2s from a regression of excess returns on MBS duration and the short rate are 14.5% (in line with the data) and a mere 1%, respectively.

Our work is also related to several strands in the asset pricing literature. We make use of the framework developed by Vayanos and Vila (2009). In their model, the term structure of interest rates is determined by the interaction of preferred habitat investors and risk-averse arbitrageurs, who demand higher risk premia as their exposure to long-term bonds increases. Thus, the net supply of bonds matters. Greenwood and Vayanos (2014) use this theoretical framework to study the implications of a change in the maturity structure of government debt supply, similar to the one undertaken in 2011 by the Federal Reserve during "Operation Twist". Our paper is different in at least three respects. First, in our model the variation in the net supply of bonds is driven endogenously by changing MBS duration and not exogenously by the government. Second, the supply factor in Greenwood and Vayanos (2014) explains low frequency variation in risk premia, because movements in maturity-weighted government debt to GDP occur at a lower frequency than movements in the short rate. Our duration factor, on the other hand, explains variations in risk premia at a higher frequency than movements in the level of interest rates. Finally, Greenwood and Vayanos (2014) posit that the government adjusts the maturity structure of its debt in a way that stabilizes bond markets. For instance, when interest rates are high, the government will finance itself with shorter maturity debt and thereby reduce the quantity of interest rate risk held by agents. Our mechanism goes exactly in the opposite direction: Because of the negative convexity in MBS, the supply effect amplifies interest rates shocks.

Corporate debt constitutes another important class of fixed income instruments, which is shown to be negatively correlated with the supply of Government debt (see, e.g., Greenwood, Hanson, and Stein (2010) and Bansal, Coleman, and Lundblad (2011)). For example, Greenwood, Hanson, and Stein (2010) show that firms choose their debt maturity in a way that tends to offset the variations in the supply and maturity of Government debt. However, the authors find no relationship between corporate debt and MBS supply, which provides us with an additional motivation to focus on the latter.

This paper contributes to the literature on equilibrium term structure models, e.g., Le, Singleton, and Dai (2010), Xiong and Yan (2010), Ehling, Gallmeyer, Heyerdahl-Larsen, and Illeditsch (2013), Bansal and Shaliastovich (2013), Hong, Sraer, and Yu

(2013), and Le and Singleton (2013), among others. Contrary to these papers, in which bond risk premia are determined by macroeconomic fundamentals or differences in beliefs about these fundamentals, we focus on the aggregate demand and supply of bonds as the main driver of risk premia. Le and Singleton (2013) also show that time-varying market prices of risk, in addition to time-varying quantities of risk, are required to explain bond risk premia. In our model, duration drives the variation in the market price of risk and, hence, bond risk premia. Recent empirical work by Duffee (2011), Le and Singleton (2013) and Joslin, Priebsch, and Singleton (2014) documents the existence of unspanned factors that are unrelated to the cross section of yields but have a significant impact on risk premia. In our model, duration is spanned by the cross section of yields by construction, but not by the short rate factor that accounts for most of the variation in the shape of the yield curve. Empirically, we find that MBS duration is not spanned by the first three principal components that drive virtually all variation in the cross section of yields. Finally, Joslin, Le, and Singleton (2013) and Fenou and Fontaine (2013) develop term structure models that include additional lags in the dynamics of yield factors. Similarly, in our paper, MBS duration depends on both current and past yields.

There is an important literature that studies the optimal prepayment in the MBS market (see, e.g., Schwartz and Torous (1989), Stanton (1995), Stanton and Wallace (1998), Longstaff (2005) and, more recently, Agarwal, Driscoll, and Laibson (2013) for examples of prepayment models and optimal prepayment decisions). Furthermore, there is evidence that households' prepayment is too sluggish and that this non-optimal prepayment can be best explained using micro-level data (see, e.g., Campbell's 2006 AFA Presidential Address for an overview). In this paper we are interested in how the aggregate properties of prepayment affect the risk of mortgage-related portfolios of financial institutions. Boyarchenko, Fuster, and Lucca (2014) study the spread between MBS rates and Treasuries and find that prepayment risk explains well the cross-section but non-prepayment risk related factors explain the time-series. Closest to us are Gabaix, Krishnamurthy, and Vigneron (2007) who study the effect of limits to arbitrage in the MBS market. The authors show that, while mortgage prepayment risk resembles a wash on an aggregate level, it nevertheless carries a positive risk premium because it is the risk exposure of financial intermediaries that matters. Our paper is based on a similar premise. Different from these authors, however, we do not study prepayment risk, but changes in interest rate risk of MBS that are driven by the prepayment probability and their effect on the term structure of interest rates.

Finally, our paper is related to the literature that studies the effect of the recent Fed's purchase of long-term assets on interest rates (see, e.g., Gagnon, Raskin, Remache, and Sack (2010), Krishnamurthy and Vissing-Jorgensen (2011), D'Amico, English, Lopez-Salido, and Nelson (2012), D'Amico and King (2013)). While our model mainly focusses on the relationship between long-term yields and volatilities and MBS duration, other papers present evidence for alternative channels. For example, Krishnamurthy and Vissing-Jorgensen (2011) report that QE1, which involved major purchases of agency MBS, led not only to large reductions in mortgage rates, but also helped drive down Treasury yields and caused a drop in the default risk premium of corporate bonds. They interpret their findings as evidence for a long-term safety channel. D'Amico and King (2013) emphasize a scarcity channel, i.e. a localized effect of supply shocks on yields of nearby maturities.

The remainder of the paper is organized as follows. Section 1 sets up a model of the term structure of interest rates and MBS duration. Section 2 describes the data used and Section 3 empirically tests our hypotheses. Section 4 concludes. Proofs are deferred to the Appendix. An Online Appendix contains additional institutional details on the MBS market.

1 Model

In this section we build a stylized equilibrium dynamic term structure model in which the supply of fixed income securities is driven by MBS duration. Motivated by the size of the MBS market relative to the Treasury market and the magnitude of the changes in MBS duration, we focus on modeling the MBS channel only.[6,7] In particular, we interpret the shortening of MBS duration due to the increased probability of future refinancing as a negative shock to the supply of long-term bonds: Before refinancing happens and investors have access to new mortgage pools, the interest rate risk profile of mortgage-related securities available to investors resembles that of relatively short maturity bonds. This induces additional buying pressure on Treasuries if investors want to keep their interest rate risk exposure constant.[8]

The aggregate duration of outstanding MBS is primarily driven by two forces: (i) changes in the level of interest rates that affect the prepayment probability of each outstanding mortgage, and (ii) actual prepayment that changes the composition of the aggregate mortgage pool. In the following, we posit a simple one-factor Vasicek (1977) model for the short-rate and a reduced-form model of aggregate prepayment in the spirit of Gabaix, Krishnamurthy, and Vigneron (2007) that captures these two effects.

The literature has adopted different ways to describe the prepayment behavior. Dunn and McConnell (1981) and Brennan and Schwartz (1985) pioneered the application of contingent claim techniques to the problem by modelling prepayment as an optimal decision by borrowers who minimize the value of their loans. This approach is used more recently in Longstaff (2005). However, micro-level evidence suggests that individual household prepayment is often non-optimal relative to the option pricing approach and prepayment depends both on observable factors such as the incentive to refinance, sea-

[6]The three largest categories in the fixed income market in terms of total amount outstanding are MBS, Treasuries, and corporate debt with MBS constituting the largest fraction of around 25% (Treasuries and corporate bonds each account for around 20%). For the period 1997 to 2012 the total amount outstanding in MBS is larger by a factor of 1.35 compared to the amount outstanding in Treasuries; see SIFMA (2013). The relative size of the MBS compared to the Treasury market is also economically relevant. As discussed above, taking into account the value of outstanding mortgage debt, we calculate that a one monthly standard deviation shock to MBS duration is a dollar duration equivalent of a USD 368bn shock to the supply of 10-year Treasuries.

[7]Another justification for our exclusive focus on MBS duration is the lack of relation between MBS duration and other potential sources of variation in aggregate duration. We find little commonality between MBS duration and convexity on the one hand, and the Government debt supply factor proposed by Greenwood and Vayanos (2014) on the other hand. Greenwood, Hanson, and Stein (2010) argue that changes in the maturity structure of corporate debt, that could potentially constitute another important source of supply shocks, tend to offset the variations in the supply and maturity of Government debt but are not related to MBS.

[8]Market participants can invest in new mortgage loans by buying corresponding MBS. Up to 90 days before those MBS are issued, investors have access to them through the "to-be-announced" (TBA) market. We thank Douglas McManus for an insightful discussion; see also Vickery and Wright (2010).

sonality or the level of house prices in general as well as on non-observable factors such as the media effect (for example prepayment rates can directly react to specific news stories relating to interest rates and the mortgage market, see e.g., Soo (2014)). Non-optimality has been a major topic in the literature on MBS valuation (see, e.g., Campbell (2006) or Amromin, Huang, and Sialm (2007)). To address this in the context of a contingent claim analysis, Stanton and Wallace (1998) add an exogenous delay to refinancing, while Schwartz and Torous (1989) and Stanton (1995) are examples for purely reduced-form econometric models that aim to capture the empirical behavior.[9] Our motivation for using a reduced-form approach is twofold. First, we avoid making strong assumptions regarding the optimal prepayment. Second, the incentive to prepay on aggregate is well explained by interest rates themselves. Boudoukh, Whitelaw, Richardson, and Stanton (1997) for example show that for a given MBS coupon the level of long-term interest rates is a very good proxy for the likelihood that a mortgage will be prepaid.

1.1 Bond market

Time is continuous and goes from zero to infinity. We denote the time t price of a zero coupon bond paying one dollar at maturity $t+\tau$ by Λ_t^τ, and its yield by $y_t^\tau = -\frac{1}{\tau}\log\Lambda_t^\tau$. The short rate r_t is the limit of y_t^τ when $\tau \to 0$. We take r_t as exogenous and assume that its dynamics under the physical probability measure are given by

$$dr_t = \kappa\left(\theta - r_t\right)dt + \sigma dB_t, \tag{1}$$

where θ is the long run mean of r_t, κ is the speed of mean reversion, and σ is the volatility of the short rate.[10]

At each date t, there exists a continuum of zero-coupon bonds with time to maturity $\tau \in (0, T]$ in total supply of s_t^τ. Bonds are held by financial institutions. We think about them as representing a range of investors such as investment banks, hedge funds, and

[9] Parameters in these models change over time, which represents an unhedgeable risk in the cash flows of MBS. This is the focus in Gabaix, Krishnamurthy, and Vigneron (2007).

[10] Similar to Collin-Dufresne and Harding (1999), we use a Vasicek process primarily because of its simplicity.

fund managers, who trade actively in fixed income markets and act as marginal investors there. Financial institutions are competitive and have mean-variance preferences over the instantaneous change in the value of their bond portfolio. If x_t^τ denotes their holdings in maturity-τ bonds at time t, the investors' budget constraint becomes

$$dW_t = \left(W_t - \int_0^T x_t^\tau \Lambda_t^\tau d\tau \right) r_t dt + \int_0^T x_t^\tau \Lambda_t^\tau \frac{d\Lambda_t^\tau}{\Lambda_t^\tau} d\tau, \tag{2}$$

and their optimization problem is given by

$$\max_{\{x_t^\tau\}_{\tau \in (0,T]}} \mathrm{E}_t \left[dW_t \right] - \frac{\alpha}{2} \mathrm{Var}_t \left[dW_t \right], \tag{3}$$

where α is their absolute risk aversion. Since financial institutions have to take the other side of the trade in the bond market, the market clearing condition is given by

$$x_t^\tau = s_t^\tau, \quad \forall t \text{ and } \tau. \tag{4}$$

1.2 MBS duration

The supply of bonds is driven by households' mortgage liabilities. Without explicitly modeling it, we think about a continuum of households who do not themselves invest in bonds but take fixed rate mortgage loans that are then sold on the market as MBS. Households refinance their mortgages when the incentive to do so is sufficiently high. Prepaying a mortgage is equivalent to exercising an American option. As shown in Richard and Roll (1989), the difference between the fixed rate paid on a mortgage and the current mortgage rate is a good measure of the moneyness of this prepayment option. Because households can have mortgages with different characteristics, we focus on the average mortgage coupon (interest payment) on outstanding mortgages, c_t. Following Schwartz and Torous (1989) we approximate the current mortgage rate by the long-term interest rate $y_t^{\bar{\tau}}$ with reference maturity $\bar{\tau}$.[11,12] In sum, we define the refinancing incentive as $c_t - y_t^{\bar{\tau}}$.

On aggregate, refinancing activity does not change the size of the mortgage pool: when a mortgage is prepaid, another mortgage is issued. However, the average coupon c_t, is affected by prepayment because the coupon of the newly issued mortgages is a function of the current level of mortgage rates. We assume that the evolution of the average coupon is a function of the refinancing incentive:

$$dc_t = -\kappa_c \left(c_t - y_t^{\bar{\tau}}\right) dt, \tag{5}$$

with $\kappa_c > 0$. This means that a higher refinancing incentive leads to more prepayments, and new mortgages decrease the average coupon by more. Because our focus is the feedback between MBS market and interest rates, we also assume that on aggregate, there is no additional uncertainty about refinancing. The upper panel of Figure 1 provides empirical motivation for equation (5). We plot the difference between long-term interest rates and the average MBS coupon, together with the subsequent change in the average coupon. The two series are closely aligned with the coupon reacting with a slight delay to a change in the refinancing incentive.

The distinctive feature of mortgage-related securities is that their duration depends primarily on the likelihood that they will be refinanced in the future. The MBS coupon and the level of interest rates proxy for the expected level of prepayments and the moneyness of the option (see Boudoukh, Whitelaw, Richardson, and Stanton (1997)). We thus assume that the aggregate dollar duration of outstanding mortgages is a function of the refinancing incentive:

$$D_t = \theta_D - \eta_y \left(c_t - y_t^{\bar{\tau}}\right), \tag{6}$$

[11] An argument against this choice is mentioned in Krishnamurthy (2010) who studies the spread between mortgage rates and interest rate swaps. He notes that especially during autumn 2008 there was a large disconnect between the two which can be attributed to a flight-to-liquidity episode from relatively illiquid mortgages to more liquid government bonds. Since these considerations are outside the scope of the model, we leave this to future research.

[12] We use $\bar{\tau} = 10$ years. According to Hancock and Passmore (2010), it is common industry practice to use either the 5- or 10-year swap rate as a proxy for MBS duration.

where duration, $D_t \equiv -dMBS_t/dy_t^{\bar{\tau}}$, is the observable sensitivity of the aggregate mortgage portfolio value to the changes in a reference long-maturity rate $y_t^{\bar{\tau}}$, and $\theta_D, \eta_y > 0$ are constants. The middle panel of Figure 1 provides empirical motivation for equation (6). We plot the difference between long-term interest rates and the average MBS coupon, together with aggregate MBS duration. The two series are again very closely aligned. Overall, we note that our model captures well the key stylized properties of aggregate refinancing activity.

[Insert Figure 1 here.]

Combining equations (5) and (6) gives us the dynamics of D_t:

$$dD_t = \kappa_D \left(\theta_D - D_t\right) dt + \eta_y dy_t^{\bar{\tau}}, \tag{7}$$

where $\kappa_D = \kappa_c$. Hence, dollar duration is driven both by changes in long-term interest rates and refinancing activity. In particular, there are two sources of mean reversion in aggregate duration: the mean-reversion in interest rates, $\eta_y E_t\left(dy_t^{\bar{\tau}}\right)$, that affects the dollar duration of each individual mortgage, and the renewal in the aggregate pool of mortgages, $\kappa_D\left(\theta_D - D_t\right) dt$. The parameter $\eta_y = dD_t/dy_t^{\bar{\tau}}$ is the negative of the dollar convexity: When $\eta_y > 0$, lower interest rates increase the probability of borrowers prepaying their mortgages in the future, leading to a lower duration. The lower panel of Figure 1 plots the MBS convexity series. Comparative statics with respect to η_y allow us to derive predictions regarding the effect of negative convexity on interest rate volatility.[13]

Why do shocks to mortgage duration matter? First, note that while lower interest rates trigger a certain amount of refinancing of the most in-the-money mortgages, they also increase the probability of future prepayment and thus decrease the duration of all outstanding mortgages. In addition, there is ample empirical evidence that shows

[13] A model where η_y itself follows a stochastic process would not fall into one of the standard tractable classes of models. Appendix C presents a version of the model that accommodates time-varying convexity. While this model implies a quadratic instead of an affine term structure, it leads to identical qualitative predictions.

that households' refinancing is gradual and sluggish (see Campbell (2006)). The progressive nature of refinancing ($\kappa_c < \infty$) leaves financial institutions who invest in MBS on aggregate short of duration exposure.[14] Second, there is no evidence suggesting that households are actively hedging the changes in the duration of their mortgage liabilities by trading bonds. Interest rate instruments are traded primarily by financial institutions and, moreover, trading volume in Treasuries and MBS is very high.[15] Consistent with this evidence, we assume that the supply of bonds is held by financial institutions who actively manage the duration of their portfolios and act as marginal investors in the bond market. The mortgage choice of households affects the supply of fixed income securities, s_t^τ, through the duration of mortgages, D_t, but in addition to this channel households are not present on either side of the market-clearing condition (4). As a result, changes in MBS duration matter for bond prices, and the MBS feedback channel is not a wash.

1.3 Equilibrium term structure

Before solving for equilibrium yields, we determine the market price of interest rate risk.

Lemma 1. *Given (1)-(4), the unique market price of interest rate risk is proportional to the dollar duration of the total supply of bonds:*

$$\lambda_t = \alpha\sigma \frac{d\left(\int_0^T s_t^\tau \Lambda_t^\tau d\tau\right)}{dr_t}. \tag{8}$$

Lemma 1 follows from the absence of arbitrage and implies that, regardless of the specific maturity composition of the supply of bonds, the market price of interest rate risk is determined by its total quantity. This means that in order to derive the equilibrium

[14]Gabaix, Krishnamurthy, and Vigneron (2007) make a related point that from the perspective of financial intermediaries who are the marginal investors in MBS, this sluggishness creates an unhedgeable risk which is priced.

[15]Financial intermediaries and institutional investors hold approximately 25% of the total amount outstanding in Treasuries, and daily trading volume is almost 10% of the total amount outstanding. In addition, these financial intermediaries hold around 30% of the total amount outstanding in MBS, and daily trading is almost 25% of the total amount outstanding. Data are for the period 1997 to 2012; see Securities Industry and Financial Markets Association (2013) and Flow of Funds Tables of the Federal Reserve. The Online Appendix contains a summary of holdings and size of the MBS market.

term structure, it is not necessary to explicitly model the full dynamics of the supply of bonds, but it is sufficient to capture its duration.

In this paper we are interested in only one source of variation in the duration of bond supply, namely the changes in MBS dollar duration, and therefore we assume that:

$$\lambda_t = \alpha\sigma \frac{dMBS_t}{dr_t}. \tag{9}$$

Using a simple chain rule, $\frac{dMBS_t}{dr_t} = \frac{dMBS_t}{dy_t^{\bar{\tau}}}\frac{dy_t^{\bar{\tau}}}{dr_t}$, we rewrite (9) in terms of the sensitivity to the reference long-maturity rate $y_t^{\bar{\tau}}$:

$$\lambda_t = -\alpha\sigma_y^{\bar{\tau}} D_t, \tag{10}$$

where $\sigma_y^{\bar{\tau}} \equiv \frac{dy_t^{\bar{\tau}}}{dr_t}\sigma$, the volatility of $y_t^{\bar{\tau}}$, is a constant to be determined in equilibrium.

We look for an equilibrium in which yields are affine in the short rate and the duration factor. Under the conjectured affine term structure, the physical dynamics of MBS duration (7) can be written as

$$dD_t = (\delta_0 - \delta_r r_t - \delta_D D_t)\, dt + \eta_y \sigma_y^{\bar{\tau}} dB_t, \tag{11}$$

where δ_0, δ_r and δ_D are constants to be determined in equilibrium. In turn, equations (1), (10) and (11) together imply that the dynamics of the short rate and the MBS duration factor under the risk-neutral measure are

$$dr_t = \left[\kappa\left(\theta - r_t\right) + \alpha\sigma\sigma_y^{\bar{\tau}} D_t\right] dt + \sigma dB_t^{\mathbb{Q}} \text{ and} \tag{12}$$

$$dD_t = \left(\delta_0 - \delta_r r_t - \delta_D^{\mathbb{Q}} D_t\right) dt + \eta_y \sigma_y^{\bar{\tau}} dB_t^{\mathbb{Q}}, \tag{13}$$

where $\delta_D^{\mathbb{Q}} \equiv \delta_D - \alpha\eta_y\left(\sigma_y^{\bar{\tau}}\right)^2$.

We now have all the ingredients to solve for the equilibrium term structure.

Theorem 1. *In the term structure model described by (12) and (13), equilibrium yields are affine and given by*

$$y_t^{\tau} = \mathcal{A}(\tau) + \mathcal{B}(\tau) r_t + \mathcal{C}(\tau) D_t, \tag{14}$$

where the functional forms of $\mathcal{A}(\tau)$, $\mathcal{B}(\tau)$, and $\mathcal{C}(\tau)$ are given by (A-29)-(A-31), and the parameters $\sigma_y^{\bar{\tau}}$, δ_r, δ_D, and δ_0 satisfy

$$\sigma_y^{\bar{\tau}} = \frac{\sigma \mathcal{B}(\bar{\tau})}{1-\eta_y \mathcal{C}(\bar{\tau})}, \ \delta_r = \frac{\kappa \eta_y \mathcal{B}(\bar{\tau})}{1-\eta_y \mathcal{C}(\bar{\tau})}, \ \delta_D = \frac{\kappa_D}{1-\eta_y \mathcal{C}(\bar{\tau})}, \ \textit{and } \delta_0 = \delta_r \theta + \delta_D \theta_D. \tag{15}$$

Equations (15) have a solution whenever α is below a threshold $\bar{\alpha} > 0$.

Theorem 1 motivates the inclusion of mortgage market factors in the term structure analysis. Equation (14) implies that long-term yields depend on two separate factors, the short rate and the aggregate dollar duration of mortgages, even though the model has only one shock. The following corollary clarifies the relationship between the factors and the shape of the yield curve in the model.

Corollary 1. *The theoretical R^2s of univariate regressions of the duration factor D_t on the short rate factor r_t, the long-term yield $y_t^{\bar{\tau}}$, and the slope $y_t^{\bar{\tau}} - r_t$ are given by*

$$R^2_{D,r} = 1 - \frac{\delta_D}{\kappa + \delta_D}, \tag{16}$$

$$R^2_{D,y} = 1 - \frac{\delta_D}{\kappa\left(1 + \frac{\mathcal{C}(\bar{\tau})}{\mathcal{B}(\bar{\tau})}\frac{\delta_r}{\kappa}\right)^2 + \delta_D}, \textit{ and} \tag{17}$$

$$R^2_{D,y-r} = 1 - \frac{\delta_D}{\kappa\left(1 + \frac{\mathcal{C}(\bar{\tau})}{\mathcal{B}(\bar{\tau})-1}\frac{\delta_r}{\kappa}\right)^2 + \delta_D}. \tag{18}$$

As is evident from Corollary 1, neither the short rate nor the long-term yield fully explain MBS duration provided that $\delta_D \neq 0$, which from (15) is equivalent to $\kappa_D \neq 0$. This is because, in addition to the current mortgage rate, aggregate duration depends also on the history of mortgage rates that determines the coupon of outstanding mortgages. As a result, the dynamics of aggregate duration are driven not only by changes in interest rates, but also by the renewal of the pool of mortgages.[16] Also from Corollary 1, the slope of the yield curve, defined as $y_t^{\bar{\tau}} - r_t$, is correlated with duration. However, the effect that duration has on the term premium tends to reduce $R^2_{D,y-r}$. To see this note

[16]Formally, when $\kappa_D \neq 0$, interest rates in our model are non-Markovian with respect to the short rate r_t alone. However, their history dependence can be summarized by an additional Markovian factor, namely the duration D_t.

that a negative shock to the short rate steepens the slope ($\mathcal{B}(\bar{\tau}) < 1$), because interest rates are expected to mean revert. At the same time, the corresponding drop in duration reduces the term premium due to $\mathcal{C}(\bar{\tau}) > 0$. As the two effects tend to cancel each other out, the correlation between duration and the slope of the yield curve is pushed closer to 0.

1.4 Model implications

Our model has a series of implications regarding the effect of MBS risk on bond risk premia, as well as bond price and yield volatilities. We summarize them in three propositions which will guide our empirical analysis.

Proposition 1. *The dollar duration of MBS positively predicts excess bond returns for all maturities and the effect is stronger for longer maturities. Moreover, this continues to hold when we control for the short rate.*

The market price of interest rate risk depends on the quantity of this risk that financial institutions hold to clear the supply. In turn, bonds with higher exposure to interest rate risk are more affected. As a result, MBS duration should predict excess bond returns and the effect is stronger for longer maturity bonds.[17]

More formally, for the theoretical slope coefficient $\beta^{\tau,h}$ of the regression of excess returns on bonds with maturity τ over horizon h on the MBS duration factor D_t, we verify that $\lim_{\tau \to h} \beta^{\tau,h} = 0$ and $d\beta^{\tau,h}/d\tau > 0$ and hence, $\beta^{\tau,h}$ is positive and increasing across maturities.

Consider the following bivariate regression where we control for the short rate (or more generally the level of interest rates)

$$rx_{t,t+h,\tau} = \alpha_{D,r} + \left(\beta_1^{\tau,h}, \beta_2^{\tau,h}\right) \begin{pmatrix} D_t \\ r_t \end{pmatrix} + \epsilon_{t+h}.$$

[17]Note that the effect of MBS dollar duration on the level of yields is not necessarily monotonic in maturity. A yield depends on the average of risk premia over the life of the bond. Higher risk premia increase yields. However, because of mean reversion in interest rates and duration, we expect risk premia at longer horizons to be lower. We are not testing this implication empirically, because duration itself depends on yields, thus causing an endogeneity problem for identification.

The slope coefficient on duration $\beta_1^{\tau,h}$ remains positive and increasing in maturity, while the slope coefficient on the short rate $\beta_2^{\tau,h}$ is negative and decreasing (i.e., becoming more negative) in maturity. This is the case because the level of interest rates does not contain any information about the current market price of risk beyond that already encoded in duration. However, including the short rate allows to control for the mean reversion in interest rates, and therefore to better predict the changes in duration over the return horizon h.

Proposition 2. *The volatility of all yields is increasing in the negative dollar convexity of MBS and the effect is strongest for intermediary maturities, i.e., is hump-shaped.*

Higher MBS convexity implies that the quantity of duration risk and therefore the market price of risk is more sensitive to changes in interest rates. Moreover, because MBS convexity is negative, portfolio rebalancing by investors amplifies, rather than offsets, the effect that the initial shock to the short rate has on long-term interest rates. As a result, interest rate volatility is higher.

The link between convexity and volatility has a term structure dimension. Short-maturity yields are close to the short rate and therefore are not significantly affected by the variations in the market price of risk. For long maturities, we expect the duration of MBS to revert to its long-term mean. At the limit, yields at the infinite horizon should not be affected by current changes in the short rate and MBS duration at all.[18] As a result, the effect of MBS convexity on yield volatilities has a hump-shaped term structure.

More formally, bond yield volatilities σ_y^τ for all maturities $0 < \tau < \infty$ are increasing in the negative convexity parameter $\eta_y : d\sigma_y^\tau / d\eta_y > 0$. In addition, $\lim_{\tau \to 0} \sigma_y^\tau = \sigma$ and $\lim_{\tau \to \infty} \sigma_y^\tau = 0$, where neither limit depends on η_y. Therefore, the effect of negative convexity on yield volatilities tends to zero at very short and very long maturities, and hence it must be hump-shaped.

Proposition 3. *The effect of negative dollar convexity on the volatility of bond returns is positive.*

[18] This is just an application of a more general argument by Dybvig, Ingersoll, and Ross (1996) on why the yield volatility curve for long maturities should be downward sloping.

Proposition 3 is a simple corollary to Proposition 2. According to the latter, the volatility of all interest rates σ_y^τ is increasing in negative convexity, and as a result the volatility of bond returns $\sigma_y^\tau \tau$ is also increasing in negative convexity. Note that if we consider maturities long enough, the effect of negative convexity on bond return volatility across maturities is hump-shaped: Even though longer maturity bonds have higher duration and therefore are more sensitive to changes in interest rates, the effect of negative convexity on interest rate volatility tends to zero at very long maturities faster than the increase in bond duration.

Since negative convexity, η_y, is constant in the model we posit above, Propositions 2 and 3 are comparative static results. In Appendix C we show that our framework can be extended to a model with stochastic convexity. Importantly, in this less parsimonious version of the model we can derive time series predictions that are equivalent to Propositions 2 and 3.

2 Data

We use data from several sources. From Bloomberg, we collect data on interest rate swaps. Furthermore, we use Treasury data from the Federal Reserve Board and agency bond index data from Datastream. Data are weekly and span the time period from January 1997 through December 2012 for a total of 835 observations.

2.1 Mortgage data

We use estimates of MBS duration, convexity, and average coupon from Barclays. The Barclays US MBS index covers mortgage-backed pass-through securities guaranteed by Ginnie Mae, Fannie Mae, and Freddie Mac. The index is comprised of pass-throughs backed by conventional fixed rate mortgages and is formed by grouping the universe of over one million agency MBS pools into generic pools based on agency, program (i.e., 30-year, 15-year, etc.), coupon (e.g., 6.0%, 6.5%, etc.), and vintage year (e.g., 2011, 2012, etc.). A generic pool is included in the index if it has a weighted-average contractual maturity greater than one-year and more than USD 250m outstanding.

The middle panel of Figure 1 depicts MBS duration. The average duration in our sample is around 4.5 years which is mainly due to the prepayment option in fixed rate mortgages which reduces the duration of MBS. The lower panel plots MBS convexity which is negative throughout the whole sample and the average convexity is around -1.5. We note that virtually all of the variation in dollar duration and dollar convexity of the index is driven by, respectively, duration and convexity themselves and not by the changes in the price. Therefore, for simplicity we use duration and convexity as our factors.

2.2 Interest rates and excess returns

We use the Gürkaynak, Sack, and Wright (2007, GSW henceforth), zero coupon yield data available from the Federal Reserve Board. Unlike the Fama and Bliss discount bond database from CRSP, the GSW data is available at the weekly frequency. We use the raw data to calculate annual Treasury bond excess returns for 2- to 10-year bonds. We also download interest rate swap data from Bloomberg from which we bootstrap a zero coupon yield curve.

We denote the return on a τ-year bond with price Λ_t^τ by $r_{t+1}^\tau = \log \Lambda_{t+1}^{\tau-1} - \log \Lambda_t^\tau$. The annual excess bond return is then defined as $rx_{t+1}^\tau \equiv r_{t+1}^\tau - y_t^1$, where y_t^1 is the one-year yield. As we have weekly data, the annual excess returns are overlapping by 51 weeks. From the same data, we also construct a tent-shaped factor from forward rates, the Cochrane and Piazzesi (2005) factor (CP factor, labeled cp_t).

We calculate the slope of the term structure as the difference between the 10-year and the one-year zero coupon yield (labeled slope_t).

2.3 Bond volatility

Using the GSW yields ranging from six months to ten years, we estimate a time-varying term structure of yield volatility. We sample the data at the weekly frequency and take log yield changes. We then construct rolling window measures of realized volatility which represent the conditional bond yield volatility. The resulting term structure of unconditional volatility exhibits a hump shape consistent with the stylized facts reported in Dai and Singleton (2010), with the volatility peak being at the two-year maturity.

Mueller, Vedolin, and Yen (2012) calculate measures of model-free implied and realized bond market volatilities for the one-month horizon using Treasury futures and options data from the Chicago Mercantile Exchange (CME). We use their data for the 30-year Treasury bond and henceforth label the realized volatility trv_t and the implied volatility tiv_t. Moreover, the difference between the expected variance under the risk-neutral (tiv^2) and physical probability measure (trv^2) is defined as the ex ante variance risk premium ($\mathrm{vrp}_t \equiv \mathrm{tiv}_t^2 - \mathrm{trv}_t^2$). We also construct a measure of volatility of variance for both the realized and implied proxies. Rolling volatility is calculated using a 52 week window and we label the two time series tivov_t and trvov_t, for implied and realized volatility of variance, respectively.

From Bloomberg, we also get implied volatility for at-the-money swaptions for different maturities ranging from one to ten years and we fix the tenor to ten years. We label these swaption implied volatilities by $\mathrm{iv}^{\tau y 10 y}$, where $\tau = 1, \ldots, 10$.

Another way to capture bond price volatility risk is through returns on at-the-money straddles, which are portfolios mainly exposed to volatility risk (see Collin-Dufresne and Goldstein (2002)). We construct monthly straddle returns using at-the-money options written on the 30-year Treasury bond futures. We label this time series $\mathrm{straddle}_t$.

2.4 Other variables

Motivated by Hu, Pan, and Wang (2013) who report a link between swaption implied volatility and measures of liquidity in bond markets, we use their proxy of noise illiquidity, which measures an average yield pricing error from the Nelson, Siegel and Svensson model (see Svensson (2004)). To this end, we construct a daily measure of noise illiquidity from bond data available from Datastream. We also construct eight principal components from the cross section of 132 macro factors (see Ludvigson and Ng (2009)).

3 Empirical analysis

In this section we test the predictions of the model in the data and find strong support for them. To the best of our knowledge, the literature has not proposed an alternative explanation for the predictive power of MBS duration and convexity. Nevertheless, we also control for other well-known determinants of bond risk premia and interest rate volatility which are not explicitly accounted for in our model in order to address a potential omitted variable problem. We find that not only MBS duration and convexity remain statistically significant, but also the economic size of the coefficients stays stable across different specifications.

For simplicity, we use duration and convexity instead of dollar duration and dollar convexity in our benchmark results given that they drive virtually all the variation in the dollar time series. However, the results remain unchanged if we use dollar duration and dollar convexity. All regressions are standardized, meaning that we first de-mean and then divide each variable by its standard deviation to make slope coefficients comparable across different regressors. With each estimated coefficient, we report t-statistics adjusted for Newey and West (1987) standard errors. The lag length is determined using the Stock and Watson (2007) rule.

3.1 Bond risk premia

Hypothesis 1. *A regression of bond excess returns on the duration of MBS yields a positive slope coefficient for all maturities. Moreover the coefficients are increasing in bond maturity and remain significant when we control for the level of interest rates.*

This hypothesis is equivalent to Proposition 1. To test Hypothesis 1, we run linear regressions of annual excess returns on the duration factor. The regression is as follows:

$$rx^{\tau}_{t+1} = \beta^{\tau}_1 \text{duration}_t + \beta^{\tau}_2 \text{level}_t + \epsilon^{\tau}_{t+1},$$

where level_t is the first principal component of the yield curve at time t. The univariate results are depicted in the upper two panels of Figure 2, which plot the estimated slope coefficients of duration, i.e., $\hat{\beta}_1$ (upper left panel), and the associated adjusted R^2 (upper right panel). Both univariate and multivariate results are presented in Table 1.

[Insert Figure 2 and Table 1 here.]

The univariate regression results indicate that MBS duration is a significant predictor of bond excess returns at all maturities. The coefficient has the expected positive sign and is increasing with maturity. The estimated coefficients are also economically significant, especially for longer maturities. For example, for any one standard deviation increase in duration, there is a $0.381 \times 7.19\% = 273$ basis point increase in expected 10-year bond excess returns. Adjusted R^2s range from 2% for the shortest maturity to 14% for the longest maturity.

Because duration is in part determined by the current level of interest rates, we include it as a control in our multivariate test. In the data, the unconditional correlation between MBS duration and the yield level is equal to 0.55. Despite this positive correlation, the model predicts that the slope coefficients on the two variables have the opposite sign. On the one hand, higher duration implies a higher market price of interest rate risk and predicts higher bond returns. On the other hand, for a given duration, a higher level of interest rates does not contain additional information about the current market price of risk, but forecasts lower interest rates and lower duration in the future. In line with the model predictions we find that the slope coefficient on duration remains positive and increasing with maturity, while the slope coefficient on the level of interest rates has a negative sign and is decreasing with maturity.

3.2 Bond yield volatility

Proposition 2 states that bond yield volatility is increasing in the negative convexity of MBS. In addition, the effect should be strongest at intermediate maturities, meaning that we should observe a hump-shaped pattern. We therefore test the following hypothesis in the data.

Hypothesis 2. *A regression of conditional yield volatility on the negative convexity of MBS results in a positive slope coefficient for all maturities. Moreover, the coefficients are the largest for intermediate maturities.*

We run the following two univariate regressions from conditional bond yield volatility and swaption implied volatility onto convexity:

$$\mathrm{vol}_t^\tau / \mathrm{iv}_t^{\tau y 10y} = \beta_1^\tau \mathrm{convexity}_t + \epsilon_t^\tau,$$

where vol_t^τ is the conditional bond yield volatility at time t of a bond with maturity $\tau = 1, \ldots, 10$ years and $\mathrm{iv}_t^{\tau y 10y}$ is the τ-year maturity implied volatility from swaptions on the 10-year swap rate. The reason to include swaption implied volatility is that hedging activity could potentially affect both. For example, Wooldridge (2001) notes that non-government securities were routinely hedged in the Treasury market until the financial crisis of 1998 when investors started hedging their interest rate exposure in the swaptions market. This point is also made in Perli and Sack (2003), Duarte (2008), or Feldhütter and Lando (2008). More recently, Begenau, Piazzesi, and Schneider (2013) and Landier, Sraer, and Thesmar (2013) also report that financial institutions hold large positions in interest rate derivatives to retain their interest rate risk exposure.

The univariate results are presented in the lower two panels of Figure 2 and Table 2. In line with our model predictions, we find a significant effect from convexity onto bond yield volatility and the effect is most pronounced for intermediate maturities. The estimated slope coefficients produce the hump shaped feature similar to the one observed in the unconditional averages of yield volatility. Adjusted R^2s range from 6% for the shortest maturities, increase to 7% for the two and three year maturities and decrease again to 3% for longer maturities. Estimated coefficients are not only statistically but also economically significant: For the two-year maturity, any one standard deviation change in MBS convexity is associated with a 120 basis point increase in bond yield volatility.

[Insert Table 2 here.]

The same picture emerges for implied volatilities from swaptions reported in Panel B: Higher convexity induces higher volatility on swaptions. All estimation coefficients are statistically significant with t-statistics ranging from 2.09 to 3.82.

An obvious concern with our regression results is that negative convexity could itself depend on volatility. Note, that it is a priori unclear in which direction volatility affects convexity as this depends on whether a particular mortgage is in-, out-, or at-the-money.[19] For an at-the-money MBS, an increase in volatility will lead to an increase in negative convexity. Discount (i.e., small negative to positive convexity) and premium (negative convexity) mortgages will in general have a much lower sensitivity to changes in volatility, and the effect could go in the opposite direction.[20]

We run Granger causality tests between MBS convexity and volatility and present the results in Figure 3. In the left panel, we plot p-values from F-tests that assess the null hypothesis whether negative convexity does not Granger cause volatility. On the right panel, we plot the p-values of the reversed Granger regression, i.e., we test the null hypothesis whether volatility does not Granger cause negative convexity. We note that for standard confidence levels, we can reject the null of no Granger causality up to a maturity of seven years. On the other hand, volatility does not seem to Granger cause convexity, as we cannot reject the null hypothesis up to a maturity of five years.

[Insert Figure 3 here.]

As an additional robustness check we use the lagged values of convexity as instruments in an Instrumental Variable (IV) estimation. Running two-stage-least-square regressions, we find that the standard errors are in fact smaller if we lag the instrument further. In line with the empirical and theoretical findings in Stock, Wright, and Yogo (2002), we conclude that IV estimation performs worse than OLS and therefore report OLS results only (see also Krishnamurthy and Vissing-Jorgensen (2012)).

[19]This is analogous to the Zomma (sensitivity of an option's Gamma with respect to changes in the implied volatility) for equity options.

[20]We thank Bruce Phelps at Barclays Capital for insightful discussions on this.

3.3 Other volatility regressions

As outlined in Proposition 3, our model implies that the volatility of bond returns is positively related to MBS convexity. We formulate a testable empirical prediction in the following hypothesis:

Hypothesis 3. *A regression of conditional volatility of bond returns, both implied and realized, on the negative convexity of MBS results in a positive slope coefficient for all maturities. Moreover, the associated volatility of volatility and variance risk premia are also increasing in MBS negative convexity.*

To test the above hypothesis, we run regressions of the following type:

$$\mathrm{tiv}_t/\mathrm{trv}_t/\mathrm{vrp}_t/\mathrm{straddle}_t/\mathrm{tivov}_t/\mathrm{trvov}_t = \beta_1 \mathrm{convexity}_t + \epsilon_t,$$

where tiv_t (trv_t) is the implied (realized) volatility of the 30-year Treasury bond at time t, vrp_t is the associated variance risk premium defined as the difference between the implied and realized measure, tivov_t (trvov_t) is the volatility of implied (realized) variance. The volatility of variance regressions are motivated by the results in Perli and Sack (2003) who find that mortgage hedging has not only a direct effect on the level of volatility but also on the changes in expected volatility, which is larger whenever hedging is intense. The results of these regressions are found in Table 2, Panel C.

The data confirm our model predictions. The estimated coefficients are significant and also carry the expected sign: Higher MBS convexity implies both a larger implied and realized volatility, a larger variance risk premium, and a larger volatility of variance. In economic terms, the regression results imply that for any one standard deviation increase in MBS convexity there is a 60 (40) basis point increase in implied (realized) volatility. MBS convexity also helps to explain returns of straddle strategies. The estimated coefficient for MBS convexity is significant, albeit only at the 10% confidence level.

3.4 What do MBS duration and convexity capture beyond information in yields?

Households refinance mortgages when interest rates drop and it is therefore tempting to assume that MBS duration is a mere reflection of information already contained in the yield curve. Despite the stylized nature of our model, Corollary 1 motivates us to take a closer look at this question. The spanning properties of the model can be illustrated with a simple calibration exercise in which we set the parameters to match the properties of the observed short rate and MBS duration series together with the predictive power of duration for one year excess returns on the 10-year bond (R^2=14.5%) reported in Table 1.[21] In the calibrated model the short rate factor explains over 97% of the variation in yields across maturities, but at the same time less than 9% of the variation in MBS duration. Moreover, the model implied R^2 for a predictive regression of one year excess returns on the short rate is only about 1%. This means that in the model the factor that accounts for the predictive power is not spanned by the factor that accounts for a dominant fraction of the cross sectional variation in yields. Finally, the theoretical correlations between duration and the 10-year yield, and between duration and the slope of the yield curve are 0.44 and -0.15, respectively.

Motivated by the above observations in the context of our stylized model, we look at the empirical relationship between MBS duration and the factors that explain most of the variation in yields. First, Table 3 (Panel A) reports the unconditional correlations between MBS duration and the first three principal components (PCs) from the cross section of yields, commonly known to represent its level, slope, and curvature. While the correlation between MBS duration and the first PC is 0.55, duration is almost uncorrelated with the second and third PCs. To test more formally whether duration is spanned by these yield factors we run the following regression similar to Joslin, Priebsch, and Singleton (2014):

$$\text{duration}_t = \alpha + \beta_1 \text{level}_t + \beta_2 \text{slope}_t + \beta_3 \text{curvature}_t + \epsilon_t,$$

[21] More precisely, we simply estimate the parameters in equations (1) and (7) that describe the dynamics of the short rate and duration, and choose α in equation (10) to match the predictive power of duration on excess returns.

where $\text{level}_t, \text{slope}_t$ and curvature_t are the first three PCs. For our sample period, the adjusted R^2 of this regression is equal to 25%. This means that 75% of the variation in MBS duration arises from risks which are distinct from these PCs. Moreover, the AR(1) coefficient of residuals is equal to 0.94 and the associated Durbin and Watson statistic to 0.1, which clearly rejects the null of zero autocorrelation.

This now begs the question of whether duration contains any information beyond these principal components to predict bond returns. We tackle this question by regressing bond excess returns on MBS duration and the first three PCs. Table 3 reports the results. The economic and statistical significance of the duration factor remains very close to the results reported in Table 1.

[Insert Table 3 here.]

It is also natural to investigate how MBS convexity is related to the level, slope, and curvature of the yield curve. Running a regression from MBS convexity onto the first three yield PCs, we find that the regressors explain around 40% of the variation. The AR(1) coefficient of residuals is 0.85 and thus significantly different from zero. We therefore conclude that—just as duration—convexity is not spanned by the first three yield factors. Since in our model, convexity is mainly related to bond yield volatility, we repeat the same exercise using the PCs calculated from the cross section of bond yield volatilities. Similar to the cross section of yields, three factors are essentially enough to fully capture the dynamics of bond yield volatilities: The first three factors explain 91.8%, 6.5% and 1.15% of the overall variation, respectively. Running a regression from convexity onto the first three volatility PCs, we find that they explain little of the variation in convexity as the adjusted R^2 is only about 10%. Hence, we conclude that MBS convexity is also not spanned by yield volatilities.

3.5 MBS duration and other predictors of bond returns

The evidence on bond return predictability by MBS duration is consistent with our theoretical predictions. However, it could also be the case that duration simply captures predictability stemming from sources that are not explicitly accounted for in our model. To the best of our knowledge, we are not aware of work that establishes, either theoretically or empirically, the correlation between MBS duration and some alternative predictor of bond returns. However, we control for this possibility by including the slope of the term structure and the CP factor, two well-known determinants of bond risk premia (see Cochrane and Piazzesi (2005)), in the predictability regression that now becomes:

$$rx_{t+1}^{\tau} = \beta_1^{\tau}\text{duration}_t + \beta_2^{\tau}\text{slope}_t + \beta_3^{\tau}\text{cp}_t + \epsilon_{t+1}.$$

Results are reported in Table 4. We find that including these additional regressors does not deteriorate the significance of duration, moreover, estimated slope coefficients on MBS duration remain remarkably stable. When we add the slope of the term structure and the CP factor to the regressions, the estimated coefficients on duration remain highly significant for maturities of five years and beyond. All three regressors combined explain between 21% and 46% of the time variation of annual bond excess returns.

[Insert Table 4 here.]

We conclude that there is a strong link between bond risk premia and MBS duration. The effect is more pronounced for longer maturity bonds and remains significant when we add other predictors to the regressions.

3.6 MBS duration and macro factors

Similarly, it could be the case that MBS duration depends on macroeconomic conditions and its predictive power is due to correlations with macro factors. Ludvigson and Ng (2009) exploit information in 132 different realized macroeconomic and financial series and explore the predictive content of these series for bond risk premia. The authors find that the main principal components extracted from this panel are statistically significant even in the presence of the CP factor and substantially improve predictability. To test whether our MBS duration factor is robust to the inclusion of these macroeconomic factors, we compute the eight static macroeconomic factors, F^j, $j = 1, \ldots, 8$, for an updated data set through December 2012. In the following, we report regressions from bond risk premia onto duration and the macroeconomic factors at the monthly frequency. Table 4 presents the results.

While the effect of duration at the shortest maturities is negligible, at longer maturities, the impact remains almost unaltered compared to the results for weekly returns reported in Table 1. For longer maturities, macroeconomic factors become less significant, unlike duration. More importantly, the significance of the duration factor is virtually unchanged when moving from the univariate regression results to the regressions which include the macro factors. We summarize our findings as follows: MBS duration is a strong predictor of bond risk premia at longer horizons and its predictive power is not subsumed by macroeconomic factors.

3.7 MBS convexity and other determinants of yield and bond return volatility

In this section we control for additional determinants of yield and bond return volatility that have been documented in the literature. It is well-known that volatility tends to increase in periods of high illiquidity (see, e.g., Hu, Pan, and Wang (2013)). In our multivariate specification, we therefore add a proxy for illiquidity and a proxy of fixed-income implied volatility, similar to the VIX in equity markets.[22] We run the following regression from conditional bond yield volatility onto convexity and a set of other predictors:

$$\mathrm{vol}_t^\tau = \beta_1^\tau \mathrm{convexity}_t + \beta_2^\tau \mathrm{illiq}_t + \beta_3^\tau \mathrm{tiv}_t + \epsilon_t,$$

where vol_t^τ is the conditional bond yield volatility at time t of a bond with maturity $\tau = 1, \ldots, 10$ years, illiq_t is the illiquidity factor at time t, and tiv_t is the Treasury-implied volatility at time t. Results are reported in Table 5. We find that when we add illiquidity into the regression, convexity still remains highly statistically significant. The estimated coefficients reveal that the effect is the largest for the intermediate maturities of two-three years as indicated by the size of the coefficient. Illiquidity has the expected positive

[22] Hu, Pan, and Wang (2013) document a strong link between their illiquidity proxy and a fixed-income implied volatility index, the Bank of America/Merrill Lynch MOVE index. Note that the MOVE is calculated from rather illiquid over-the-counter Treasury options while tiv_t is calculated using extremely liquid Treasury future options.

sign as bond volatility tends to be high when markets are illiquid. However, the effect becomes insignificant as we add lagged values of yield volatility into the regression. All three factors together explain between 50% and 63% of the time variation in bond yield volatility across different maturities. The same picture emerges for implied volatilities from swaptions: Estimated slope coefficients on convexity are robust to the inclusion of other regressors.

[Insert Tables 5]

We repeat the exercise for the volatility of bond returns for which we run regressions of the following type:

$$\text{tiv}_t/\text{trv}_t/\text{vrp}_t/\text{straddle}_t/\text{tivov}_t/\text{trvov}_t/\text{iv}^{\tau y 10y} = \beta_1 \text{convexity}_t + \beta_2 \text{illiq}_t + \beta_3 \text{lagged}_{t-1} + \epsilon_t,$$

where tiv_t (trv_t) is the implied (realized) volatility of the 30-year Treasury bond at time t, vrp_t is the associated variance risk premium defined as the difference between the implied and realized measure, tivov_t (trvov_t) is the volatility of implied (realized) variance, $\text{iv}^{\tau y 10y}$ is the τ-year maturity implied volatility from swaptions on the 10-year swap rate, illiq_t is the illiquidity measure and lagged_{t-1} are the LHS variables lagged for one period that are included given the high persistence. The results of these regressions are found in Tables 5.

When we add additional regressors, $\hat{\beta}_1$ remains significant. Illiquidity is significant and carries the expected positive sign: Higher illiquidity implies higher volatility, a higher associated variance risk premium, and higher returns on a straddle strategy.

Overall, we conclude that both MBS duration and convexity are significantly linked to first and second moments of bond yields and bond returns. The empirical results are in line with the theoretical predictions and also hold when adding other control variables to the regressions.

3.8 The impact of MBS duration and convexity over time

Figure 4 plots the ratio of outstanding mortgages and GDP from 1990 to 2012 together with the amount outstanding in mortgages and Treasuries. As one can see, the importance of the mortgage market vis-à-vis both GDP and Treasuries has increased over the past 20 years, although both ratios peak in 2010 and since then have somewhat declined. If mortgage markets have become more important over time, one would expect the effect of MBS duration and convexity to increase over time as well. To control for the variation in mortgage volume, we run similar regressions as before but interact the duration and convexity measures with the mortgage to GDP ratio. The results are reported in Table 6.

[Insert Figure 4 and Table 6 here.]

The results indicate that MBS duration and convexity are still highly significant predictors of bond excess returns and bond yield volatility. For the bond yield volatility regressions, we find that the interaction term is able to explain 24% of the time variation in long term bond yield volatility. Overall, we conclude that MBS duration and convexity are robust predictors for levels and second moments of bond yields.

3.9 Robustness

Interest rate swaps: Interest rate risk is primarily hedged in either the Treasury or interest rate swap market and the main focus in the previous section has been on Treasury data. The reason for this is twofold. First, interest rate swap data contain a considerable credit risk component (see Feldhütter and Lando (2008)) which is outside the scope of our model to explain. Second, after the Lehman default in 2008, prices of interest rate swaps (especially at longer maturities) got possibly distorted due to a decline in arbitrage capital (see Krishnamurthy (2010)). In particular, our data sample also covers the time period where the swap spread, defined as the difference between the fixed rate on a fixed-for-floating 10-year swap and 10-year Treasury rate, turned negative. This is another feature in the data that goes beyond what the model is designed to capture.

For robustness reasons, we also run bond risk premia regressions using swap rather than Treasury data and we report estimated coefficients in Table 7.[23] We note that the size and significance of the estimated coefficients are almost identical to those reported for Treasuries. Adding explanatory factors such as the slope of the term structure or the CP factor does not deteriorate the significance of MBS duration.

[Insert Table 7 here.]

Bond portfolios: One issue with using annual bond excess returns is the short sample period. In our data sample of 16 years, we have a maximum of 16 independent observations. To address this issue, we use actual bond returns for different maturity bins available from CRSP. Data is monthly and represents an equally-weighted average of holding period returns for each bond in the portfolio. We calculate excess returns by subtracting the T-bill rate. Because of the large impact of monetary policy on T-bills in the past couple of years, we also use the one-month Eurodollar deposit rate as an alternative as suggested by Duffee (1996). Results are reported in Table 8.

[Insert Table 8 here.]

We note that while the adjusted R^2s are almost halved compared to the previous results, the estimated coefficients for duration are still highly significant, even in the multivariate regressions. Moreover, these results hold whether we use the T-bill or the Eurodollar rate to construct the excess returns.

4 Conclusion

This paper studies the feedback from MBS risk on bond risk premia and interest rate volatility. We build an equilibrium model of bond supply shocks driven by changes in MBS duration and embed it into an otherwise standard one factor term structure model. Despite its simple structure, our model has interesting implications for first and second

[23] We bootstrap a zero coupon curve from swap rates and calculate excess returns that are directly comparable to the Treasury excess returns we use in the benchmark results.

moments of interest rates: Our model is able to replicate the predictive power of MBS duration for bond excess returns and can accommodate a hump shaped term structure of bond yield volatilities.

We then empirically test our model predictions and find our hypotheses confirmed. For example, we find a strong positive link between MBS duration and bond excess returns. The relationship is not only statistically significant but also economically relevant. This relationship remains highly significant and stable when we add other standard regressors. We then proceed to study the relationship between MBS hedging and bond yield volatilities and bond variance risk premia. In line, with our theoretical predictions, we find that MBS convexity significantly affects bond yield volatilities and it produces the predicted hump shaped feature. A higher MBS convexity not only significantly increases yield volatility but also measures of bond return volatility, and bond variance risk premia.

While MBS duration and convexity are naturally related to information in the term structure of bond yields, we provide novel evidence that duration and convexity are not spanned by the usual bond yield factors (principal components), as well as theoretical motivation why this could be the case. An investigation of supply factors within a multi-factor reduced form term structure model and their relation to higher order yield factors is an exciting avenue which we leave to future research.

References

AGARWAL, S., J. C. DRISCOLL, AND D. I. LAIBSON (2013): "Optimal Mortgage Refinancing: A Closed Form Solution," *Journal of Money, Credit, and Banking*, 45, 591–622.

AMROMIN, G., J. HUANG, AND C. SIALM (2007): "The Tradeoff Between Mortgage Prepayments and Tax-Deferred Retirement Savings," *Journal of Public Economics*, 91, 2014–2040.

BANSAL, R., W. J. COLEMAN, AND C. T. LUNDBLAD (2011): "Endogenous Liquidity Supply," Working Paper, Duke University.

BANSAL, R., AND I. SHALIASTOVICH (2013): "A Long-Run Risk Explanation of Predictability Puzzles in Bond and Currency Markets," *Review of Financial Studies*, 26, 1–33.

BEGENAU, J., M. PIAZZESI, AND M. SCHNEIDER (2013): "Banks' Risk Exposure," Working Paper, Stanford University.

BOUDOUKH, J., R. F. WHITELAW, M. RICHARDSON, AND R. STANTON (1997): "Pricing Mortgage-Backed Securities in a Multifactor Interest Rate Environment: A Multivariate Density Estimation Approach," *Review of Financial Studies*, 10, 405–446.

BOYARCHENKO, N., A. FUSTER, AND D. O. LUCCA (2014): "Understanding Mortgage Spreads," Working Paper, Federal Reserve Bank of New York.

BRENNAN, M. J., AND E. S. SCHWARTZ (1985): "Determinants of GNMA Mortgage Prices," *Journal of the American Real Estate and Urban Economics Association*, 13, 209–228.

CAMPBELL, J. Y. (2006): "Household Finance," *Journal of Finance*, 61, 1553–1604.

CHANG, Y., D. MCMANUS, AND B. RAMAGOPAL (2005): "Does Mortgage Hedging Raise Long-Term Interest Rate Volatility?," *Journal of Fixed Income*, 14, 57–66.

CHENG, P., AND O. SCAILLET (2007): "Linear-Quadratic Jump Diffusion Modelling," *Mathematical Finance*, 17, p. 575 – 598.

COCHRANE, J. H., AND M. PIAZZESI (2005): "Bond Risk Premia," *American Economic Review*, 95, 138–160.

COLLIN-DUFRESNE, P., AND R. S. GOLDSTEIN (2002): "Do Bonds Span the Fixed Income Markets? Theory and Evidence for Unspanned Stochastic Volatility," *Journal of Finance*, 57, 1685–1730.

COLLIN-DUFRESNE, P., AND J. P. HARDING (1999): "A Closed-Form Formula for Valuing Mortgages," *Journal of Real Estate Finance and Economics*, 19, 133–146.

DAI, Q., AND K. SINGLETON (2003): "Term Structure Dynamics in Theory and Reality," *Review of Financial Studies*, 16, 631–678.

D'AMICO, S., W. ENGLISH, D. LOPEZ-SALIDO, AND E. NELSON (2012): "The Federal Rerseves Large Scale Asset Purchase Programs: Rationale and Effects," *Economic Journal*, 122, 415–446.

D'AMICO, S., AND T. B. KING (2013): "Flow and Stock Effects of Large-Scale Treasury Purchases: Evidence on the Importance of Local Supply," *Journal of Financial Economics*, 108, 425–448.

Duarte, J. (2008): "The Causal Effect of Mortgage Refinancing on Interest Rate Volatility: Empirical Evidence and Theoretical Implications," *Review of Financial Studies*, 21, 1698–1731.

Duffee, G. R. (1996): "Idiosyncratic Variation of Treasury Bill Yields," *Journal of Finance*, 51(2), 527–551.

Duffee, G. R. (2011): "Information in (and not in) the term structure," *Review of Financial Studies*, 24(9), 2895–2934.

Dunn, K. B., and J. J. McConnell (1981): "Valuation of GNMA Mortgage-Backed Securities," *Journal of Finance*, 36, 599–616.

Dybvig, P. H., J. E. Ingersoll, and S. A. Ross (1996): "Long Forward and Zero-Coupon Rates Can Never Fall," *Journal of Business*, 69, 1–25.

Ehling, P., M. F. Gallmeyer, C. Heyerdahl-Larsen, and P. K. Illeditsch (2013): "Disagreement about Inflation and the Yield Curve," Working Paper, Norwegian Business School.

Feldhütter, P., and D. Lando (2008): "Decomposing Swap Spreads," *Journal of Financial Economics*, 88, 375–405.

Fenou, B., and J.-S. Fontaine (2013): "Gaussian Term Structure Models and Bond Risk Premia," Working Paper, Bank of Canada.

Gabaix, X., A. Krishnamurthy, and O. Vigneron (2007): "Limits of Arbitrage: Theory and Evidence from Mortgage-Backed Securities Market," *Journal of Finance*, 62, 557–595.

Gagnon, J. E., M. Raskin, J. Remache, and B. P. Sack (2010): "Large-Scale Asset Purchases by the Federal Reserve: Did They Work?," Working Paper, Board of Governors of the Federal Reserve System.

Greenwood, R., S. G. Hanson, and J. C. Stein (2010): "A Gap-Filling Theory of Corporate Debt Maturity Choice," *Journal of Finance*, 65, 993–1028.

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

Gürkaynak, R., B. Sack, and J. Wright (2007): "The U.S. Treasury Yield Curve: 1961 to Present," *Journal of Monetary Economics*, 54, 2291–2304.

Hancock, D., and W. Passmore (2010): "Did the Federal Reserve's MBS Purchase Program Lower Mortgage Rates?," Working Paper, Board of Governors of the Federal Reserve System.

Hanson, S. G. (2013): "Mortgage Convexity," forthcoming, Journal of Financial Economics.

Hong, H., D. Sraer, and J. Yu (2013): "Reaching for Maturity," Working Paper, Princeton University.

Hu, G. X., J. Pan, and J. Wang (2013): "Noise as Information for Illiquidity," *Journal of Finance*, 68, 2223–2772.

Joslin, S., A. Le, and K. J. Singleton (2013): "Gaussian Macro-Finance Term Structure Models with Lags," *Journal of Financial Econometrics*, 11, 581–609.

Joslin, S., M. Priebsch, and K. J. Singleton (2014): "Risk Premiums in Dynamic Term Structure Models with Unspanned Macro Risks," *forthcoming, Journal of Finance*.

Krishnamurthy, A. (2010): "How Debt Markets Have Malfunctioned in the Crisis," *Journal of Economic Perspectives*, 24, 3–28.

Krishnamurthy, A., and A. Vissing-Jorgensen (2011): "The Effects of Quantitative Easing on Interest Rates: Channels and Implications for Policy," *Brookings Papers on Economic Activity*, 2, 215–287.

——— (2012): "The Aggregate Demand for Treasury Debt," *Journal of Political Economy*, 120, 233–267.

Landier, A., D. Sraer, and D. Thesmar (2013): "Banks' Exposure to Interest Rate Risk and the Transmission of Monetary Policy," Working Paper, Toulouse School of Economics.

Le, A., and K. J. Singleton (2013): "The Structure of Risks in Equilibrium Affine Models of Bond Yields," Working Paper, UNC.

Le, A., K. J. Singleton, and Q. Dai (2010): "Discrete-Time Affine Q Term Structure Models with Generalized Market Prices of Risk," *Review of Financial Studies*, 23, 2184 – 2227.

Li, C., and M. Wei (2013): "Term Structure Modeling with Supply Factors and the Federal Reserve's Large Scale Asset Purchase Programs," Working Paper, Board of Governors of the Federal Reserve System.

Longstaff, F. (2005): "Borrower Credit and the Valuation of Mortgage-Backed Securities," *Real Estate Economics*, 33, 619–661.

Lou, D., H. Yan, and J. Zhang (2013): "Anticipted and Repeated Shocks in Liquid Markets," *Review of Financial Studies*, 26, 1891–1912.

Ludvigson, S. C., and S. Ng (2009): "Macro Factors in Bond Risk Premia," *Review of Financial Studies*, 22, 5027–5067.

Mueller, P., A. Vedolin, and Y.-M. Yen (2012): "Bond Variance Risk Premiums," Working Paper, London School of Economics.

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

Perli, R., and B. P. Sack (2003): "Does mortgage hedging amplify movements in long-term interest rates?," *Journal of Fixed Income*, 13, 7–17.

Richard, S. F., and R. Roll (1989): "Prepayments on Fixed-rate Mortgage-Backed Securities," *Journal of Portfolio Management*, 15, 73–82.

Schwartz, E. S., and W. Torous (1989): "Prepayment and the Valuation of Mortgage-Backed Securities," *Journal of Finance*, 44, 375–392.

Securities Industry and Financial Markets Association (2013): "Data Statistics," http://www.sifma.org/research/statistics.aspx.

Soo, C. K. (2014): "Quantifying Animal Spirits: News Media and Sentiment in the Housing Market," Working Paper, University of Michigan.

Stanton, R. S. (1995): "Rational Prepayment and the Valuation of Mortgage-Backed Securities," *Review of Financial Studies*, 8, 677–708.

Stanton, R. S., and N. Wallace (1998): "Mortgage Choice: What's the Point?," *Real Estate Economics*, 26, 173–205.

Stock, J. H., and M. W. Watson (2007): "Introduction to Econometrics," Addison Wesley.

Stock, J. H., J. Wright, and M. Yogo (2002): "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments," *Journal of Business and Economic Statistics*, 20, 518–529.

Svensson, L. (2004): "Estimating and Interpreting Forward Rates: Sweden 1992–1994," National Bureau of Economic Research Working Paper No. 4871.

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

Vayanos, D., and J.-L. Vila (2009): "A Preferred-Habitat Model of the Term Structure of Interest Rates," Working Paper, London School of Economics.

Vickery, J., and J. Wright (2010): "TBA Trading and Liquidity in the Agency MBS Market," Working Paper, Federal Reserve Bank of New York.

Wooldridge, P. D. (2001): "The Emergence of New Benchmark Yield Curves," BIS Quarterly Review.

Xiong, W., and H. Yan (2010): "Heterogeneous Expectations and Bond Markets," *Review of Financial Studies*, 23, 1433–1466.

Appendix A Preliminary results

Appendix A.1 Properties of useful functions

This appendix introduces five functions necessary to derive our main results and studies their properties.

Lemma 2. *For any $\tau > 0$, the function $g_1(x) \equiv \frac{1-e^{-x\tau}}{x\tau}$ for all $x \neq 0$ and $g_1(0) \equiv 1$ is positive, decreasing, and convex for all $x \in \mathbb{R}$. Moreover, for arbitrary $x, y \in \mathbb{R}$,*

$$\frac{1-e^{-x\tau}}{x\tau} + \left(-\frac{1-e^{-x\tau}-x\tau e^{-x\tau}}{x^2\tau}\right)(y-x) < \frac{1-e^{-y\tau}}{y\tau}. \tag{A-1}$$

Proof. The derivative of g_1 is given by

$$g_1'(x) = -\frac{1-e^{-x\tau}-x\tau e^{-x\tau}}{x^2\tau}, \tag{A-2}$$

which has the opposite sign as $F(x) \equiv 1-e^{-x\tau}-x\tau e^{-x\tau}$. But $\lim_{x\to 0} F(x) = 0$, and $F'(x) = x\tau^2 e^{-x\tau}$, which is negative for $x < 0$ and positive for $x > 0$. Hence, $F(x) \geq 0$ for all $x \in \mathbb{R}$, and $g_1'(x) \leq 0$ for all $x \in \mathbb{R}$; g_1 is a decreasing function. However, the limit of g_1 when $x \to \infty$ is $\lim_{x\to\infty} g_1(x) = 0$. Since g_1 is a decreasing function and it converges to zero when $x \to \infty$, it must be that $g_1(x) > 0$ for all $x \in \mathbb{R}$.

Regarding convexity, (A-2) implies

$$g_1''(x) = \frac{2e^{-x\tau}}{x^3\tau}\left[e^{x\tau} - \left(1 + x\tau + \frac{x^2\tau^2}{2}\right)\right],$$

but $1 + z + \frac{z^2}{2}$ are the first three terms of the power series of e^z, and it is well-known that $1 + z + \frac{z^2}{2} < e^z$ for $z > 0$ and $1 + z + \frac{x^2}{2} > e^z$ for $z < 0$. Therefore, $g_1''(x) > 0$ and thus g_1 is convex.

Finally, convexity of g_1 is equivalent to the function lying above all of its tangents. From (A-2), (A-1) is describing exactly this inequality for the point of tangency x and an arbitrary y. □

Lemma 3. *For any $h > 0$ and $\kappa \in \mathbb{R}$, the function*

$$g_2(x) \equiv \frac{e^{-\kappa h} - e^{-xh}}{\kappa - x} \tag{A-3}$$

is negative and increasing for all $x \in \mathbb{R}$.

Proof. We have $e^{-\kappa h} > e^{-xh}$ iff $\kappa < x$, which implies negativity. Also, differentiation gives

$$g_2'(x) = \frac{e^{-(\kappa-x)h} - [1-(\kappa-x)h]}{(\kappa-x)^2} e^{-xh}.$$

but since $e^z > 1 + z$ for all $z \in \mathbb{R}$, we have $e^{-(\kappa-x)h} \geq 1-(\kappa-x)h$ and thus $g_2'(x) \geq 0$ for all x, which concludes the proof. □

Lemma 4. *Suppose $\kappa, h > 0$ constants that satisfy $\kappa h < 1$. The function*

$$g_3(x) \equiv \frac{\kappa e^{-\kappa h} - x e^{-xh}}{\kappa - x} \tag{A-4}$$

is positive and decreasing for all $x < \frac{1}{h}$.

Proof. It is easy to confirm that the function $G(x) = x e^{-xh}$ satisfies (i) $G(0) = 0$ and $G(x) > 0$ iff $x > 0$, and (ii) $G'(x) \geq 0$ for $x \leq \frac{1}{h}$ and $G'(x) < 0$ otherwise. Therefore, as long as $\kappa, x < \frac{1}{h}$, the numerator and the denominator of (A-4) have the same sign and thus $g_3(x) > 0$. Next we differentiate (A-4) to obtain

$$\frac{dg_3}{dx} = \frac{\left(\kappa e^{-\kappa h} - x e^{-xh}\right) - (1 - xh)(\kappa - x) e^{-xh}}{(\kappa - x)^2}. \tag{A-5}$$

Denoting the numerator of (A-5) by $H(x)$ and differentiating, we obtain

$$H'(x) = (\kappa - x) h (2 - xh) e^{-xh}$$

while $H(\kappa) = 0$. Hence, $H(x)$ increases for $x < \kappa$ where it reaches zero, and afterwards it decreases. That is, the numerator of (A-5) is negative for all $x < \frac{1}{h}$ and so is $\frac{dg_3}{dx}$; g_3 is decreasing. □

Lemma 5. *For any $\tau > 0$, the function*

$$g_4(x, y) \equiv \frac{g_1(x) - g_1(y)}{x - y} = \frac{\frac{1 - e^{-x\tau}}{x\tau} - \frac{1 - e^{-y\tau}}{y\tau}}{x - y}, \tag{A-6}$$

$x, y \in \mathbb{R}$, is symmetric, negative, and increasing in both arguments. Moreover, if $x < x' < y' < y$ while $x + y = x' + y'$, $g_4(x, y) < g_4(x', y')$.

Proof. Lemma 2 implies that the numerator of g_4 is positive if and only if $x < y$, hence $g_4(x, y) < 0$ for all $x, y \in \mathbb{R}$. Symmetry, i.e. $g_4(x, y) = g_4(y, x)$, is obvious, and for g_4 being increasing, it means we only need to show that $\frac{dg_4}{dx} > 0$ for a fixed y. Differentiating (A-6) w.r.t. x, we have

$$\frac{dg_4}{dx} = -\frac{1}{(x - y)^2}\left[\frac{1 - e^{-x\tau}}{x\tau} - \frac{1 - e^{-x\tau} - x\tau e^{-x\tau}}{x^2\tau}(y - x) - \frac{1 - e^{-y\tau}}{y\tau}\right].$$

Lemma 2 also implies that the term inside the bracket is negative, and hence g_4 is increasing in x and y. Moreover, for an arbitrary constant y we have $\lim_{x\to\infty} g_4(x, y) = 0$, so if g_4 is increasing in x, it must be that for all $x, y \in \mathbb{R}$, $g_4(x, y) < 0$. □

Lemma 6. *Fix $\tau > 0$. The function*

$$g_5(x, y) = \frac{1 - e^{-x\tau}}{x} + \frac{x(y - x)}{y}\left(-\frac{1 - e^{-x\tau} - x\tau e^{-x\tau}}{x^2}\right) - \frac{1 - e^{-y\tau}}{y}, \tag{A-7}$$

$x, y \in \mathbb{R}^+$, satisfies $g_5(x, y) = 0$ if $x = y$ and $g_5(x, y) > 0$ whenever $x \neq y$.

Proof. If $x = y$, the first and last terms of g_5 are equivalent and the middle one is zero, hence $g_5(x, y) = 0$. Next we differentiate g_5 with respect to y while keeping x fixed to obtain

$$\frac{dg_5(x,y)}{dy} = \frac{(1 - e^{-y\tau} - y\tau e^{-y\tau}) - (1 - e^{-x\tau} - x\tau e^{-x\tau})}{y^2} = \frac{F(y) - F(x)}{y^2},$$

where F is defined in the proof of Lemma 2. As shown there, F is increasing on $\mathbb{R}^+$, so $0 < x < y$ implies the numerator is positive and thus $\frac{dg_5(x,y)}{dy} > 0$. On the other hand, $0 < y < x$ implies the numerator is negative and $\frac{dg_5(x,y)}{dy} < 0$. Therefore, g_5 is decreasing in y before x, reaches zero, then increasing, i.e., is positive for all $y \neq x$. □

Appendix A.2 Covariance matrix of $(r_t, D_t)^\top$

In this appendix we derive the variance-covariance matrix of $(r_t, D_t)^\top$ under $\mathbb{P}$ from (1) and (11). Following the standard technique, applying Ito's lemma to $e^{\kappa t} r_t$ and combining it with (1), we obtain

$$d\left(e^{\kappa t} r_t\right) = \left[\kappa e^{\kappa t} r_t + e^{\kappa t}\kappa(\theta - r_t)\right] dt + e^{\kappa t}\sigma dB_t = e^{\kappa t}\kappa\theta dt + e^{\kappa t}\sigma dB_t.$$

Integrating both sides between zero and t and rearranging gives

$$r_t = \theta\left(1 - e^{-\kappa t}\right) + r_0 e^{-\kappa t} + \sigma e^{-\kappa t}\int_0^t e^{\kappa s} dB_s. \tag{A-8}$$

As r_t shows up in the drift of D_t under P whenever $\delta_r \neq 0$, we cannot simply use the same calculation for D_t. Instead, let us define

$$\bar{D}_t = D_t + \frac{\delta_r}{\delta_D - \kappa} r_t; \tag{A-9}$$

we then have

$$d\bar{D}_t = dD_t + \frac{\delta_r}{\delta_D - \kappa} dr_t = \delta_D\left(\bar{\theta}_D - \bar{D}_t\right) dt + \bar{\sigma}_D dB_t,$$

where in the last step we use (15) and introduce the notation

$$\bar{\theta}_D = \theta_D + \frac{\delta_r}{\delta_D - \kappa}\theta \text{ and } \bar{\sigma}_D = \eta_y\sigma_y^{\bar{\tau}} + \frac{\delta_r}{\delta_D - \kappa}\sigma = \delta_D\frac{\eta_y\sigma_y^{\bar{\tau}}}{\delta_D - \kappa}.$$

That is, $\bar{D}_t$ is a Vasicek process with a speed of mean reversion δ_D. Applying the same steps as for r_t, we also obtain

$$\bar{D}_t = \bar{\theta}_D\left(1 - e^{-\delta_D t}\right) + \bar{D}_0 e^{-\delta_D t} + \bar{\sigma}_D e^{-\delta_D t}\int_0^t e^{\delta_D s} dB_s, \tag{A-10}$$

where, importantly, the Brownian increments dB_s are the same as in (A-8).

From (A-8) and (A-10) we can compute the conditional means and variances of r and D and the conditional covariance between them t time ahead. First, as the expectation of the increments of the Brownian motion is zero, we have

$$E_0\left[r_t\right]=\theta\left(1-e^{-\kappa t}\right)+r_0 e^{-\kappa t} \text{ and } E_0\left[\bar{D}_t\right]=\bar{\theta}_D\left(1-e^{-\delta_D t}\right)+\bar{D}_0 e^{-\delta_D t}; \tag{A-11}$$

this, together with (A-10), also yields

$$\begin{aligned} E_0\left[D_t\right] &= E_0\left[\bar{D}_t\right]-\frac{\delta_r}{\delta_D-\kappa} E_0\left[r_t\right] \\ &= \bar{\theta}_D\left(1-e^{-\delta_D t}\right)+D_0 e^{-\delta_D t}-\frac{\delta_r}{\delta_D-\kappa}\left[\theta\left(1-e^{-\kappa t}\right)+r_0\left(e^{-\kappa t}-e^{-\delta_D t}\right)\right]. \end{aligned} \tag{A-12}$$

Next, from (A-8), we have

$$Var_0\left[r_t\right]=\sigma^2 e^{-2\kappa t} E_0\left[\left(\int_0^t e^{\kappa s} dB_s\right)^2\right]=\sigma^2 e^{-2\kappa t} \int_0^t e^{2\kappa s} ds=\frac{\sigma^2}{2\kappa}\left(1-e^{-2\kappa t}\right), \tag{A-13}$$

where we use that $dB_s \sim N(0, ds)$ i.i.d. over time. Similarly, from (A-10) we obtain

$$Var_0\left[\bar{D}_t\right]=\bar{\sigma}_D^2 e^{-2\delta_D t} E_0\left[\left(\int_0^t e^{\delta_D s} dB_s\right)^2\right]=\bar{\sigma}_D^2 e^{-2\delta_D t} \int_0^t e^{2\delta_D s} ds=\frac{\bar{\sigma}_D^2}{2\delta_D}\left(1-e^{-2\delta_D t}\right). \tag{A-14}$$

Finally, for the covariance, we have

$$\begin{aligned} Cov_0\left[r_t, \bar{D}_t\right] &= \sigma\bar{\sigma}_D e^{-(\kappa+\delta_D)t} E_0\left[\int_0^t e^{\kappa s} dB_s \int_0^t e^{\delta_D s} dB_s\right] \\ &= \sigma\bar{\sigma}_D e^{-(\kappa+\delta_D)t} \int_0^t e^{(\kappa+\delta_D)s} ds=\frac{\sigma\bar{\sigma}_D}{\kappa+\delta_D}\left(1-e^{-(\kappa+\delta_D)t}\right). \end{aligned} \tag{A-15}$$

From (A-9), (A-13), and (A-15) we then have

$$\begin{aligned} Cov_0\left[r_t, D_t\right] &= Cov_0\left[r_t, \bar{D}_t\right]-\frac{\delta_r}{\delta_D-\kappa} Var_0\left[r_t\right] \\ &= \frac{\sigma\bar{\sigma}_D}{\kappa+\delta_D}\left(1-e^{-(\kappa+\delta_D)t}\right)-\frac{\delta_r}{\delta_D-\kappa}\frac{\sigma^2}{2\kappa}\left(1-e^{-2\kappa t}\right), \end{aligned} \tag{A-16}$$

and from (A-9), (A-13), (A-14), and (A-16) we get

$$\begin{aligned} &Var_0\left[D_t\right]=Var_0\left[\bar{D}_t\right]-2\frac{\delta_r}{\delta_D-\kappa} Cov_0\left[r_t, D_t\right]-\left(\frac{\delta_r}{\delta_D-\kappa}\right)^2 Var_0\left[r_t\right] \\ &=\frac{\bar{\sigma}_D^2}{2\delta_D}\left(1-e^{-2\delta_D t}\right)-2\frac{\delta_r}{\delta_D-\kappa}\frac{\sigma\bar{\sigma}_D}{\kappa+\delta_D}\left(1-e^{-(\kappa+\delta_D)t}\right)+\left(\frac{\delta_r}{\delta_D-\kappa}\right)^2\frac{\sigma^2}{2\kappa}\left(1-e^{-2\kappa t}\right). \end{aligned} \tag{A-17}$$

A notable special case that we use to determine the coefficients of the predictive regressions is the unconditional variance-covariance matrix of $(r_t, D_t)^\top$. Taking the limit $t \to \infty$ in (A-13), (A-16) and (A-17) yields

$$V = \begin{pmatrix} \frac{\sigma^2}{2\kappa} & \frac{\delta_r \sigma^2}{2\kappa(\kappa+\delta_D)} \\ \frac{\delta_r \sigma^2}{2\kappa(\kappa+\delta_D)} & \frac{\delta_r^2 \sigma^2}{2\kappa^2(\kappa+\delta_D)} \end{pmatrix}, \tag{A-18}$$

implying that in general when $\delta_D \neq 0$ the two factors are not collinear.

Appendix B Proofs and derivations

Proof of Lemma 1. For notational simplicity let us write bond prices in the form

$$\frac{d\Lambda_t^\tau}{\Lambda_t^\tau} = \mu_t^\tau dt - \sigma_t^\tau dB_t. \tag{A-19}$$

Substituting (A-19) into intermediaries' budget constraint, (2), we get

$$dW_t = \left[r_t W_t + \int_0^T x_t^\tau \Lambda_t^\tau \left(\mu_t^\tau - r_t\right) d\tau\right] dt - \left[\int_0^T x_t^\tau \Lambda_t^\tau \sigma_t^\tau d\tau\right] dB_t,$$

therefore (3) simplifies to

$$\max_{\{x_t^\tau\}_{\tau\in(0,T]}} \int_0^T x_t^\tau \Lambda_t^\tau \left(\mu_t^\tau - r_t\right) d\tau - \frac{\alpha}{2}\left[\int_0^T x_t^\tau \Lambda_t^\tau \sigma_t^\tau d\tau\right]^2. \tag{A-20}$$

Because markets are complete, by no-arbitrage, there exists a unique market price of interest rate risk across all bonds that satisfies

$$\lambda_t = \frac{E_t\left(\frac{d\Lambda_t^\tau}{\Lambda_t^\tau}\right)/dt - r_t}{\frac{1}{\Lambda_t^\tau}\frac{d\Lambda_t^\tau}{dr_t}\sigma} = \frac{\mu_t^\tau - r_t}{-\sigma_t^\tau}, \tag{A-21}$$

and introducing

$$x_t = \frac{d\left(\int_0^T x_t^\tau \Lambda_t^\tau d\tau\right)}{dr_t} = \int_0^T x_t^\tau \frac{d\Lambda_t^\tau}{dr_t} d\tau = -\frac{1}{\sigma}\int_0^T x_t^\tau \Lambda_t^\tau \sigma_t^\tau d\tau, \tag{A-22}$$

for the total exposure of interest rate risk borne by intermediaries, their maximization problem (A-20) reduces to

$$\max_{x_t} \lambda_t x_t - \frac{\alpha\sigma}{2} x_t^2. \tag{A-23}$$

The first order condition of (A-23) together with the market clearing condition (4) determine the equilibrium market price of risk and provides (8). $\square$

Proof of Theorem 1. We conjecture that equilibrium yields in the model defined by (12) and (13) are in the form (14), i.e., bond prices are

$$\Lambda_t^\tau = e^{-[\tau A(\tau)+\tau B(\tau) r_t + \tau C(\tau) D_t]}. \tag{A-24}$$

Applying Ito's Lemma to (A-24), substituting in (12) and (13), and imposing the condition that the bond price drift under $\mathbb{Q}$ must be $r_t \Lambda_t^\tau dt$, we obtain an equation affine in the factors r_t and D_t. Collecting the r_t, D_t, and constant terms, respectively, we get a set of ODEs:

$$\begin{aligned}
1 &= \tau\mathcal{B}'(\tau) + \mathcal{B}(\tau) + \kappa\tau\mathcal{B}(\tau) + \delta_r\tau\mathcal{C}(\tau), && \text{(A-25)}\\
0 &= \tau\mathcal{C}'(\tau) + \mathcal{C}(\tau) + \delta_D^{\mathbb{Q}}\tau\mathcal{C}(\tau) - \alpha\sigma\sigma_y^{\bar{\tau}}\tau\mathcal{B}(\tau), \text{ and} && \text{(A-26)}\\
0 &= \tau\mathcal{A}'(\tau) + \mathcal{A}(\tau) - \kappa\theta\tau\mathcal{B}(\tau) - \delta_0\tau\mathcal{C}(\tau) + \frac{1}{2}\sigma^2\tau^2\mathcal{B}^2(\tau) + \frac{1}{2}\eta_y^2\left(\sigma_y^{\bar{\tau}}\right)^2\tau^2\mathcal{C}(\tau)^2. && \text{(A-27)}
\end{aligned}$$

with terminal conditions $\mathcal{A}(0) = \mathcal{C}(0) = 0$ and $\mathcal{B}(0) = 1$. Combining (A-25) and (A-26), we write the following second order ODE for $\mathcal{C}$:

$$0 = \tau\mathcal{C}''(\tau) + 2\mathcal{C}'(\tau) + \left(\kappa + \delta_D^{\mathbb{Q}}\right)\left(\tau\mathcal{C}'(\tau) + \mathcal{C}(\tau)\right) + \left(\kappa\delta_D^{\mathbb{Q}} + \alpha\sigma\sigma_y^{\bar{\tau}}\delta_r\right)\tau\mathcal{C}(\tau) - \alpha\sigma\sigma_y^{\bar{\tau}}. \quad \text{(A-28)}$$

Solving (A-28) for $\mathcal{C}$, from there deriving $\mathcal{B}$ and $\mathcal{A}$, and applying the terminal conditions, yields the following solution:

$$\mathcal{C}(\tau) = -\frac{\alpha\sigma\sigma_y^{\bar{\tau}}}{(\kappa+\varepsilon) - \left(\delta_D^{\mathbb{Q}} - \varepsilon\right)}\left[\frac{1 - e^{-(\kappa+\varepsilon)\tau}}{(\kappa+\varepsilon)\tau} - \frac{1 - e^{-\left(\delta_D^{\mathbb{Q}} - \varepsilon\right)\tau}}{\left(\delta_D^{\mathbb{Q}} - \varepsilon\right)\tau}\right], \quad \text{(A-29)}$$

$$\mathcal{B}(\tau) = \frac{1 - e^{-(\kappa+\varepsilon)\tau}}{(\kappa+\varepsilon)\tau} - \frac{\varepsilon}{(\kappa+\varepsilon) - \left(\delta_D^{\mathbb{Q}} - \varepsilon\right)}\left[\frac{1 - e^{-(\kappa+\varepsilon)\tau}}{(\kappa+\varepsilon)\tau} - \frac{1 - e^{-\left(\delta_D^{\mathbb{Q}} - \varepsilon\right)\tau}}{\left(\delta_D^{\mathbb{Q}} - \varepsilon\right)\tau}\right] \quad \text{(A-30)}$$

and

$$\begin{aligned}
\mathcal{A}(\tau) = {} & \frac{1}{(\kappa+\varepsilon)}\left[\kappa\theta\frac{(\kappa+\varepsilon) - \delta_D^{\mathbb{Q}}}{(\kappa+\varepsilon) - \left(\delta_D^{\mathbb{Q}} - \varepsilon\right)} - \delta_0\frac{\alpha\sigma\sigma_y^{\bar{\tau}}}{(\kappa+\varepsilon) - \left(\delta_D^{\mathbb{Q}} - \varepsilon\right)}\right]\omega\left((\kappa+\varepsilon)\tau\right) \qquad \text{(A-31)} \\
& + \frac{1}{\left(\delta_D^{\mathbb{Q}} - \varepsilon\right)}\left[\kappa\theta\frac{\varepsilon}{(\kappa+\varepsilon) - \left(\delta_D^{\mathbb{Q}} - \varepsilon\right)} + \delta_0\frac{\alpha\sigma\sigma_y^{\bar{\tau}}}{(\kappa+\varepsilon) - \left(\delta_D^{\mathbb{Q}} - \varepsilon\right)}\right]\omega\left(\left(\delta_D^{Q} - \varepsilon\right)\tau\right) \\
& + \frac{1}{2}\sigma^2\frac{\left[(\kappa+\varepsilon) - \delta_D^{\mathbb{Q}}\right]^2 + \left[\alpha\eta_y\left(\sigma_y^{\bar{\tau}}\right)^2\right]^2}{(\kappa+\varepsilon)^2\left[(\kappa+\varepsilon) - \left(\delta_D^{\mathbb{Q}} - \varepsilon\right)\right]^2}\left[\frac{1}{2}\omega\left(2(\kappa+\varepsilon)\tau\right) - \omega\left((\kappa+\varepsilon)\tau\right)\right] \\
& + \frac{1}{2}\sigma^2\frac{\varepsilon^2 + \left[\alpha\eta_y\left(\sigma_y^{\bar{\tau}}\right)^2\right]^2}{\left(\delta_D^{\mathbb{Q}} - \varepsilon\right)^2\left[(\kappa+\varepsilon) - \left(\delta_D^{\mathbb{Q}} - \varepsilon\right)\right]^2}\left[\frac{1}{2}\omega\left(2\left(\delta_D^{\mathbb{Q}} - \varepsilon\right)\tau\right) - \omega\left(\left(\delta_D^{\mathbb{Q}} - \varepsilon\right)\tau\right)\right] \\
& + \sigma^2\frac{\left[(\kappa+\varepsilon) - \delta_D^{\mathbb{Q}}\right]\varepsilon - \left[\alpha\eta_y\left(\sigma_y^{\bar{\tau}}\right)^2\right]^2}{(\kappa+\varepsilon)\left(\delta_D^{\mathbb{Q}} - \varepsilon\right)\left[(\kappa+\varepsilon) - \left(\delta_D^{\mathbb{Q}} - \varepsilon\right)\right]^2}\left[\omega\left(\left(\kappa + \delta_D^{\mathbb{Q}}\right)\tau\right) - \omega\left((\kappa+\varepsilon)\tau\right) - \omega\left(\left(\delta_D^{\mathbb{Q}} - \varepsilon\right)\tau\right)\right],
\end{aligned}$$

where the function $\omega(.)$ is defined as $\omega(x) = 1 - \frac{1-e^{-x}}{x}$ for all $x \neq 0$ and $\omega(0) = 0$, and where

$$\varepsilon = \frac{\delta_D^{\mathbb{Q}} - \kappa - \sqrt{\left(\delta_D^{\mathbb{Q}} - \kappa\right)^2 - 4\alpha\sigma\sigma_y^{\bar{\tau}}\delta_r}}{2} \quad \text{(A-32)}$$

as long as it exists, i.e., the determinant is positive.

Next we pin down the endogenous parameters of the model. First, from (1), (11), and (14) the volatility of the reference yield has to solve

$$\sigma_y^{\bar{\tau}} = \mathcal{B}(\bar{\tau})\sigma + \mathcal{C}(\bar{\tau})\eta_y\sigma_y^{\bar{\tau}}. \quad \text{(A-33)}$$

Moreover, again from (14), we have

$$dy_t^{\bar{\tau}} = \mathcal{B}(\bar{\tau})\, dr_t + \mathcal{C}(\bar{\tau})\, dD_t. \tag{A-34}$$

Plugging (A-34) into (11) and using (1), we get

$$dD_t = \kappa_D(\theta_D - D_t)\, dt + \mathcal{B}(\bar{\tau})\,\eta_y\left[\kappa(\theta - r_t)\, dt + \sigma dB_t\right] + \mathcal{C}(\bar{\tau})\,\eta_y dD_t,$$

and collecting all dD_t terms on the LHS yields

$$\left[1 - \eta_y \mathcal{C}(\bar{\tau})\right] dD_t = \left[\kappa_D(\theta_D - D_t) + \kappa\eta_y \mathcal{B}(\bar{\tau})(\theta - r_t)\right] dt + \mathcal{B}(\bar{\tau})\,\sigma dB_t. \tag{A-35}$$

Matching the r_t, D_t, and constant terms in the drift of dD_t from (A-35) with those in (11), we obtain (15).

We make use of the following result:

Lemma 7. *As long as ε exists, we have (i) $\kappa+\varepsilon < \delta_D^{\mathbb{Q}} - \varepsilon$ always; and (ii) ε has the same sign as $\delta_D^{\mathbb{Q}} - \kappa$. Finally, (iii) $\kappa+\varepsilon$ and $\delta_D^{\mathbb{Q}} - \varepsilon$ are always "between" κ and $\delta_D^{\mathbb{Q}}$. That is, if $\kappa < \delta_D^{\mathbb{Q}}$, we have*

$$\kappa < \kappa + \varepsilon < \frac{\kappa + \delta_D^{\mathbb{Q}}}{2} < \delta_D^{\mathbb{Q}} - \varepsilon < \delta_D^{\mathbb{Q}}; \tag{A-36}$$

if $\kappa > \delta_D^{\mathbb{Q}}$, we have

$$\delta_D^{\mathbb{Q}} < \kappa + \varepsilon < \frac{\kappa + \delta_D^{\mathbb{Q}}}{2} < \delta_D^{\mathbb{Q}} - \varepsilon < \kappa. \tag{A-37}$$

Proof. First, notice that (A-32) and (15) together imply ε can be rewritten as

$$\varepsilon = \frac{\left(\delta_D^{\mathbb{Q}} - \kappa\right) - \sqrt{\left(\delta_D^{\mathbb{Q}} - \kappa\right)^2 - 4\kappa\alpha\eta_y \left(\sigma_y^{\bar{\tau}}\right)^2}}{2}. \tag{A-38}$$

From (A-38), we have

$$\kappa + \varepsilon = \frac{\delta_D^{\mathbb{Q}} + \kappa - \sqrt{\left(\delta_D^{\mathbb{Q}} - \kappa\right)^2 - 4\kappa\alpha\eta_y \left(\sigma_y^{\bar{\tau}}\right)^2}}{2}$$

and

$$\delta_D^{\mathbb{Q}} - \varepsilon = \frac{\delta_D^{\mathbb{Q}} + \kappa + \sqrt{\left(\delta_D^{\mathbb{Q}} - \kappa\right)^2 - 4\kappa\alpha\eta_y \left(\sigma_y^{\bar{\tau}}\right)^2}}{2},$$

and since the square-root is non-negative, we always have

$$\kappa + \varepsilon < \frac{\kappa + \delta_D^{\mathbb{Q}}}{2} < \delta_D^{\mathbb{Q}} - \varepsilon.$$

Second, revisiting (A-38), if $\kappa > \delta_D^{\mathbb{Q}}$, both components of the RHS are negative and thus $\varepsilon < 0$. On the other hand, since $4\kappa\alpha\eta_y \left(\sigma_y^{\bar{\tau}}\right)^2 > 0$, we have

$$\left|\delta_D^{\mathbb{Q}} - \kappa\right| > \sqrt{\left(\delta_D^{\mathbb{Q}} - \kappa\right)^2 - 4\kappa\alpha\eta_y \left(\sigma_y^{\bar{\tau}}\right)^2}.$$

Therefore, if $\delta_D^{\mathbb{Q}} > \kappa$, the first component of ε is positive and greater than the second, and thus $\varepsilon > 0$.

Third, we can write

$$\kappa + \varepsilon - \delta_D^{\mathbb{Q}} = \frac{-\left(\delta_D^{\mathbb{Q}} - \kappa\right) - \sqrt{\left(\delta_D^{\mathbb{Q}} - \kappa\right)^2 - 4\kappa\alpha\eta_y\left(\sigma_y^{\bar{\tau}}\right)^2}}{2},$$

and with similar reasoning as above, $\kappa > \delta_D^{\mathbb{Q}}$ implies $\kappa - \delta_D^{\mathbb{Q}} + \varepsilon > 0$, i.e. $\delta_D^{\mathbb{Q}} < \kappa + \varepsilon$ and $\delta_D^{\mathbb{Q}} - \varepsilon < \kappa$. Combining the three results gives inequalities (A-36) and (A-37). □

To complete the proof of the Theorem, we need to provide sufficient conditions such that the set of equations given by (15) has a solution. First, we show that all meaningful $\sigma_y^{\bar{\tau}}$ solutions of (A-33) are non-negative. Notice that with the help of (A-6), (A-30) and (A-29) can be written as

$$\mathcal{C}(\tau) = -\alpha\sigma\sigma_y^{\bar{\tau}} g_4\left(\kappa + \varepsilon, \delta_D^{\mathbb{Q}} - \varepsilon\right) \text{ and} \tag{A-39}$$

$$\mathcal{B}(\tau) = -\left[\frac{e^{-(\kappa+\varepsilon)\tau} - e^{-\left(\delta_D^{\mathbb{Q}}-\varepsilon\right)\tau}}{(\kappa+\varepsilon) - \left(\delta_D^{\mathbb{Q}} - \varepsilon\right)} + \delta_D^{\mathbb{Q}}\, g_4\left(\kappa + \varepsilon, \delta_D^{\mathbb{Q}} - \varepsilon\right)\right]. \tag{A-40}$$

Lemma 5 and (A-39) together imply that $\mathcal{C}(\tau)$ and $\sigma_y^{\bar{\tau}}$ have the same sign; therefore, the second term of the RHS of (A-33) is always non-negative. Regarding (A-40), as the function $x \mapsto e^{-x}$ is decreasing, the first term inside the bracket is negative. On the other hand, according to Lemma 5, the second term has the opposite sign as $\delta_D^{\mathbb{Q}}$. But $\delta_D^{\mathbb{Q}}$ must be positive, otherwise the duration process under $\mathbb{Q}$, (13), would explode. Hence, both terms inside the bracket are negative, i.e., $\mathcal{B}(\tau) \geq 0$. Going back to (A-33), we have shown that both components of the RHS are non-negative, and thus in all meaningful solutions $\sigma_y^{\bar{\tau}} \geq 0$. Notice that this also implies $\mathcal{C}(\tau) \geq 0$ for all $\tau \geq 0$, and from (15) it also means $0 < 1 - \eta_y\mathcal{C}(\bar{\tau}) \leq 1$.

Second, a sufficient condition for the existence of a solution to (A-33) is that its LHS is smaller than the RHS for $\sigma_y^{\bar{\tau}} = 0$ but greater than the RHS when $\sigma_y^{\bar{\tau}}$ is large enough. It is easy to see that $\sigma_y^{\bar{\tau}} = 0$ leads to $\mathcal{C}(\tau) = 0$ for all $\tau \geq 0$, and yields

$$\mathcal{B}(\tau) = \frac{1 - e^{-\kappa\tau}}{\kappa\tau}.$$

Therefore, the RHS of (A-33) is zero while the LHS equals

$$\mathcal{B}(\bar{\tau})\,\sigma + \mathcal{C}(\bar{\tau})\,\eta_y\sigma_y^{\bar{\tau}} = \frac{1 - e^{-\kappa\bar{\tau}}}{\kappa\bar{\tau}}\sigma > 0.$$

For inequality in the other direction, notice that Lemmas 5 and 7 together imply

$$g_4\left(\kappa, \delta_D^{\mathbb{Q}}\right) < g_4\left(\kappa + \varepsilon, \delta_D^{\mathbb{Q}} - \varepsilon\right) < 0, \tag{A-41}$$

regardless of the order of κ and $\delta_D^{\mathbb{Q}}$. Further, Lemma 5 states that g_4 is increasing in both arguments, therefore

$$g_4\left(\kappa, \kappa_D^{\mathbb{Q}}\right) < g_4\left(\kappa, \delta_D^{\mathbb{Q}}\right) < 0, \tag{A-42}$$

where $\kappa_D^{\mathbb{Q}} \equiv \kappa_D - \alpha\eta_y \left(\sigma_y^{\bar{\tau}}\right)^2 \leq \delta_D^{\mathbb{Q}}$ always, because $\kappa_D \leq \delta_D$ holds due to (15) and $\mathcal{C}(\tau) \geq 0$. Combining (A-39), (A-41), and (A-42), we obtain

$$0 < \mathcal{C}(\tau) < -\alpha\sigma\sigma_y^{\bar{\tau}} g_4\left(\kappa, \kappa_D^{\mathbb{Q}}\right). \tag{A-43}$$

We also approximate $\mathcal{B}(\tau)$ from above with the help of Lemma 5. First, if we assume $\kappa < \delta_D^{\mathbb{Q}}$, which also implies $\varepsilon > 0$ and $\kappa < \delta_D^{\mathbb{Q}} - \varepsilon$ according to Lemma 7, since g_4 is negative and increasing in both arguments, we have

$$g_4(\kappa + \varepsilon, \kappa) < g_4\left(\kappa + \varepsilon, \delta_D^{\mathbb{Q}} - \varepsilon\right) < 0.$$

Rearranging and using (A-30), we obtain

$$\mathcal{B}(\tau) = \frac{1 - e^{-(\kappa+\varepsilon)\tau}}{(\kappa+\varepsilon)\tau} - \varepsilon\, g_4\left(\kappa + \varepsilon, \delta_D^{\mathbb{Q}} - \varepsilon\right) < \frac{1 - e^{-\kappa\tau}}{\kappa\tau}. \tag{A-44}$$

If, on the other hand, $\delta_D^{\mathbb{Q}} < \kappa$, we rewrite (A-30) as

$$\mathcal{B}(\tau) = \frac{1 - e^{-\left(\delta_D^{\mathbb{Q}} - \varepsilon\right)\tau}}{\left(\delta_D^{\mathbb{Q}} - \varepsilon\right)\tau} + \left(\kappa + \varepsilon - \delta_D^{\mathbb{Q}}\right) g_4\left(\kappa + \varepsilon, \delta_D^{\mathbb{Q}} - \varepsilon\right).$$

Notice that Lemma 7 in this case yields $\kappa + \varepsilon - \delta_D^{\mathbb{Q}} > 0$, and since g_4 is negative, we get

$$\mathcal{B}(\tau) < \frac{1 - e^{-\left(\delta_D^{\mathbb{Q}} - \varepsilon\right)\tau}}{\left(\delta_D^{\mathbb{Q}} - \varepsilon\right)\tau} < \frac{1 - e^{-\delta_D^{\mathbb{Q}}\tau}}{\delta_D^{\mathbb{Q}}\tau} \leq \frac{1 - e^{-\kappa_D^{\mathbb{Q}}\tau}}{\kappa_D^{\mathbb{Q}}\tau}, \tag{A-45}$$

where in the last two steps we use $\delta_D^{\mathbb{Q}} - \varepsilon > \delta_D^{\mathbb{Q}} \geq \kappa_D^{\mathbb{Q}}$. Combining (A-44) and (A-45), we obtain that under any circumstances we have

$$\mathcal{B}(\tau) < \max\left\{\frac{1 - e^{-\kappa\tau}}{\kappa\tau}, \frac{1 - e^{-\kappa_D^{\mathbb{Q}}\tau}}{\kappa_D^{\mathbb{Q}}\tau}\right\}, \tag{A-46}$$

and, together with (A-43),

$$\mathcal{B}(\tau)\sigma + \mathcal{C}(\tau)\eta_y\sigma_y^{\bar{\tau}} < \left[\max\left\{\frac{1 - e^{-\kappa\tau}}{\kappa\tau}, \frac{1 - e^{-\kappa_D^{\mathbb{Q}}\tau}}{\kappa_D^{\mathbb{Q}}\tau}\right\} - \alpha\eta_y\left(\sigma_y^{\bar{\tau}}\right)^2 g_4\left(\kappa, \kappa_D^{\mathbb{Q}}\right)\right]\sigma. \tag{A-47}$$

We want to give a sufficient condition for the LHS of (A-33) to be larger than the RHS when $\sigma_y^{\bar{\tau}}$ is large enough to make $\kappa_D^{\mathbb{Q}} = 0$, that is, $\left(\sigma_y^{\bar{\tau}}\right)^2 = \frac{\kappa_D}{\alpha\eta_y}$. For this it is sufficient if we make $\sigma_y^{\bar{\tau}}$ larger than the RHS of (A-47), which, after some algebra, is equivalent to

$$\sigma_y^{\bar{\tau}} > \left[1 - \frac{\kappa_D}{\kappa}\left(\frac{1 - e^{-\kappa\tau}}{\kappa\tau} - 1\right)\right]\sigma,$$

because $\kappa_D^{\mathbb{Q}} = 0$ makes the RHS of (A-45) equal to 1. Taking squares of both sides and using $\left(\sigma_y^{\bar{\tau}}\right)^2 = \frac{\kappa_D}{\alpha \eta_y}$ again, after some algebra we obtain

$$\alpha < \frac{\kappa_D}{\eta_y \left(\frac{\kappa+\kappa_D}{\kappa} - \frac{\kappa_D}{\kappa}\frac{1-e^{-\kappa\bar{\tau}}}{\kappa\bar{\tau}}\right)^2 \sigma^2}. \tag{A-48}$$

Defining $\bar{\alpha}$ as the RHS of (A-48), which is certainly positive, (A-33) has at least one solution whenever $0 \leq \alpha < \bar{\alpha}$. □

Proof of Corollary 1. When running a linear regression in the form $Y_t = \alpha + \gamma X_t + \epsilon_t$, the theoretical slope coefficient is simply $\gamma = Cov\left[X_t, Y_t\right]/Var[X_t]$, whereas the R^2 of the regression is

$$R^2 = \frac{\gamma^2 Var[X_t]}{Var[Y_t]} = \frac{Cov^2\left[X_t, Y_t\right]}{Var[X_t]Var[Y_t]}. \tag{A-49}$$

Applying this to $Y_t = D_t$ and $X_t = r_t$ and using (A-18) yields (16). Applying (A-49) to $Y_t = D_t$ and $X_t = y_t^{\bar{\tau}}$ and combining it with (14), we obtain

$$R^2_{D,y} = \frac{Cov^2\left[y_t^{\bar{\tau}}, D_t\right]}{Var[y_t^{\bar{\tau}}]Var[D_t]} = \frac{\left(\mathcal{B}\left(\bar{\tau}\right)Cov\left[r_t, D_t\right] + \mathcal{C}\left(\bar{\tau}\right)Var\left[D_t\right]\right)^2}{\left(\mathcal{B}^2\left(\bar{\tau}\right)Var[r_t] + 2\mathcal{B}\left(\bar{\tau}\right)\mathcal{C}\left(\bar{\tau}\right)Cov[r_t, D_t] + \mathcal{C}^2\left(\bar{\tau}\right)Var[D_t]\right)Var[D_t]}.$$

Equation (A-18) and rearranging gives (17). Finally, applying (A-49) to $Y_t = D_t$ and $X_t = y_t^{\bar{\tau}} - r_t$ and combining it with (14), we obtain

$$R^2_{D,y-r} = \frac{\left(\left(\mathcal{B}\left(\bar{\tau}\right) - 1\right)Cov\left[r_t, D_t\right] + \mathcal{C}\left(\bar{\tau}\right)Var\left[D_t\right]\right)^2}{\left(\left(\mathcal{B}\left(\bar{\tau}\right) - 1\right)^2 Var[r_t] + 2\left(\mathcal{B}\left(\bar{\tau}\right) - 1\right)\mathcal{C}\left(\bar{\tau}\right)Cov[r_t, D_t] + \mathcal{C}^2\left(\bar{\tau}\right)Var[D_t]\right)Var[D_t]}.$$

Using (A-18) and rearranging gives (18). □

Proof of Proposition 1. The excess return over horizon $(t, t+h)$ on a maturity-τ bond is

$$rx_{t,t+h,\tau} = \log\Lambda_{t+h,\tau-h} - \log\Lambda_{t,\tau} + \log\Lambda_{t,h} = -\left(\tau - h\right)y_{t+h,\tau-h} + \tau y_{t,\tau} - hy_{t,h}. \tag{A-50}$$

From (14) we have

$$\begin{aligned} &E_t\left[y_{t+h,\tau-h}\right] = \mathcal{A}\left(\tau-h\right) + \mathcal{C}\left(\tau-h\right)E_t\left[D_{t+h}\right] + \mathcal{B}\left(\tau-h\right)E_t\left[r_{t+h}\right] \\ &= ... + e^{-\delta_D h}\mathcal{C}\left(\tau-h\right)D_t + \left[e^{-\kappa h}\mathcal{B}\left(\tau-h\right) - \frac{\delta_r}{\delta_D - \kappa}\left(e^{-\kappa h} - e^{-\delta_D h}\right)\mathcal{C}\left(\tau-h\right)\right]r_t, \end{aligned}$$

where in the second equality we use (A-11) and (A-12) for the time period $(t, t+h)$ and ignore the additive constant terms.

We take the expectation of the sides in (A-50) as of time t and plug in the above result. Omitting additive constants again, we obtain that the forecastable part of excess returns is

$$\begin{aligned} E_t\left[rx_{t,t+h,\tau}\right] = ... + \left[\tau\mathcal{C}\left(\tau\right) - h\mathcal{C}\left(h\right) - e^{-\delta_D h}\left(\tau-h\right)\mathcal{C}\left(\tau-h\right)\right]D_t \qquad \text{(A-51)} \\ + \left[\tau\mathcal{B}\left(\tau\right) - h\mathcal{B}\left(h\right) - e^{-\kappa h}\left(\tau-h\right)\mathcal{B}\left(\tau-h\right) - \delta_r\frac{e^{-\kappa h} - e^{-\delta_D h}}{\kappa - \delta_D}\left(\tau-h\right)\mathcal{C}\left(\tau-h\right)\right]r_t. \end{aligned}$$

When running a multivariate regression of $rx_{t,t+h,\tau}$ on D_t and r_t in the form

$$rx_{t,t+h,\tau} = \alpha_{D,r} + \left(\beta_1^{\tau,h}, \beta_2^{\tau,h}\right) \begin{pmatrix} D_t \\ r_t \end{pmatrix} + \epsilon_{t+h},$$

the vector of theoretical coefficients becomes $(\beta_1^\tau, \beta_2^\tau)^\top = V^{-1} Cov\left[rx_{t+h,\tau}, (r_t, D_t)^\top\right]$, where V is given by (A-18). Using (A-51), after some algebra we get

$$\begin{aligned} \beta_1^\tau &= \tau \mathcal{C}(\tau) - h\mathcal{C}(h) - e^{-\delta_D h}(\tau - h)\mathcal{C}(\tau - h) \text{ and} \\ \beta_2^\tau &= \tau \mathcal{B}(\tau) - h\mathcal{B}(h) - e^{-\kappa h}(\tau - h)\mathcal{B}(\tau - h) - \delta_r \frac{e^{-\kappa h} - e^{-\delta_D h}}{\kappa - \delta_D}(\tau - h)\mathcal{C}(\tau - h), \end{aligned} \tag{A-52}$$

i.e. the loadings in (A-51). On the other hand, when running a univariate regression of $rx_{t+h,\tau}$ on D_t in the form $rx_{t,t+h,\tau} = \alpha_D + \beta^\tau D_t + \epsilon_{t+h}$, the slope coefficient is

$$\beta^\tau = \frac{Cov\left[rx_{t,t+h,\tau}, D_t\right]}{Var\left[D_t\right]} = \beta_1^\tau + \frac{Cov\left[r_t, D_t\right]}{Var\left[D_t\right]}\beta_2^\tau = \beta_1^\tau + \frac{\sigma}{\eta_y \sigma_y^{\bar{\tau}}}\beta_2^\tau, \tag{A-53}$$

where the last equality is due to (15) and (A-18).

We start by looking at β_1^τ. First, it is easy to see from (A-29) that $\lim_{\tau \to h} \beta_1^\tau = 0$. Second, substituting (A-29) into (A-52), differentiating with respect to τ, and rearranging, we obtain

$$\frac{d\beta_1^\tau}{d\tau} = -\alpha\sigma\sigma_y^{\bar{\tau}} \frac{\left[e^{-(\kappa+\varepsilon)h} - e^{-\delta_D h}\right] e^{-(\kappa+\varepsilon)(\tau-h)} - \left[e^{-\left(\delta_D^{\mathbb{Q}}-\varepsilon\right)h} - e^{-\delta_D h}\right] e^{-\left(\delta_D^{\mathbb{Q}}-\varepsilon\right)(\tau-h)}}{(\kappa+\varepsilon) - \left(\delta_D^{\mathbb{Q}} - \varepsilon\right)}. \tag{A-54}$$

We focus on the parameter set that satisfies $\kappa < \delta_D$, that is, the speed of mean-reversion of the short rate is smaller than the speed of mean-reversion of the duration, which is true for the calibrated real-world parameters. We have the following result:

Lemma 8. *$\kappa < \delta_D$ implies $\delta_D^{\mathbb{Q}} - \varepsilon < \delta_D$.*

Proof. Suppose instead $\delta_D \leq \delta_D^{\mathbb{Q}} - \varepsilon$; together with $\kappa < \delta_D$ we get $\kappa < \delta_D^{\mathbb{Q}} - \varepsilon$. Revisiting Lemma 7, it must be that (A-36) holds, i.e. $\varepsilon > 0$. But then $\delta_D^{\mathbb{Q}} - \varepsilon < \delta_D^{\mathbb{Q}} < \delta_D$; contradiction. $\square$

Combining Lemmas 7 and 8, we obtain that $\kappa + \varepsilon < \delta_D^{\mathbb{Q}} - \varepsilon < \delta_D$. As $0 < h \leq \tau$, we have both $0 < e^{-\left(\delta_D^{\mathbb{Q}}-\varepsilon\right)(\tau-h)} < e^{-(\kappa+\varepsilon)(\tau-h)}$ and $0 < e^{-(\kappa+\varepsilon)h} < e^{-\left(\delta_D^{\mathbb{Q}}-\varepsilon\right)h} < e^{-\delta_D h}$. In turn, the latter implies $0 < e^{-\left(\delta_D^{\mathbb{Q}}-\varepsilon\right)h} - e^{-\delta_D h} < e^{-(\kappa+\varepsilon)h} - e^{-\delta_D h}$. Therefore, the two terms in the numerator of the LHS are positive, but both (positive) components of the first expression are larger than the corresponding component from the second. Hence, the numerator and the total RHS of (A-54) are both positive: β_1^τ is increasing in τ. Since we also have $\beta_1^h = 0$, we conclude that β_1^τ is positive and increasing across maturities.

Next we look at β_2^{τ}. First, it is easy to see from (A-30) and (A-29) that $\lim_{\tau\to h}\beta_2^{\tau}=0$. Second, substituting (A-29) and (A-30) into (A-52), differentiating with respect to τ, and rearranging, we obtain

$$\frac{d\beta_2^{\tau}}{d\tau}=\kappa\alpha\eta_y\left(\sigma_y^{\bar{\tau}}\right)^2\frac{\left[g_2\left(\delta_D\right)-g_2\left(\kappa+\varepsilon\right)\right]e^{-(\kappa+\varepsilon)(\tau-h)}-\left[g_2\left(\delta_D\right)-g_2\left(\delta_D^{\mathbb{Q}}-\varepsilon\right)\right]e^{-\left(\delta_D^{\mathbb{Q}}-\varepsilon\right)(\tau-h)}}{\left(\kappa+\varepsilon\right)-\left(\delta_D^{\mathbb{Q}}-\varepsilon\right)},\tag{A-55}$$

where $g_2\left(.\right)$ is defined in (A-3). As $\kappa+\varepsilon<\delta_D^{\mathbb{Q}}-\varepsilon<\delta_D$, Lemma 3 implies $g_2\left(\kappa+\varepsilon\right)<g_2\left(\delta_D^{\mathbb{Q}}-\varepsilon\right)<g_2\left(\delta_D\right)$, i.e. $0<g_2\left(\delta_D\right)-g_2\left(\delta_D^{\mathbb{Q}}-\varepsilon\right)<g_2\left(\delta_D\right)-g_2\left(\kappa+\varepsilon\right)$. On the other hand $\kappa+\varepsilon<\delta_D^{\mathbb{Q}}-\varepsilon$ yields $0<e^{-\left(\delta_D^{\mathbb{Q}}-\varepsilon\right)h}<e^{-(\kappa+\varepsilon)h}$. Hence, the numerator is positive, and thus the total RHS of (A-55) is negative: β_2^{τ} is decreasing in τ. Since we also have $\beta_2^{h}=0$, we conclude that β_2^{τ} is negative and decreasing (i.e., becoming more negative) across maturities.

Total univariate slope is given by (A-53). First, it is easy to confirm that $\lim_{\tau\to h}\beta^{\tau}=0$. Second, using (A-54) and (A-55) and rearranging, we obtain

$$\begin{aligned}\frac{d\beta^{\tau}}{d\tau} &= \frac{d\beta_1^{\tau}}{d\tau}+\frac{\sigma}{\eta_y\sigma_y^{\bar{\tau}}}\frac{d\beta_2^{\tau}}{d\tau} \qquad\qquad (\text{A-56})\\ &= -\alpha\sigma\sigma_y^{\bar{\tau}}\frac{\left[g_3\left(\kappa+\varepsilon\right)-g_3\left(\delta_D\right)\right]e^{-(\kappa+\varepsilon)(\tau-h)}-\left[g_3\left(\delta_D^{\mathbb{Q}}-\varepsilon\right)-g_3\left(\delta_D\right)\right]e^{-\left(\delta_D^{\mathbb{Q}}-\varepsilon\right)(\tau-h)}}{\left(\kappa+\varepsilon\right)-\left(\delta_D^{\mathbb{Q}}-\varepsilon\right)},\end{aligned}$$

where $g_3\left(.\right)$ is defined in (A-4). Lemma 4 implies that for small $h>0$, the function g_3 is decreasing, hence $g_3\left(\delta_D\right)<g_3\left(\delta_D^{\mathbb{Q}}-\varepsilon\right)<g_3\left(\kappa+\varepsilon\right)$, i.e. $0<g_3\left(\delta_D^{\mathbb{Q}}-\varepsilon\right)-g_3\left(\delta_D\right)<g_3\left(\kappa+\varepsilon\right)-g_3\left(\delta_D\right)$. On the other hand $\kappa+\varepsilon<\delta_D^{\mathbb{Q}}-\varepsilon$ implies $0<e^{-\left(\delta_D^{\mathbb{Q}}-\varepsilon\right)h}<e^{-(\kappa+\varepsilon)h}$. Hence, the numerator and the total RHS of (A-56) are positive: β^{τ} is increasing in τ. Since we also have $\beta^{h}=0$, we conclude that β^{τ} is positive and increasing across maturities.

Finally, from (14), the effect of duration on yields is given by $\mathcal{C}\left(\tau\right)$. From the Proof of Theorem 1, $\mathcal{C}\left(\tau\right)\geq 0$. Moreover, from (A-29) it is easy to show that

$$\lim_{\tau\to 0}\mathcal{C}\left(\tau\right)=\lim_{\tau\to\infty}\mathcal{C}\left(\tau\right)=0,$$

with $\mathcal{C}\left(\tau\right)$ increasing for small but decreasing for large τ values, which implies that the effect is either increasing across maturities if T is small, or first increasing then decreasing if T is sufficiently large. This completes the proof. □

Proof of Propositions 2 and 3. From (14), (A-29), and (A-30), bond yield volatility is given by

$$\begin{aligned}\sigma_y^{\tau} &= \mathcal{B}\left(\tau\right)\sigma+\mathcal{C}\left(\tau\right)\eta_y\sigma_y^{\bar{\tau}} \qquad\qquad (\text{A-57})\\ &= \frac{1-e^{-(\kappa+\varepsilon)\tau}}{\left(\kappa+\varepsilon\right)\tau}\sigma-\frac{\varepsilon+\alpha\eta_y\left(\sigma_y^{\bar{\tau}}\right)^2}{\left(\kappa+\varepsilon\right)-\left(\delta_D^{Q}-\varepsilon\right)}\left[\frac{1-e^{-(\kappa+\varepsilon)\tau}}{\left(\kappa+\varepsilon\right)\tau}-\frac{1-e^{-\left(\delta_D^{Q}-\varepsilon\right)\tau}}{\left(\delta_D^{Q}-\varepsilon\right)\tau}\right]\sigma.\end{aligned}$$

Due to the complexity the feedback mechanism introduces into the endogenous parameters, we cannot compute the exact effect of convexity $-\eta_y$ on yield volatilities in closed form. Instead, we derive its effect by considering (A-57) around $\alpha = 0$. From (A-57) we write $\sigma_y^\tau \approx h_0(\tau) + \alpha\eta_y h_1(\tau)$, where

$$h_0(\tau) \equiv \left(\mathcal{B}(\tau)\sigma + \mathcal{C}(\tau)\eta_y\sigma_y^{\bar{\tau}}\right)|_{\alpha=0} \text{ and } h_1(\tau) \equiv \frac{1}{\eta_y}\frac{d\left(\mathcal{B}(\tau)\sigma + \mathcal{C}(\tau)\eta_y\sigma_y^{\bar{\tau}}\right)}{d\alpha}|_{\alpha=0}.$$

We start with h_0. It is straightforward from (A-32) and (15) that taking the limit $\alpha \to 0$ yields

$$\lim_{\alpha\to 0}\varepsilon = 0 \text{ and } \lim_{\alpha\to 0}\delta_D^Q = \lim_{\alpha\to 0}\delta_D = \kappa_D, \tag{A-58}$$

hence

$$h_0(\tau) = \frac{1-e^{-\kappa\tau}}{\kappa\tau}\sigma. \tag{A-59}$$

and as a special case, the volatility of the reference-maturity yield is

$$\lim_{\alpha\to 0}\sigma_y^{\bar{\tau}} = h_0(\bar{\tau}) = \frac{1-e^{-\kappa\bar{\tau}}}{\kappa\bar{\tau}}\sigma. \tag{A-60}$$

Second, differentiating (A-57) with respect to α, we get

$$\begin{aligned}\frac{1}{\sigma}\frac{d\sigma_y^\tau}{d\alpha} &= -\frac{d(\kappa+\varepsilon)}{d\alpha}\frac{1-e^{-(\kappa+\varepsilon)\tau} - (\kappa+\varepsilon)\tau e^{-(\kappa+\varepsilon)\tau}}{(\kappa+\varepsilon)^2\tau} \\ &\quad -\left(\varepsilon + \alpha\eta_y\left(\sigma_y^{\bar{\tau}}\right)^2\right)\frac{d}{d\alpha}\left(\frac{\frac{1-e^{-(\kappa+\varepsilon)\tau}}{(\kappa+\varepsilon)\tau} - \frac{1-e^{-\left(\delta_D^Q-\varepsilon\right)\tau}}{\left(\delta_D^Q-\varepsilon\right)\tau}}{(\kappa+\varepsilon) - \left(\delta_D^Q-\varepsilon\right)}\right) \\ &\quad -\left(\frac{d\varepsilon}{d\alpha} + \eta_y\left(\sigma_y^{\bar{\tau}}\right)^2 + 2\alpha\eta_y\sigma_y^{\bar{\tau}}\frac{d\sigma_y^{\bar{\tau}}}{d\alpha}\right)\frac{\frac{1-e^{-(\kappa+\varepsilon)\tau}}{(\kappa+\varepsilon)\tau} - \frac{1-e^{-\left(\delta_D^Q-\varepsilon\right)\tau}}{\left(\delta_D^Q-\varepsilon\right)\tau}}{(\kappa+\varepsilon) - \left(\delta_D^Q-\varepsilon\right)}.\end{aligned} \tag{A-61}$$

As

$$\frac{d(\kappa+\varepsilon)}{d\alpha} = \frac{d\varepsilon}{d\alpha} = \frac{\varepsilon}{(\kappa+\varepsilon) - \left(\delta_D^Q-\varepsilon\right)}\frac{d\delta_D^Q}{d\alpha} - \frac{\kappa\eta_y\left(\sigma_y^{\bar{\tau}}\right)^2}{(\kappa+\varepsilon) - \left(\delta_D^Q-\varepsilon\right)},$$

we get

$$\lim_{\alpha\to 0}\frac{d(\kappa+\varepsilon)}{d\alpha} = \lim_{\alpha\to 0}\frac{d\varepsilon}{d\alpha} = \frac{\kappa\eta_y}{\kappa_D-\kappa}\lim_{\alpha\to 0}\left(\sigma_y^{\bar{\tau}}\right)^2 = \frac{\kappa\eta_y\sigma^2}{\kappa_D-\kappa}\left(\frac{1-e^{-\kappa\bar{\tau}}}{\kappa\bar{\tau}}\right)^2, \tag{A-62}$$

where in the last step we used (A-60). Hence, after some algebra, (A-61) yields

$$\begin{aligned}\frac{1}{\sigma}\lim_{\alpha\to 0}\frac{d\sigma_y^\tau}{d\alpha} &= \frac{\kappa_D\eta_y\sigma^2}{(\kappa_D-\kappa)^2}\left(\frac{1-e^{-\kappa\bar{\tau}}}{\kappa\bar{\tau}}\right)^2\left[\frac{1-e^{-\kappa\tau}}{\kappa\tau} - \frac{\kappa_D-\kappa}{\kappa_D}\frac{1-e^{-\kappa\tau}-\kappa\tau e^{-\kappa\tau}}{\kappa\tau} - \frac{1-e^{-\kappa_D\tau}}{\kappa_D\tau}\right] \\ &= \frac{\kappa_D\eta_y\sigma^2}{(\kappa_D-\kappa)^2}\left(\frac{1-e^{-\kappa\bar{\tau}}}{\kappa\bar{\tau}}\sigma\right)^2 g_5(\kappa,\kappa_D),\end{aligned}$$

and thus

$$h_1(\tau) = \frac{\kappa_D\sigma^3}{(\kappa_D-\kappa)^2}\left(\frac{1-e^{-\kappa\bar{\tau}}}{\kappa\bar{\tau}}\right)^2 g_5(\kappa,\kappa_D),$$

where g_5 is defined in (A-7). But $g_5 \geq 0$ always according to Lemma 6, hence $h_1(\tau) \geq 0$, which implies that bond yield volatilities are increasing in negative convexity: $\frac{d\sigma_y^\tau}{d\eta_y} > 0$.

Third, we trivially verify that

$$\lim_{\tau \to 0} \sigma_y^\tau = \sigma \text{ and } \lim_{\tau \to \infty} \sigma_y^\tau = 0,$$

independent of η_y. Therefore, the effect of negative convexity on yield volatilities tends to zero at very short and very long maturities, and hence it must be hump-shaped.

Regarding bond return volatilities, given by $\sigma_y^\tau \tau$, we have $\lim_{\tau \to 0} \sigma_y^\tau \tau = 0$ and

$$\lim_{\tau \to \infty} \sigma_y^\tau \tau = \frac{\delta_D}{(\kappa + \varepsilon)\left(\delta_D^Q - \varepsilon\right)} \sigma = \frac{\sigma}{\kappa},$$

again independent of η_y, where the last equality is due to (A-32). Hence, the effect of negative convexity on bond return volatilities tends to zero at very short and very long maturities. However, as η_y increases σ_y^τ, it also increases $\sigma_y^\tau \tau$, and thus the effect must be hump-shaped.

□

Appendix C Time-varying convexity

Here we present a tractable way to relax the assumption of constant MBS convexity and capture the non-linearities inherent to the prepayment option. This version of the model allows for an additional degree of freedom and provides a better statistical description of MBS duration and convexity series. However, the qualitative implications of the model are identical to the ones outlined in Section 1. More precisely, we allow the sensitivity of outstanding MBS to the short rate to be quadratic:

$$\frac{dMBS_t}{dr_t} = z_t + \phi z_t^2, \quad \text{and} \tag{A-63}$$

$$dz_t = -\kappa_z^{\mathbb{Q}} z_t dt + \sigma_z dB_t^{\mathbb{Q}}. \tag{A-64}$$

In the data, when interest rates and MBS duration decrease, the negative convexity of MBS increases. In other words MBS duration has negative skewness. The skewness of the monthly series of MBS duration in our sample is equal to -1.32 compared to the 10 year yield which displays only a moderate skewness of -0.06. The parameter ϕ can be calibrated to match this feature of the data. From an economic point of view negative skewness corresponds to the asymmetry in MBS duration response to changes in interest rates: it reacts more to falling than to rising interest rates.

In the model described by (1), (9) and (A-63)-(A-64) yields are given by

$$y_t^\tau = \mathcal{A}_z(\tau) + \mathcal{B}_z(\tau) r_t + \mathcal{C}_z(\tau) z_t + \mathcal{D}_z(\tau) z_t^2,$$

The key to quadratic closed form solution is that while quadratic terms appear under $\mathbb{Q}$ in the dynamics of r_t, z_t is still affine under $\mathbb{Q}$ (and therefore not affine under $\mathbb{P}$); see also Cheng and Scaillet (2007). Bond prices are given by

$$\Lambda_t^\tau = e^{-\left[A_z(\tau) + B_z(\tau) r_t + C_z(\tau) z_t + D_z(\tau) z_t^2\right]},$$

where $A_z(\tau) \equiv \mathcal{A}_z(\tau)\tau$, $B_z(\tau) \equiv \mathcal{B}_z(\tau)\tau$, $C_z(\tau) \equiv \mathcal{C}_z(\tau)\tau$, and $D_z(\tau) \equiv \mathcal{D}_z(\tau)\tau$. No-arbitrage pricing of bonds results in the following system of ODEs (where we remove the time-dependence and subscript to simplify the notation):

$$0 = A' - \kappa\theta B + \frac{1}{2}\sigma^2 B^2 + \frac{1}{2}\sigma_z^2\left(C^2 - 2D\right) + \sigma\sigma_z BC,$$

and

$$1 = B' + \kappa B,$$

and

$$0 = C' + \left(\kappa_z^{\mathbb{Q}} + 2\sigma_z^2 D\right)C - \alpha\sigma^2 B + 2\sigma\sigma_z BD,$$

and

$$0 = D' + 2\kappa_z^{\mathbb{Q}} D + 2\sigma_z^2 D^2 - \alpha\sigma^2\phi B,$$

together with the boundary conditions $A(0) = B(0) = C(0) = D(0) = 0$. The solution to the system above can be written in terms of J- and Y-type Bessel functions. To simplify, we can also solve for $A_z(\tau)$, $B_z(\tau)$, $C_z(\tau)$, and $D_z(\tau)$ recursively using a discrete time approximation of the dynamics of the state variables.

By Itô's lemma the second order dollar sensitivity of outstanding MBS to short rate shocks ($\frac{d^2 MBS_t}{dr_t^2} \equiv -\gamma$) is equal to:

$$\sigma_z + 2\phi\sigma_z z_t,$$

implying time-varying convexity. The instantaneous volatility of maturity-τ yield is given by

$$\mathcal{B}_z(\tau)\sigma + \mathcal{C}_z(\tau)\sigma_z + 2\mathcal{D}_z(\tau)\sigma z_t.$$

The code that calculates $A_z(\tau)$, $B_z(\tau)$, $C_z(\tau)$, and $D_z(\tau)$ and allows to verify Propositions 1, 2, and 3 in the context of stochastic convexity is available upon request.

Appendix D Tables

Table 1
Bond risk premia regressions: Treasuries

This table reports estimated coefficients from regressing annual bond excess returns constructed from Treasuries, rx_{t+1}^{τ}, onto a set of variables:

$$rx_{t+1}^{\tau} = \beta_1^{\tau}\text{duration}_t + \beta_2^{\tau}\text{level}_t + \epsilon_{t+1}^{\tau},$$

where level_t is the first principal component from bond yields. t-Statistics presented in parentheses are calculated using Newey and West (1987). All variables are standardized to have mean zero and a standard deviation of one. Data is weekly and runs from 1997 through 2011.

	2y	3y	4y	5y	6y	7y	8y	9y	10y
duration	0.116	0.167	0.212	0.251	0.285	0.314	0.340	0.362	0.381
	(1.63)	(2.28)	(2.88)	(3.38)	(3.80)	(4.16)	(4.46)	(4.72)	(4.96)
Adj. R^2	1.34%	2.78%	4.48%	6.29%	8.11%	9.87%	11.53%	13.07%	14.51%
duration	0.067	0.161	0.243	0.312	0.370	0.419	0.459	0.492	0.520
	(0.83)	(1.94)	(2.91)	(3.78)	(4.54)	(5.18)	(5.72)	(6.16)	(6.54)
level	0.133	0.014	-0.083	-0.164	-0.230	-0.282	-0.322	-0.353	-0.375
	(1.48)	(0.16)	(-0.96)	(-1.89)	(-2.63)	(-3.20)	(-3.60)	(-3.89)	(-4.07)
Adj. R^2	2.75%	2.67%	4.96%	8.50%	12.56%	16.63%	20.40%	23.71%	26.53%

Table 2
Bond volatility regressions

Panel A and B report estimated coefficients from regressing bond yield volatility, vol_t^τ, and τ-year maturity implied volatility of swaptions written on the 10-year swap rate, $\text{iv}_t^{\tau y10y}$, onto convexity:

$$\text{vol}_t^\tau/\text{iv}_t^{\tau y10y} = \beta_1^\tau \text{convexity}_t + \epsilon_t^\tau,$$

where $\tau = 1, \ldots, 10$. Panel C reports estimated coefficients from regressing treasury implied volatility (tiv), realized volatility (trv), the bond variance risk premium (vrp), returns on a monthly straddle strategy (straddle), implied and realized measures of volatility of volatility (tivov/trvov) onto convexity. t-Statistics presented in parentheses are calculated using Newey and West (1987). All variables are standardized to have mean zero and a standard deviation of one. Data is weekly and runs from 1997 through 2011.

Panel A: Bond Yield Volatility

	1y	2y	3y	4y	5y	6y	7y	8y	9y	10y
convexity	0.248	0.264	0.258	0.247	0.231	0.213	0.195	0.180	0.166	0.154
	(3.24)	(3.38)	(3.32)	(3.19)	(3.01)	(2.81)	(2.61)	(2.44)	(2.28)	(2.15)
Adj. R^2	6.17%	6.96%	6.66%	6.10%	5.34%	4.54%	3.82%	3.22%	2.75%	2.38%

Panel B: Swaption Implied Volatility

	1y	2y	3y	4y	5y	6y	7y	8y	9y	10y
convexity	0.283	0.289	0.308	0.314	0.313	0.240	0.323	0.251	0.256	0.319
	(3.25)	(3.41)	(3.53)	(3.66)	(3.77)	(2.09)	(3.82)	(2.19)	(2.23)	(3.69)
Adj. R^2	8.02%	8.36%	9.51%	9.88%	9.77%	5.77%	10.45%	6.32%	6.55%	10.18%

Panel C: Bond Return Volatility

	iv	rv	vrp	straddle	tivov	trvov
convexity	0.186	0.146	0.187	0.090	0.152	0.159
	(2.47)	(1.75)	(3.29)	(1.93)	(2.69)	(2.59)
Adj. R^2	3.45%	2.12%	3.48%	1.10%	2.32%	2.52%

Table 3
Duration and yield factors

Panel A reports the unconditional correlation between MBS duration and the first three principal components (PCs) of bond yields. Panel B reports estimated coefficients from regressing bond excess returns onto duration and the PCs.

$$rx^{\tau}_{t+1} = \beta^{\tau}_1 \text{duration}_t + \beta^{\tau}_2 \text{level}_t + \beta^{\tau}_3 \text{slope}_t + \beta^{\tau}_4 \text{curvature}_t + \epsilon^{\tau}_{t+1},$$

where $\text{level}_t, \text{slope}_t$ and curvature_t represent the first three principal components of bond yields. t-Statistics presented in parentheses are calculated using Newey and West (1987). All variables are standardized to have mean zero and a standard deviation of one. Data is weekly and runs from 1997 through 2011.

Panel A: Unconditional Correlations

PC1	PC2	PC3
0.55	0.12	-0.14

Panel B: Bond Excess Return Regression

	2y	3y	4y	5y	6y	7y	8y	9y	10y
duration	0.064	0.144	0.216	0.278	0.331	0.374	0.409	0.438	0.462
	(0.87)	(1.88)	(2.84)	(3.70)	(4.45)	(5.05)	(5.54)	(5.93)	(6.24)
level	0.420	0.318	0.233	0.164	0.108	0.063	0.028	0.001	-0.020
	(3.78)	(3.01)	(2.26)	(1.59)	(1.03)	(0.58)	(0.25)	(0.00)	(-0.17)
slope	0.334	0.374	0.403	0.428	0.449	0.466	0.480	0.491	0.497
	(2.68)	(2.94)	(3.20)	(3.50)	(3.81)	(4.13)	(4.43)	(4.70)	(4.93)
curvature	0.318	0.284	0.270	0.262	0.254	0.246	0.237	0.228	0.218
	(2.75)	(2.46)	(2.37)	(2.38)	(2.42)	(2.47)	(2.51)	(2.52)	(2.51)
Adj. R^2	14.54%	14.36%	17.19%	21.49%	26.33%	31.07%	35.36%	39.04%	42.08%

Table 4
Bond risk premia regressions with controls

This table reports estimated coefficients from regressing bond excess returns, rx^{τ}_{t+1}, onto a set of variables:

$$rx^{\tau}_{t+1} = \beta^{\tau}_1 \text{duration}_t + \beta^{\tau}_2 \text{slope}_t + \beta^{\tau}_3 \text{cp}_t + \sum_{i=1}^{8} \text{F}^i_t + \epsilon_t,$$

where slope_t is the slope at time t, cp_t is the CP factor at time t, and F^i_t are the Ludvigson and Ng (2009) macro factors. t-Statistics presented in parentheses are calculated using Newey and West (1987). All variables are standardized to have mean zero and a standard deviation of one. Data is monthly and runs from 1997 through 2011.

	2y	3y	4y	5y	6y	7y	8y	9y	10y
duration	0.127	0.214	0.279	0.331	0.376	0.414	0.446	0.474	0.498
	(0.96)	(1.60)	(2.12)	(2.60)	(3.03)	(3.43)	(3.80)	(4.12)	(4.41)
slope	-0.079	0.042	0.130	0.198	0.253	0.298	0.333	0.361	0.381
	(-0.51)	(0.27)	(0.86)	(1.37)	(1.82)	(2.22)	(2.59)	(2.90)	(3.18)
cp factor	0.370	0.280	0.223	0.183	0.155	0.133	0.117	0.105	0.095
	(2.58)	(1.99)	(1.63)	(1.40)	(1.23)	(1.10)	(0.99)	(0.91)	(0.84)
Adj. R^2	16.34%	15.59%	17.23%	19.87%	22.93%	26.07%	29.06%	31.75%	34.05%
duration	0.081	0.142	0.189	0.229	0.264	0.297	0.327	0.356	0.382
	(0.63)	(1.04)	(1.38)	(1.67)	(1.96)	(2.23)	(2.51)	(2.78)	(3.04)
slope	-0.612	-0.507	-0.410	-0.317	-0.229	-0.145	-0.067	0.004	0.066
	(-3.48)	(-2.91)	(-2.42)	(-1.94)	(-1.45)	(-0.95)	(-0.45)	(0.03)	(0.48)
cp factor	0.584	0.548	0.519	0.496	0.474	0.451	0.428	0.404	0.381
	(5.86)	(5.88)	(5.73)	(5.56)	(5.37)	(5.15)	(4.90)	(4.61)	(4.28)
F^1	-0.588	-0.535	-0.476	-0.418	-0.361	-0.307	-0.257	-0.211	-0.168
	(-5.21)	(-4.88)	(-4.52)	(-4.15)	(-3.76)	(-3.34)	(-2.91)	(-2.45)	(-1.98)
F^2	-0.111	-0.111	-0.110	-0.108	-0.105	-0.101	-0.096	-0.089	-0.082
	(-2.75)	(-2.60)	(-2.57)	(-2.57)	(-2.56)	(-2.52)	(-2.44)	(-2.31)	(-2.13)
F^3	0.322	0.386	0.427	0.449	0.456	0.452	0.442	0.427	0.410
	(3.58)	(4.13)	(4.56)	(4.80)	(4.88)	(4.82)	(4.66)	(4.42)	(4.14)
F^4	0.016	0.015	0.024	0.033	0.041	0.047	0.050	0.053	0.055
	(0.17)	(0.15)	(0.23)	(0.32)	(0.40)	(0.47)	(0.53)	(0.57)	(0.62)
F^5	0.025	0.018	0.012	0.007	0.004	0.002	0.001	-0.001	-0.002
	(0.49)	(0.37)	(0.25)	(0.16)	(0.09)	(0.05)	(0.01)	(-0.01)	(-0.04)
F^6	-0.067	-0.091	-0.095	-0.090	-0.080	-0.068	-0.056	-0.045	-0.034
	(-0.83)	(-1.12)	(-1.20)	(-1.16)	(-1.07)	(-0.93)	(-0.78)	(-0.63)	(-0.49)
F^7	-0.116	-0.121	-0.123	-0.125	-0.125	-0.124	-0.121	-0.117	-0.111
	(-1.44)	(-1.48)	(-1.52)	(-1.56)	(-1.58)	(-1.58)	(-1.56)	(-1.53)	(-1.48)
F^8	-0.014	-0.017	-0.020	-0.021	-0.020	-0.018	-0.015	-0.011	-0.007
	(-0.22)	(-0.25)	(-0.28)	(-0.29)	(-0.29)	(-0.26)	(-0.22)	(-0.17)	(-0.11)
Adj. R^2	39.04%	37.12%	37.41%	38.76%	40.58%	42.52%	44.37%	46.01%	47.36%

Table 5
Bond volatility regressions with controls

Panel A and B report estimated coefficients from regressing bond yield volatility, vol_t^τ, and τ-year maturity implied volatility of swaptions written on the 10-year swap rate, $\text{iv}_t^{\tau y10y}$, onto convexity and illiquidity:

$$\text{vol}_t^\tau/\text{iv}_t^{\tau y10y} = \beta_1^\tau \text{convexity}_t + \beta_2^\tau \text{illiq}_t + \epsilon_t^\tau,$$

where $\tau = 1, \ldots, 10$ and illiq_t is the illiquidity factor at time t. Panel C reports estimated coefficients from regressing treasury implied volatility (tiv), realized volatility (trv), the bond variance risk premium (vrp), returns on a monthly straddle strategy (straddle), implied and realized measures of volatility of volatility (tivov/trvov) onto convexity and illiquidity. t-Statistics presented in parentheses are calculated using Newey and West (1987). All variables are standardized to have mean zero and a standard deviation of one. Data is weekly and runs from 1997 through 2011.

Panel A: Bond Yield Volatility

	1y	2y	3y	4y	5y	6y	7y	8y	9y	10y
convexity	0.301	0.312	0.307	0.299	0.286	0.271	0.257	0.243	0.231	0.220
	(4.51)	(4.50)	(4.49)	(4.46)	(4.38)	(4.27)	(4.15)	(4.01)	(3.87)	(3.73)
illiq	0.478	0.438	0.450	0.474	0.505	0.535	0.560	0.579	0.592	0.601
	(6.52)	(5.37)	(5.20)	(5.17)	(5.19)	(5.18)	(5.10)	(4.95)	(4.78)	(4.61)
Adj. R^2	28.64%	25.82%	26.54%	28.21%	30.40%	32.68%	34.69%	36.25%	37.34%	38.02%

Panel B: Swaption Implied Volatility

	1y	2y	3y	4y	5y	6y	7y	8y	9y	10y
convexity	0.272	0.267	0.258	0.266	0.266	0.259	0.248	0.236	0.227	0.221
	(2.81)	(2.61)	(2.63)	(2.94)	(3.23)	(3.47)	(3.63)	(3.70)	(3.71)	(3.71)
illiq	0.234	0.201	0.216	0.260	0.299	0.327	0.343	0.351	0.355	0.356
	(4.14)	(3.12)	(3.34)	(4.07)	(4.26)	(3.73)	(3.19)	(2.85)	(2.65)	(2.54)
Adj. R^2	11.31%	9.87%	9.95%	12.15%	14.14%	15.38%	15.85%	15.91%	15.82%	15.69%

Panel C: Bond Return Volatility

	iv	rv	vrp	straddle	tivov	trvov
convexity	0.267	0.232	0.262	0.096	0.213	0.212
	(3.65)	(2.87)	(4.41)	(1.95)	(4.11)	(3.86)
illiq	0.684	0.700	0.628	0.106	0.545	0.473
	(14.06)	(15.57)	(6.11)	(2.52)	(4.86)	(6.14)
Adj. R^2	49.28%	50.18%	42.06%	1.30%	31.55%	24.55%

Table 6
Regressions with interaction terms

This table reports estimated coefficients from regressing annual bond excess returns (Panel A) and bond yield volatility (Panel B) onto MBS duration and convexity interacted with the ratio between total mortgages outstanding and GDP. t-Statistics presented in parentheses are calculated using Newey and West (1987). All variables are standardized to have mean zero and a standard deviation of one. Data is quarterly and runs from 1997 through 2011.

`Panel A: Bond Excess Return Regression`

	2y	3y	4y	5y	6y	7y	8y	9y	10y
duration × ratio	0.131	0.216	0.270	0.305	0.326	0.339	0.346	0.350	0.351
	(0.83)	(1.39)	(1.77)	(2.01)	(2.15)	(2.22)	(2.25)	(2.25)	(2.24)
Adj. R^2	1.70%	4.68%	7.31%	9.29%	10.65%	11.50%	11.99%	12.23%	12.31%
duration × ratio	0.112	0.147	0.165	0.169	0.166	0.160	0.153	0.145	0.139
	(1.60)	(2.07)	(2.27)	(2.30)	(2.23)	(2.11)	(1.99)	(1.87)	(1.76)
level	0.225	0.138	0.062	-0.006	-0.066	-0.115	-0.156	-0.187	-0.211
	(3.18)	(2.00)	(0.90)	(-0.08)	(-0.92)	(-1.59)	(-2.11)	(-2.50)	(-2.78)
Adj. R^2	3.41%	1.62%	1.57%	2.38%	3.62%	5.02%	6.38%	7.56%	8.52%

`Panel B: Bond Yield Volatility Regression`

	1y	2y	3y	4y	5y	6y	7y	8y	9y	10y
convexity × ratio	0.404	0.422	0.439	0.453	0.461	0.467	0.472	0.477	0.484	0.492
	(2.48)	(2.77)	(2.89)	(2.99)	(3.08)	(3.16)	(3.23)	(3.30)	(3.37)	(3.43)
Adj. R^2	16.35%	17.82%	19.30%	20.49%	21.23%	21.76%	22.25%	22.80%	23.44%	24.17%
convexity × ratio	0.347	0.364	0.382	0.396	0.406	0.413	0.419	0.425	0.432	0.440
	(3.25)	(3.83)	(4.10)	(4.30)	(4.47)	(4.62)	(4.76)	(4.87)	(4.97)	(5.06)
illiq	0.635	0.647	0.639	0.623	0.608	0.596	0.587	0.579	0.572	0.566
	(10.02)	(10.44)	(10.08)	(8.91)	(7.82)	(7.14)	(6.74)	(6.52)	(6.40)	(6.36)
Adj. R^2	55.58%	58.52%	59.04%	58.25%	57.17%	56.25%	55.62%	55.26%	55.12%	55.14%

Table 7
Bond risk premia regressions: swaps

This table reports estimated coefficients from regressing annual bond excess returns constructed from interest rate swaps, rx^{τ}_{t+1}, onto a set of variables:

$$rx^{\tau}_{t+1} = \beta^{\tau}_{1}\text{duration}_t + \beta^{\tau}_{2}\text{slope}_t + \beta^{\tau}_{3}\text{cp}_t + \epsilon^{\tau}_{t+1},$$

where slope_t is the slope of the term structure at time t and cp_t is the CP factor at time t. t-Statistics presented in parentheses are calculated using Newey and West (1987). All variables are standardized to have mean zero and a standard deviation of one. Data is weekly and runs from 1997 through 2011.

	2y	3y	4y	5y	6y	7y	8y	9y	10y
duration	0.100	0.146	0.202	0.252	0.296	0.335	0.367	0.395	0.421
	(1.17)	(1.76)	(2.52)	(3.26)	(3.95)	(4.57)	(5.09)	(5.56)	(5.99)
Adj. R^2	0.99%	2.13%	4.08%	6.33%	8.78%	11.23%	13.48%	15.60%	17.68%
duration	0.050	0.157	0.251	0.327	0.388	0.436	0.474	0.505	0.531
	(0.54)	(1.73)	(2.92)	(4.04)	(5.02)	(5.84)	(6.50)	(7.07)	(7.53)
level	0.135	-0.031	-0.133	-0.204	-0.248	-0.273	-0.288	-0.297	-0.299
	(1.37)	(-0.33)	(-1.48)	(-2.33)	(-2.88)	(-3.20)	(-3.39)	(-3.50)	(-3.53)
Adj. R^2	2.43%	2.08%	5.49%	9.82%	13.99%	17.57%	20.55%	23.10%	25.32%

Table 8
Bond portfolio regressions

This table reports estimated coefficients from regressing bond portfolio excess returns onto duration and level.

$$\text{rxpf}^{\tau}_{t+1} = \beta^{\tau}_1 \text{duration}_t + \beta^{\tau}_2 \text{level}_t + \epsilon^{\tau}_{t+j},$$

where level_t is the first principal component from bond yields and rxpf_{t+1} are monthly excess returns on the CRSP bond portfolios with maturities between 5 and 10 years and larger than 10 years. Returns are in excess of either the 1-month T-bill or Eurodollar deposit rate. All variables are standardized to have mean zero and a standard deviation of one. Data is monthly and runs from 1990 through 2011.

	T-bill		ED rate	
	$\geq$ 5y $<$ 10y	$>$ 10y	$\geq$ 5y $<$ 10y	$\geq$ 5y $<$ 10y
duration	0.159	0.163	0.156	0.161
	(2.88)	(3.17)	(2.79)	(3.11)
Adj. R^2	2.53%	2.67%	2.42%	2.60%
duration	0.199	0.199	0.196	0.196
	(2.43)	(2.72)	(2.42)	(2.70)
level	-0.054	-0.038	-0.056	-0.039
	(-0.61)	(-0.48)	(-0.64)	(-0.49)
Adj. R^2	2.55%	2.77%	2.40%	2.68%

Appendix E Figures

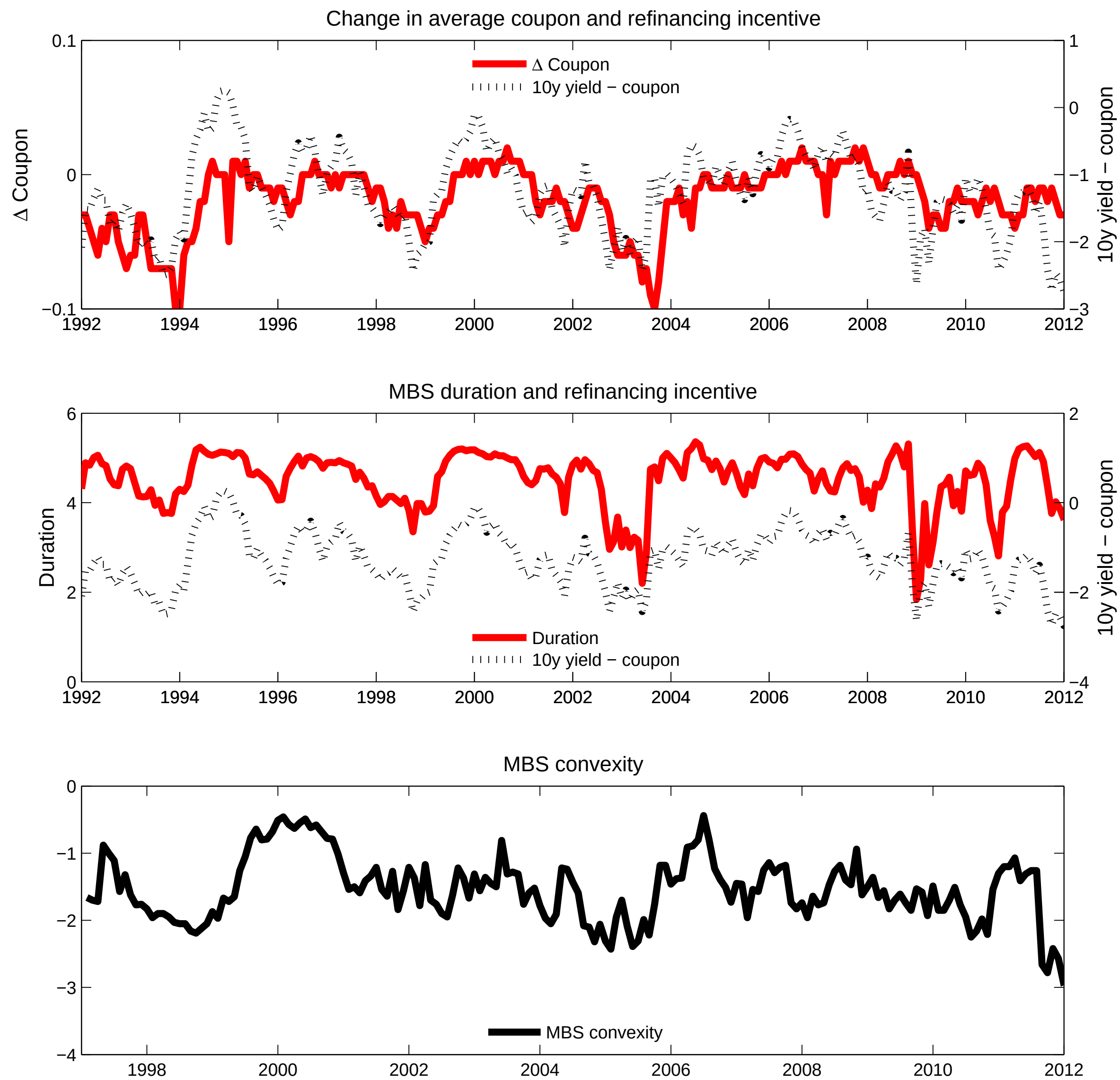

Figure 1. Average coupon, duration, and convexity

The upper panel plots the difference between the 10-year yield and the average MBS coupon together with the subsequent change in the average MBS coupon. The middle panel plots the difference between the 10-year yield and the average MBS coupon together with MBS duration. The lower panel plots MBS convexity. Data is monthly and runs from January 1990 to December 2012 (upper and middle panels) and from January 1997 to December 2012 (lower panel), respectively.

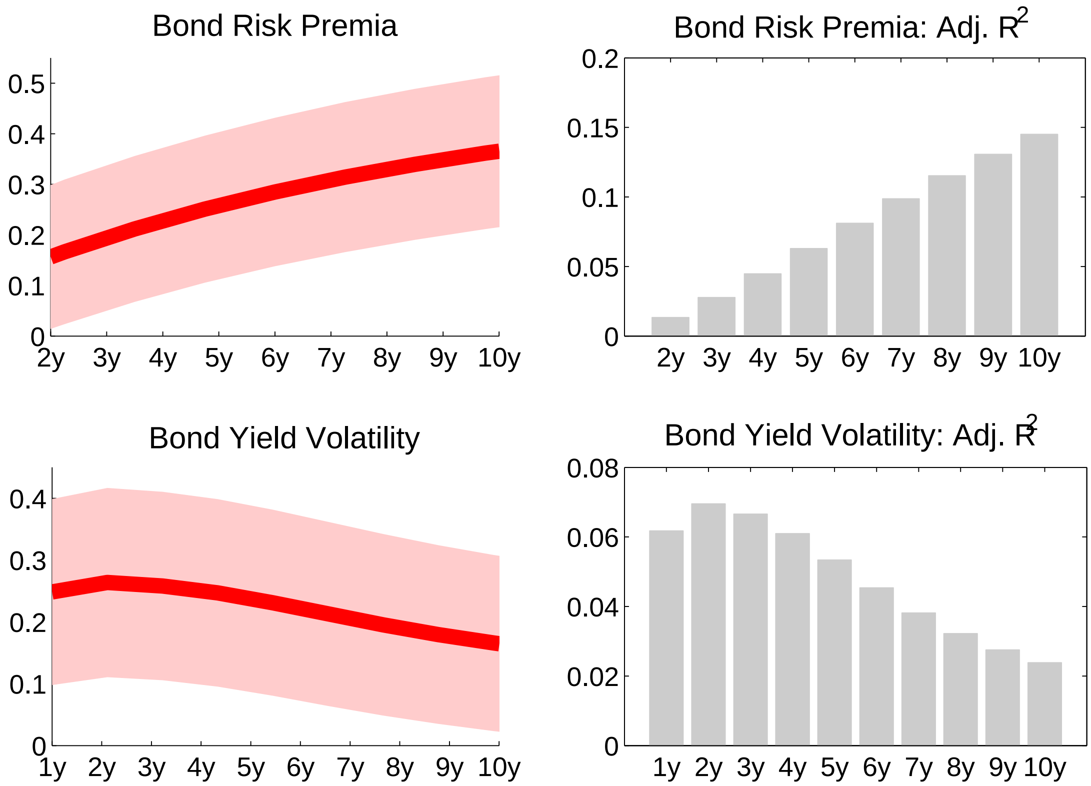

Figure 2. Univariate regression coefficients

The figure plots estimated coefficients and adjusted R^2 from univariate regressions of bond excess returns (upper panels) and bond yield volatilities (lower panels) onto MBS duration (bond excess returns) and MBS convexity (bond yield volatilities), respectively. All variables are standardized, i.e., they have a mean of zero and a standard deviation of one. Data is weekly and runs from 1997 through 2011. Shaded areas represent confidence levels on the 95% level.

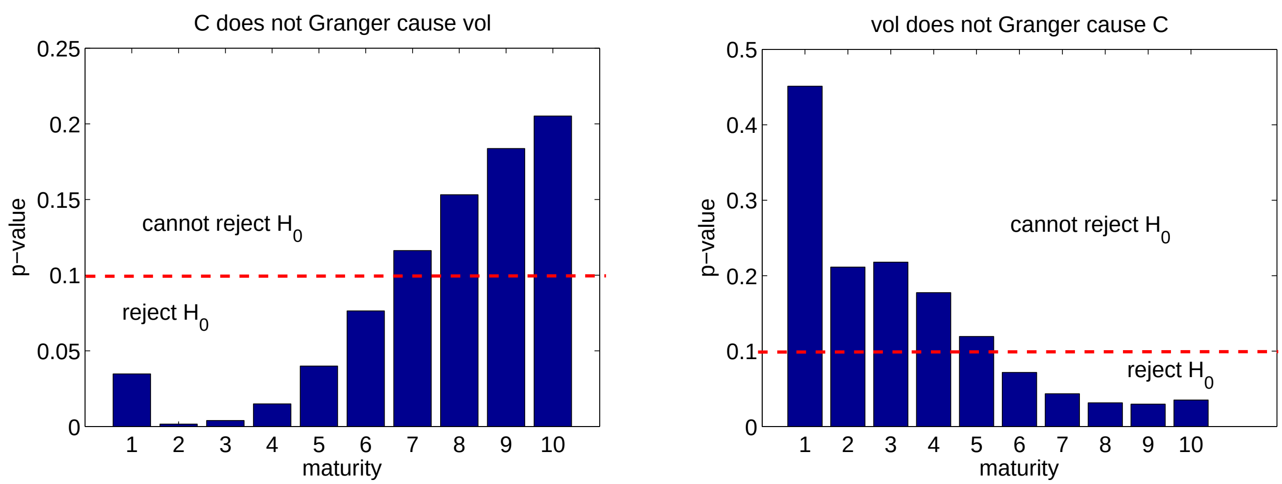

Figure 3. p-Values of Granger causality tests

These figures present results for Granger causality tests. The null hypothesis for the left (right) panel is that negative convexity (volatility) does not Granger cause bond yield volatility (convexity). The regressions are estimated on weekly data from 1997 to 2012.

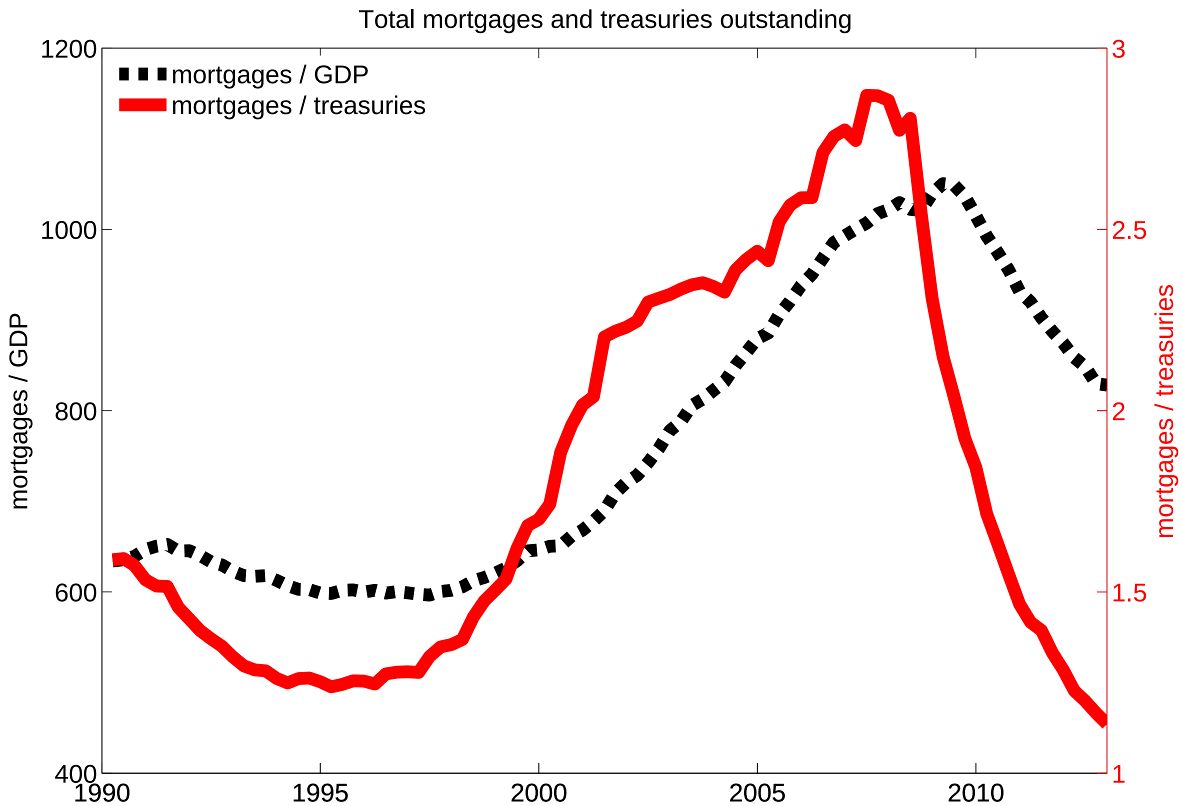

Figure 4. Total mortgages and treasuries outstanding

This figure plots total nominal value outstanding of all mortgages divided by GDP (left axis) and the total amount outstanding of mortgages divided by amount outstanding in Treasuries (right axis). Data is from the webpage of the Board of Governors of the Federal Reserve and its frequency is quarterly from 1990 to 2012.