Internet Appendix for "Mortgage Convexity"

January 2014

A Additional empirical results

A.1 Robustness checks for forecasting results

In this Appendix, I carry out the robustness checks described in the main text. First, I argue that it is difficult to reconcile my findings with the frictionless view that shifts in the quantity of duration have no affect on term premia simply by appealing to simultaneity bias. Second, I control for an exhaustive set of factors that are thought to impact term premia to address concerns that my forecasting results are driven by an omitted variable. Third, I address the standard econometric concerns that arise when estimating time-series forecasting regressions. Finally, I show that similar forecasting results obtain when using Bank of America MBS Indices instead of Barclays' Indices.

A.1.1 Simultaneity bias

First, one might be concerned about the potential for simultaneity bias if one interprets my evidence as a regression of equilibrium prices (i.e., the duration risk premium) on quantities (i.e., the quantity of duration supplied by borrowers). Let e_D denote the inverse demand elasticity and assume that $e_D \geq 0$: when the quantity of duration rises, investors demand a weakly higher term premium. Similarly, let e_S denote the inverse supply elasticity that captures how borrowers respond to variation in term premia. If we regress the equilibrium term premium on the equilibrium quantity of duration, the resulting regression coefficient b_{OLS} equals a weighted average of the inverse supply and inverse demand elasticities, where the weights depend on the volatility of supply and demand shocks

$$b_{OLS} = \frac{\sigma_D^2 e_D^2}{\sigma_D^2 e_D^2 + \sigma_S^2 e_S^2} e_S + \frac{\sigma_S^2 e_S^2}{\sigma_D^2 e_D^2 + \sigma_S^2 e_S^2} e_D.$$

If there are no demand shocks ($\sigma_D^2 = 0$) or supply is completely inelastic ($e_S \to \infty$), the regression identifies an inverse demand curve. In general, of course, a regression of price on quantities does not identify a demand curve.[1]

[1] Let π_t denote the price of interest rate risk—i.e., the term premium on long-term bonds. It is natural to assume that the demand for interest rate risk is increasing in the term premium, so $q_t^D = \overline{q} + \gamma(\pi_t - \overline{\pi}) + \varepsilon_t^D$ with $\gamma > 0$. Suppose the supply of interest rate risk is given by $q_t^S = \overline{q} + \beta(\pi_t - \overline{\pi}) + \varepsilon_t^S$. Although one might expect that $\beta < 0$—i.e., borrowers issue less long-term debt when the term premium is high, the MBS convexity channel suggests that we may have $\beta > 0$. Market equilibrium ($q_t^D = q_t^S$) implies that equilibrium term premium and quantity of duration are

$$\pi_t^* = \overline{\pi} + \frac{\varepsilon_t^S - \varepsilon_t^D}{\gamma - \beta} \text{ and } q_t^* = \overline{q} - \frac{\beta}{\gamma - \beta}\varepsilon_t^D + \frac{\gamma}{\gamma - \beta}\varepsilon_t^S.$$

First, consider what we should expect to find under frictionless null in which demand is infinitely elastic—i.e., $e_D = 0$, so demand curves are flat and the term premia is not impacted by shifts in duration supply. In this case, it is easy to see that $b_{OLS} = 0$, irrespective of the supply elasticity. Thus, one cannot reconcile the finding that $b_{OLS} > 0$ with the frictionless view that shifts in duration supply have no effect on term premia simply by appealing to simultaneity bias. This is an important conclusion: we can learn something important from a regression of prices on quantities, even if this does not allow us to explicitly recover the slope of a demand curve.

Although our regressions suggest that the aggregate demand curve for interest rate risk is not perfectly flat, we are interested in assessing the slope of this demand curve. In general, the simultaneity bias is

$$b_{OLS} - e_D = \frac{\sigma_D^2 e_D^2}{\sigma_D^2 e_D^2 + \sigma_S^2 e_S^2} \left(e_S - e_D \right).$$

If $e_S < 0$—i.e., borrowers issue short when the term premia is high—as one might generally expect, this simultaneity bias should lead me to underestimate the slope of the inverse demand curve. However, if $e_S > 0$—i.e., borrowers issue long when the term premium is high—as suggested by the MBS convexity channel in which a rise in term premia raises long yields and MBS duration, then $b_{OLS} - e_D \propto e_S - e_D$ is ambiguous and b_{OLS} may either overestimate or underestimate e_D. For instance, if $e_S > e_D > 0$—i.e., if demand is highly elastic and the MBS convexity effect is not too strong, we will have $b_{OLS} > e_D > 0$. By contrast, if $e_D > e_S > 0$—i.e., if demand is moderately elastic and if the MBS convexity effect is relatively strong, we will have $e_D > b_{OLS} > 0$. Reassuringly, the fact that the price-impact implied by my regressions is in line with prior estimates from the literature provides some comfort that this bias is small.

A.1.2 Additional controls

Second, one might be concerned that variation in DUR_t does itself not drive the term premium but is simply correlated with some omitted variable that does. In an attempt to deal with this concern, I have rerun my basic forecasting regressions controlling for a more extensive host of variables that the prior literature has argued may impact bond risk-premia. None of these controls has a significant effect on the estimated coefficients on DUR_t or meaningfully alters its statistical significance. The fact that the coefficient on DUR_t is highly robust to the addition of various controls that are thought to be associated with term premia should help to alleviate most natural concerns about omitted variable bias.

Specifically, in Tables A.1 and A.2, I show that I obtain similar results if I control for:

- **Recent movements in long-term interest rates.** One might worry that my results reflect a unrelated "reversal" or "value" effect in long-term bond markets whereby bonds perform well following past underperformance. According to this story, the reason MBS

Suppose that shocks to duration supply and demand are independent. If we regress the equilibrium term premium on the quantity of duration, we get

$$b_{OLS} = \frac{\sigma_D^2 (1/\gamma)^2}{\sigma_D^2 (1/\gamma)^2 + \sigma_S^2 (1/\beta)^2} \frac{1}{\beta} + \frac{\sigma_S^2 (1/\beta)^2}{\sigma_D^2 (1/\gamma)^2 + \sigma_S^2 (1/\beta)^2} \frac{1}{\gamma}.$$

Naturally, this is a weighted average of the inverse supply ($e_S = 1/\beta$) and inverse demand ($e_D = 1/\gamma$) elasticities, where the weights depend on the volatility of supply and demand shocks. If there are no demand shocks ($\sigma_D^2 = 0$) or supply is completely inelastic ($\beta = 0$), the regression identifies an inverse demand curve ($b_{OLS} = 1/\gamma$). Of course, a regression of price on quantities does not generally identify a demand curve.

duration predicts returns is not because of a downward-sloping demand curve for duration risk, but because MBS duration tends to be high when interest rates have recently risen. However, the predictive power of MBS duration is robust to controlling for recent movements in long-term interest rates, so the results do not appear to be simply driven by such a reversal effect.

- **The 8 macroeconomic factors that Ludvigson and Ng (2010) extract from 131 macroeconomic time series.** These series might contain information about risk premia, perhaps because they proxy for time-varying investor risk-aversion or a time-varying quantity of macroeconomic risk.

- **Implied interest rate volatility**. Another concern is that stochastic interest rate volatility might give rise to time-varying term premia; and MBS duration might be picking this up because it is a mechanical function of interest rate volatility.[2]

- **A host of financial market indicators**. I can control for corporate credit spreads, swap spreads, and various equity market variables, including past stock returns and the VIX.

The results do not simply reflect a primitive mean-reversion effect This first robustness check mentioned above is worth stressing further. We can write

$$DUR_t = DUR_{t-1} - CONV_{t-1}(y_t^{(10)} - y_{t-1}^{(10)}) + \varepsilon_t = DUR_0 - \sum_{\tau=0}^{t-1} [CONV_{t-1-\tau}(y_{t-\tau}^{(10)} - y_{t-1-\tau}^{(10)}) + \varepsilon_{t-\tau}].$$

MBS duration reflects a complex weighted average of past changes in interest rates. (Not all past changes in rates count equally—since $CONV_t$ varies over time—and there may be shocks to MBS duration that are exogenous to the path of interest rates.) Thus, one natural concern with the previous evidence is that, instead of reflecting the impact of duration supply on equilibrium *required* returns, these results may simply reflect some primitive form of mean reversion in US bond markets. Specifically, $DUR_t - DUR_{t-1}$ will be positively correlated with $y_t^{(10)} - y_{t-1}^{(10)}$(and $f_t^{(10)} - f_{t-1}^{(10)}$) or negatively correlated with $r_t^{(10)}$. Thus, if $f_t^{(10)} - f_{t-1}^{(10)}$ positively forecast returns (and $r_t^{(10)}$ negatively forecasts returns) we might mistake mean reversion in returns with an MBS convexity channel. In other words, under this competing interpretation the significant forecasting power of DUR_t is driven by an omitted variable—recent changes in rates—which itself drives *expected* returns.

To address this concern, Table A.1 shows horseraces between DUR_t and past changes in forward rates, $f_t^{(10)} - f_{t-1}^{(10)}$. DUR_t is positively related to $f_t^{(10)} - f_{t-1}^{(10)}$ because $DUR_t - DUR_{t-1}$ is strongly positively related to $f_t^{(10)} - f_{t-1}^{(10)}$. Table 4 shows that $f_t^{(10)} - f_{t-1}^{(10)}$ is a surprisingly strong predictor of bond returns from 1989-present. However, both DUR^{MBS} and $DUR\_CNTRB_t^{MBS}$ contain valuable information about future bond excess returns that is above and beyond the information contained in $f_t^{(10)} - f_{t-1}^{(10)}$. That is, there appears to be strong evidence of both the MBS convexity channel and a primitive mean reversion effect over the past 23 years. Nonetheless, these variables do fight with another to some extent—the t-statistics on each typically falls when the other is added, which is not surprising since they are positively correlated and both seem to capture fairly transient components of expected bond returns.

[2] The omitted variable bias here actually goes in the wrong direction. High volatility should be associated with high future returns and low current MBS duration—the delta of the embedded call option that MBS investors have sold short rises when volatility rises.

The finding that $f_t^{(10)} - f_{t-1}^{(10)}$ is such a strong predictor of bond returns—it is highly significant and boosts the R^2 when added to multivariate specifications—is, to the best of my knowledge, new to the literature. Simply put, $f_t^{(10)} - f_{t-1}^{(10)}$ seems to contain information about future returns that is over and above what is reflected in the *current* shape of the yield curve. This looks like medium-frequency "value" or "reversal" effect in bond markets: bonds do poorly when they have done well over the past year. However, understanding what drives this mean reversion effect is beyond the scope of the current paper.

A.1.3 Time-series robustness checks

I now address the standard econometric concerns that arise when estimating time-series forecasting regressions. First, I address several concerns stemming from the moderate length of the baseline sample, which begins in 1989. Second, I address concerns that might arise from the fact that I work with 12-month overlapping returns. In my baseline regressions, I work with 12-month returns largely for the sake of comparability with much of the recent literature on bond risk premia. However, there is very strong evidence of return predictability when I work with 1-month or 3-month returns. Third, I address concerns about finite-sample performance of Newey-West (1987) standard errors. Fourth, I show that the small-sample bias of OLS identified by Stambaugh (1999) is minimal here because the key forecasting variable, MBS duration, is not highly persistent.

Concerns about the moderate sample length One might worry about the statistical soundness of the results given that there are only 23 non-overlapping 12 month returns in the sample. First, it is worth noting that the results are also very strong when forecasting quarterly returns, and there are 93 non-overlapping 3-month periods in the sample. This can be seen in Figure 3 where I use duration to forecast quarterly bond returns. Thus, the strong results are not an artifact of using a modest number of non-overlapping 12-month returns. Second, given that the Agency MBS market only rose to prominence in the mid-1980s, even if we had a measure of effective duration dating back to the advent of this market in the 1970s, one would expect to find economically trivial effects prior to the 1980s. Finally, it is worth noting that a number of prominent studies on bond risk premia use less than 25 years of data—e.g., Campbell and Shiller (1991) work with 1952-1987 data and Fama and Bliss (1987) work with 1964-1984 data.

Concerns about the use of 12-month returns I adress concerns that might arise from the fact that I work with 12-month overlapping returns. In my baseline regressions, I work with 12-month returns largely for the sake of comparability with much of the recent literature on bond risk premia. Furthermore, there is a long-standing recognition in the literature that working with intermediate horizon returns can improve statistical power. However, there is very strong evidence of return predictability when I work with 1-month or 3-month returns. These results are shown in Table A.3 and A.4 and parallel the results for 12-month returns shown in Table 2 of the paper.

Finite sample performance of HAC standard errors It is well known that heteroskedasticity autocorrelation consistent (HAC) variance estimators such as Newey-West (1987) exhibit size distortions in finite samples. Table A.5 addresses this concern in three ways. First, I assess significance using the asymptotic theory developed by Kiefer and Vogelsang (2005), which has better finite-sample properties than traditional asymptotic theory.[3] Using these "fixed-b" asymptotics,

[3]Let m denote the bandwith for HAC standard errors—e.g., I use Newey-West (1987) standard errors with $m = 18$. The usual asymptotic theory for HAC inference is derived under the assumption that $m \to \infty$ and

the key coefficient of interest is typically significant at the 0.1% level. Second, I compute p-values using the stationary block bootstrap of Politis and Romano (1994). The stationary bootstrap re-samples blocks with randomly varying lengths, and I use an average block length of 48 months. Again, the result remains highly significant. Finally, I obtain similar t-statistics using Hansen-Hodrick (1980) standard errors allowing for 12 lags or, following the suggestion of Cochrane (2008) and Bates (2010), using parametric HAC standard errors under the assumption that regression residuals follow an $ARMA(p,q)$ process.[4]

Stambaugh bias is minimal I am estimating a classic predictive regression $y_{t+1} = \alpha + \beta \cdot x_t + u_{t+1}$, where y_{t+1} is a total return that reflects a change in an asset prices from t to $t+1$, x_t is a lagged variable related to asset prices at the end of period t, and u_{t+1} is a residual disturbance. Under the assumption that x_t obeys an $AR(1)$ process $x_{t+1} = \eta + \rho \cdot x_t + v_{t+1}$ with $Cov(u_t, v_t) = \sigma_{uv}$, Stambaugh (1999) shows that

$$E[\widehat{\beta} - \beta] = \frac{\sigma_{uv}}{\sigma_v^2} E[\widehat{\rho} - \rho] = -\frac{\sigma_{uv}}{\sigma_v^2}\left(\frac{1+3\rho}{T}\right) + O\left(1/T^2\right), \tag{1}$$

so the sign of the finite-sample bias is opposite that of σ_{uv}. When the disturbances are correlated, $\widehat{\beta}$ inherits the well known downward bias when estimating ρ in small samples (Kendall (1954)).

In the present case y_{t+1} represents excess returns on long-term Treasury bonds and x_t represents MBS duration. Suppose there is positive shock to long-term yields in a given period. This positive shock to long-term yields corresponds to negative value of u_{t+1}—excess returns are low when bond yields rise unexpectedly. Due to the negative convexity of MBS, this positive shock to bond yields will lengthen the duration of MBS securities which corresponds to a positive value of v_{t+1}. This suggests that $\sigma_{uv} < 0$ which, as shown in equation (1), leads to an upward bias in $\widehat{\beta}$ when T is small. However, one would not expect the bias to be large in this case since MBS duration is not very persistent. Instead, it exhibits a lot of high frequency variation.

As shown in Table A.5, the Stambaugh bias is tiny and is not a major concern in this setting. For the univariate regressions, I use the procedure of Amihud and Hurvich (2004) to compute bias-adjusted estimates of β and associated standard errors. Although $\phi = \sigma_{uv}/\sigma_v^2$ is large and negative here, the downward bias in ρ, $E[\widehat{\rho} - \rho] = -(1+3\rho)/T$, is tiny since the regressor is not very persistent and we have a reasonably long time-series. For multivariate regressions, I use the simulation procedure in Baker and Stein (2004) to compute bias-adjusted estimates.[5] Again, the bias is minimal.

$m/T \to 0$. Kiefer and Vogelsang (2005) instead proceed under the assumption that $m = bT$ for some $b \in (0,1]$. That is, the resulting "fixed-b" asymptotics assume that the bandwidth is a fixed fraction of the sample size.

[4] If true 1-period *expected* returns follow an $AR(1)$ process and realized returns are expected returns plus white noise, then realized 1-period returns follow an $ARMA(1,1)$ and realized k-period cumulative returns follow an $ARMA(1,k)$. Thus, once we abandon the null of no return predictability, we should expect serial correlation even in a non-overlapping forecasting regression under the null that a *given* predictor has no forecasting power. I follow the implementation of this ARMA-HAC variance estimator in Greenwood and Hanson (2012).

[5] I first simulate $y_{t+1} = \alpha + \boldsymbol{\beta}'\mathbf{x}_t + u_{t+1}$ and $\mathbf{x}_{t+1} = \boldsymbol{\theta} + \boldsymbol{\Gamma}\mathbf{x}_t + \mathbf{v}_{t+1}$ recursively, using the OLS coefficient estimates and drawing with replacement from the empirical distribution of errors, u and $\mathbf{v}$. I throw out the first 100 draws and draw T additional observations. I estimate $y_{t+1} = \alpha + \boldsymbol{\beta}'\mathbf{x}_t + u_{t+1}$ on each simulated sample, giving us a set of coefficients $\mathbf{b}^*$. The bias-adjusted estimate is $\mathbf{b}_{\text{adj}} = \mathbf{b} - \left(\overline{\mathbf{b}^*} - \mathbf{b}\right)$—i.e., I adjust the OLS estimate by subtracting off the bootstrap bias estimate: the mean of $\mathbf{b}^*$ minus the OLS estimate.

A.1.4 Alternate measures of MBS duration produce similar results

Table A.6 reproduces the key forecasting results using the Bank of America (formerly Merrill Lynch) MBS indices. This measure of effective MBS duration is available beginning in February 1991 from Datastream. The correlation between the Bank of America and Barclays MBS duration measures is 0.79. As shown in Table A.6, we obtain very similar results using the Bank of America index in place of the Barlcays index. Specifically, although the Bank of America is not as strong of a univariate return predictor, both variables perform similarly in multivariate regressions controlling for the term spread or forward rates. This suggests that the results are not driven by any special property of the Barclays indices or their specific methodology for computing effective MBS duration. Rather, the broad conclusion should hold for any reasonable, forward-looking measure of effective MBS duration.

A.2 Additional predictions

A.2.1 Asymetric persistence of duration shocks

The logic in Section 3.2.3 suggests an asymmetry in the persistence of MBS duration shocks that is clearly evident in the data. To see this, I generalize the simple $AR(1)$ model to allow for different rates of mean-reversion when duration is above or below some specific threshold, μ: $DUR_{t+1}^{MBS} = \theta + \rho_0 \cdot \left(DUR_t^{MBS} - \mu\right)^- + \rho_1 \cdot \left(DUR_t^{MBS} - \mu\right)^+ + \varepsilon_{t+1}$ where $(x)^- = \min\{x, 0\}$ and $(x)^+ = \max\{x, 0\}$. Estimating this using non-linear least squares, I obtain

$$DUR_{t+1}^{MBS} = \underset{[7.08]}{3.001} + \underset{[6.41]}{0.775} \cdot \left(DUR_t^{MBS} - \underset{[6.41]}{3.050}\right)^- + \underset{[16.77]}{0.973} \cdot \left(DUR_t^{MBS} - \underset{[6.41]}{3.050}\right)^+,$$

with $R^2 = 0.78$. The estimates imply that large contractions in MBS duration are expected to be far more transient than large extensions in MBS duration: the implied half-life of a shock is 25 months when MBS duration is high, but only 3 months when duration is low. Indeed, the coefficients on $\left(DUR_t^{MBS} - \mu\right)^+$ and $\left(DUR_t^{MBS} - \mu\right)^-$ differ statistically at the 10% level.

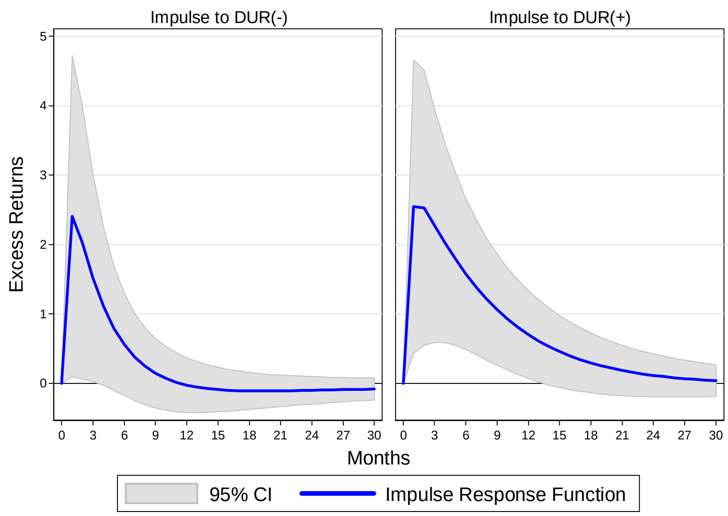
This asymmetric persistence property has implications for the relationship between MBS duration and bond yields. From the standpoint of forecasting bond returns over the following instant, I have argued that the relevant variable is the current quantity of duration. However, a large increase in MBS duration is expected to persist for longer and, thus, have a more persistent impact on bond risk premia, than a decline. And this implies that a large increase in MBS duration should have a larger impact on *current long-term yields*—which reflect expectations of future bond risk premia over multiple periods—than a large drop in duration. In other words, the onset of a prepayment drought should be associated with a larger rise in yields, while the onset of a refinancing boom should be associated with a smaller drop in current yields. Specifically, if MBS duration is already high (e.g., above the threshold μ), then a further increase in duration is expected to persist for longer and, thus, should be associated with a larger increase in long-term yields.

In the absence of a clean instrument for exogenous shocks to mortgage duration that are otherwise unrelated to changes in rates, it is difficult to test for an asymmetric response of yields.[6]

[6] Since DUR_t is function of $y_t^{(n)}$, one worries that estimating $\Delta y_t^{(n)} = a + b_1 \cdot \Delta\left(DUR_t - \gamma\right)^+ + b_0 \cdot \Delta\left(DUR_t - \gamma\right)^- + \mathbf{c}'\Delta\mathbf{x}_t + \Delta\varepsilon_t^{(n)}$ and finding $b_1 > b_0$ (as is clearly the case in the data) is recovering a mechanical fact about MBS duration as opposed to a deeper fact about equilibrium term premium. The problem is that MBS become more negatively convex at lower yields. When yields are low, a given change in $\Delta(DUR_t - \gamma)^-$ is mechanically associated with a small change in $\Delta y_t^{(n)}$. And when yields are high, a given change in $\Delta(DUR_t - \gamma)^+$

Nonetheless, this observation might partially help to explain the violent bond-market sell-offs of 1994, 1999, 2003 when MBS duration extended significantly.[7]

However, I can apply a simple VAR methodology to investigate the asymmetric persistence of MBS shocks. Specifically, let $DUR_t^+ = \max\{DUR_t - \hat{\mu}, 0\}$ and $DUR_t^- = \min\{DUR_t - \hat{\mu}, 0\}$ denote the positive and negative deviation of DUR_t from its estimated mean. I then estimate a first order monthly VAR using rx_t, DUR_t^+, DUR_t^-, and TS_t. Using the estimated coefficients from this VAR, I can then compare the return forecasting consequences from an impulse to DUR_t^+ (i.e., a shock that triggers a refincing drought) to an impulse to DUR_t^- (i.e., a shock that triggers a refinancing boom). Doing so, I obtain the following simple impulse response functions. This suggests that, starting from a situation where duration is at its long-run mean, positive and negative shocks to duration have a similar effect on excess returns in the following month. However, a positive shock is expected to persist far longer and therefore to have a larger effect on current yields.



.

B Extended Model with Multiple Maturities

In this Appendix I extend the model with only short and long-term bonds in the main text to allow for multiple bond maturities. The extension is a discrete time version of Vayanos and Vila (2009) and Greenwood and Vayanos (2013). I add the MBS convexity effect to this work horse model in a simple fashion: I assume that the aggregate supply of duration which must be held by arbitrageurs rises when interest rates rise. However, I assume that these duration shocks dissipate quickly—i.e., duration shocks are transitory and are expected to mean-revert over time. These assumptions are a simple way to capture the dynamics of MBS duration explained in the paper. Thus, the my model is similar to one developed by Malkhozov, Mueller, Vedolin, and Venter (2013) who also extend the Vayanos and Vila (2009) framework to sudy the effect of MBS duration. (They work is continuous time whereas I work in discrete time.)

---

is mechanically associated with a large change in $\Delta y_t^{(n)}$.

[7] Market participants were surprised by the pace of Fed tightening in early 1994. From January to May, long-term yields spiked and MBS duration rapidly extended from 3.35 to its sample maximum of 4.84 years. Carrol and Lappen (1994) and Fernald, Keane, and Mosser (1994) argue that this duration extension contributed to the bond market dislocations of 1994 when the market for collateralized mortgage obligations (CMOs) collapsed and several hedge funds and dealers specializing in mortgages suffered major losses.

Solving the model, I obtain a discrete-time affine model of the term structure with two state variables, the current short rate and the current level of MBS duration which depends on the past path of interest rates. The model generates the following predictions:

1. An MBS duration shock raises the *current* expected excess returns on long-term bonds over short-term bonds. Furthermore, since MBS duration is expected to quickly revert to its long-run mean, this effect is short-lived in expectation.

2. A stronger MBS convexity effect (i.e., a more negatively convex MBS universe) increases the sensitivity of long-term yields and forward rates to movements in short rates.

3. A stronger MBS convexity effect increases the volatility of long-term yields and forwards.

These results echo those from the simple model in the text. However, the multiple maturity extension delivers the following additional predictions that I can take to the data:

4. An MBS duration shock has a larger effect on the expected excess returns on longer-term bonds than on those on intermediate-term bonds. This is a natural consequence of the fact that an MBS duration shock raises the current duration risk premium in bond markets: the expected returns on long duration bonds move more than those on intermediate duration bonds.

5. Transient shocks to MBS duration have a humped-shaped effect on the yield curve and the forward rate curve. That is, a shock to MBS duration increases the curvature of the yield curve. This suggests that transient shocks to MBS duration may account for some of the predictive power of the tent-shaped combination of forward rates identified by Cochrane and Piassezi (2005) since the tent factor picks up times when curvature is high.

B.1 Set-up

Let $r_t = y_t^{(1)}$ denote the log return on 1-period riskless bonds between time t and $t+1$. I assume that the short-term rate, r_t, is exogenously pinned down by monetary policy. Alternately, one can assume that r_t is pinned down by a riskfree storage technology in perfectly elastic supply.

Let $y_t^{(n)}$ denote the log yield of n-period bonds and $p_t^{(n)} = -ny_t^{(n)}$ the log price of n-period bonds. The log excess return on n-period bonds over 1-period bonds from t to $t+1$ is

$$rx_{t+1}^{(n)} = p_{t+1}^{(n-1)} - p_t^{(n)} - r_t = ny_t^{(n)} - (n-1)\, y_{t+1}^{(n-1)} - r_t. \tag{2}$$

Suppose that there are bonds with maturiities $n = 2, ..., N$. Let $\mathbf{b}_t^0 {=} [b_t^{(2)}, b_t^{(3)}, ...b_t^{(N)}]'$ denote the $(N-1) \times 1$ vector of arbitrageurs' bond holdings and $\mathbf{rx}_{t+1}^0 {=} [rx_{t+1}^{(2)}, rx_{t+1}^{(3)}, ...rx_{t+1}^{(N)}]'$ denote the $(N-1) \times 1$ vector of excess returns over 1-period bonds.

At each date, arbitrageurs have mean-variance utility over 1-period ahead wealth. Thus, at time t, arbitrageurs solve

$$\max_{\mathbf{b}_t^0} \left\{ (\mathbf{b}_t^0)' E_t \left[\mathbf{rx}_{t+1}^0\right] - \frac{1}{2\tau} (\mathbf{b}_t^0)' Var_t \left[\mathbf{rx}_{t+1}^0\right] (\mathbf{b}_t^0)' \right\}, \tag{3}$$

so their demand for bonds at time t is

$$\mathbf{b}_t^{0*} = \tau \left(Var_t \left[\mathbf{rx}_{t+1}^0\right]\right)^{-1} E_t \left[\mathbf{rx}_{t+1}^0\right]. \tag{4}$$

Suppose that the net supply of these bonds at time t is $\mathbf{q}_t^0 = [q_t^{(2)}, q_t^{(3)}, ...q_t^{(N)}]'$. In order for the markets for bonds with maturiities $n = 2, ..., N$ to clear, we require

$$E_t[\mathbf{rx}_{t+1}^0] = \tau^{-1} Var_t[\mathbf{rx}_{t+1}^0]\mathbf{q}_t^0 = \tau^{-1} Cov_t[\mathbf{rx}_{t+1}^0, rx_{t+1}^P] \tag{5}$$

where $rx_{t+1}^P = \left(\mathbf{rx}_{t+1}^0\right)' \mathbf{q}_t^0$ is the excess return on the portfolio held by arbitrageurs. Thus, since $E_t\left[rx_{t+1}^P\right] = \left(\mathbf{q}_t^0\right)' E_t[\mathbf{rx}_{t+1}^0] = \tau^{-1}\left(\mathbf{q}_t^0\right)' Var_t[\mathbf{rx}_{t+1}^0]\left(\mathbf{q}_t^0\right) = \tau^{-1} Var_t\left[rx_{t+1}^P\right]$, we obtain a conditional-CAPM for bonds

$$E_t[\mathbf{rx}_{t+1}^0] = \frac{Cov_t[\mathbf{rx}_{t+1}^0, rx_{t+1}^P]}{Var_t\left[rx_{t+1}^P\right]} E_t\left[rx_{t+1}^P\right] = \boldsymbol{\beta}_t^0 E_t\left[rx_{t+1}^P\right]. \tag{6}$$

Naturally, the portfolio of bonds held by arbitrageurs (rx_{t+1}^P) functions as the "market portfolio" from the standpoint of the bond pricing kernel.

B.2 Closing the Model

Let the vectors $\mathbf{b}_t = [0, \mathbf{b}_t^{0\prime}]'$, $\mathbf{rx}_{t+1} = [0, \mathbf{rx}_{t+1}^{0\prime}]'$, and $\mathbf{q}_t = [0, \mathbf{q}_t^{0\prime}]$ denote the $N \times 1$ analogs of $\mathbf{b}_t^0$, $\mathbf{rx}_{t+1}^0$, $\mathbf{q}_t^0$; in each case I have simply appended an intial 0 to the analogous $(N-1) \times 1$ vector.

I suppose the short rate follows an exogenous AR(1) process

$$r_{t+1} = \overline{r} + \rho_r\left(r_t - \overline{r}\right) + \varepsilon_{r,t+1}, \tag{7}$$

where $Vab\left[\varepsilon_{r,t+1}\right] = \sigma_r^2$. Thus, we have $E_t\left[r_{t+j}\right] = \overline{r} + \rho_r^j\left(r_t - \overline{r}\right)$, so that the short rate is expected to revert to its long-run mean of $\overline{r}$ at rate ρ_r per period.

I next suppose that the net supply of bonds that arbitrageurs must hold is

$$\mathbf{q}_t = \mathbf{q}\left(s_t\right) \equiv \mathbf{q}_0 + \mathbf{q}_1 s_t \tag{8}$$

where

$$s_{t+1} = \rho_s s_t + C \cdot \varepsilon_{r,t+1}. \tag{9}$$

Thus, we have $E_t\left[s_{t+j}\right] = \rho_s^j s_t$ and $\sigma^2\left[s_t\right] = C^2\sigma_r^2/\left(1 - \rho_s^2\right)$. I assume this is pure duration shock in the sense that (i) $q_1^{(n)}$ is increasing in n, but (ii) $\mathbf{1}'\mathbf{q}_1 = \sum_{n=1}^N q_1^{(n)} = 0$: an increase in s_t shifts the maturity structure of arbitrageurs' bond portfolios, but leaves the total notional size of those portfolios unchanged. For simplicity, I assume that the

The parameter $C > 0$ captures the MBS convexity effect: duration rises when interest rates rise. Furthermore, the MBS maturity shocks I study are transient and are characterized by small values of ρ_s, perhaps near 0 at an annual horizon. In summary, s_t captures the path-dependency induced by mortgage refinancing: the amount of duration that arbitrageurs must hold depends on the past path of rates.

B.3 Equilibrium prices and yields

Recall that $p_t^{(n)} = -ny_t^{(n)}$ and $y_t^{(n)} = -p_t^{(n)}/n$. Let $\mathbf{p}_t = [p_t^{(1)}, p_t^{(2)}, p_t^{(3)}, ...p_t^{(N)}]'$ denote the $N \times 1$ vector of log bond prices and $\mathbf{y}_t = [y_t^{(1)}, y_t^{(2)}, y_t^{(3)}, ...y_t^{(N)}]'$ the $N \times 1$ vector of log bond yields. I conjecture that

$$\mathbf{p}_t = \mathbf{b}_0 + \mathbf{b}_r r_t + \mathbf{b}_s s_t \tag{10}$$

and

$$\mathbf{y}_t = \mathbf{a}_0 + \mathbf{a}_r r_t + \mathbf{a}_s s_t. \tag{11}$$

As in Vayanos and Vila (2009) and Greenwood and Vayanos (2013), we can solve for the unknown parameters $\mathbf{b}_0$, $\mathbf{b}_r$, and $\mathbf{b}_s$ (equivalently, $\mathbf{a}_0$, $\mathbf{a}_r$, and $\mathbf{a}_s$) by using the method of undetermined coefficients to match terms in the market-clearing condition

$$E_t\left[\mathbf{rx}_{t+1}\right] = \tau^{-1} Var_t\left[\mathbf{rx}_{t+1}\right]\mathbf{q}\left(s_t\right). \tag{12}$$

Specificically, letting $\mathbf{I}^{(sub)}$ denote the $N \times N$ matrix of zeros with 1s along the subdiagonal, the vector of excess returns is given by

$$\mathbf{rx}_{t+1} = \mathbf{I}^{(sub)}\mathbf{p}_{t+1} - \mathbf{p}_t - r_t\mathbf{1}, \tag{13}$$

so we need to match terms in

$$E_t\left[\mathbf{I}^{(sub)}\mathbf{p}_{t+1} - \mathbf{p}_t - r_t\mathbf{1}\right] = \tau^{-1} Var_t\left[\mathbf{I}^{(sub)}\mathbf{p}_{t+1} - \mathbf{p}_t - r_t\mathbf{1}\right]\mathbf{q}\left(s_t\right). \tag{14}$$

I now discuss the key properties of the extended model.

Bond risk premia Bond risk premia are

$$E_t[rx_{t+1}^{(n)}] = E_t[ny_t^{(n)} - (n-1)\, y_{t+1}^{(n-1)} - r_t] = \underbrace{rp_0^{(n)}}_{\text{RP (uncond)}} + \underbrace{rp_s^{(n)} \cdot s_t}_{\text{RP (cond)}}. \tag{15}$$

In words, bond risk premia can be decomposed into a unconditional piece ($rp_0^{(n)} > 0$) that is independent of of MBS duration (s_t) and a conditional piece that depends on the current level of MBS duration ($rp_s^{(n)} > 0$). It immediately follows that a high level of MBS duration at time t predicts higher excess returns on long-term bonds from t to $t+1$. Furthermore, when ρ_s is small, MBS duration shocks dissipate quickly and only have a short-lived effect on bond risk premia.

We also have the following results:

1. Both $rp_0^{(n)}$ and $rp_s^{(n)}$ are increasing in maturity, n.

2. Unconditional and conditional risk premia are both increasing in the persistence of short-rate shocks and duration shocks:$\partial(rp_0^{(n)})/\partial\rho_r > 0$, $\partial(rp_0^{(n)})/\partial\rho_s > 0$, $\partial(rp_s^{(n)})/\partial\rho_r > 0$, and $\partial(rp_s^{(n)})/\partial\rho_s > 0$.

3. Both the unconditional and conditional risk premia are increasing in the strength of the MBS convexity effect: $\partial(rp_0^{(n)})/\partial C > 0$ and $\partial(rp_s^{(n)})/\partial C > 0$.

Yields and forward Yields on n-period bonds take the following form

$$
\begin{aligned}
y_t^{(n)} &= a_0^{(n)} + a_r^{(n)} r_t + a_s^{(n)} \cdot s_t \qquad (16) \\
&= \underbrace{\left(\bar{r} - n^{-1}\frac{1-\rho_r^n}{1-\rho_r}\bar{r}\right)}_{a_0^{(n)}} + \underbrace{\left(n^{-1}\frac{1-\rho_r^n}{1-\rho_r}\right)}_{a_r^{(n)}} \cdot r_t + \underbrace{\left(n^{-1}\textstyle\sum_{j=1}^n \rho_s^{n-j} rp_s^{(j)}\right)}_{a_s^{(n)}} \cdot s_t \\
&= \underbrace{\left(\bar{r} + n^{-1}\frac{1-\rho_r^n}{1-\rho_r}(r_t - \bar{r})\right)}_{EH^{(n)}(r_t) = n^{-1}\sum_{j=1}^n E_t[r_{t+j-1}]} + \underbrace{\left(n^{-1}\textstyle\sum_{j=1}^n rp_0^{(j)}\right)}_{TP^{(n)}_{(uncond)}} + \underbrace{\left(n^{-1}\textstyle\sum_{j=1}^n \rho_s^{n-j} rp_s^{(j)}\right) \cdot s_t}_{TP^{(n)}_{(cond)}}.
\end{aligned}
$$

In words, yields can be decomposed into an expectations hypothesis term ($EH^{(n)}(r_t) = n^{-1}\sum_{j=1}^n E_t[r_{t+j-1}]$), an unconditional term premium term ($TP^{(n)}_{(uncond)} = n^{-1}\sum_{j=1}^n rp_0^{(j)}$) which is independent of MBS duration, and a conditional term premium term ($(n^{-1}\sum_{j=1}^n \rho_s^{n-j} rp_s^{(j)}) \cdot s_t$) which depends on the level of MBS duration. The expectations and unconditional term premia terms take the same form as in a discrete time Vasicek (1977) model (see e.g., Campbell, Lo, and MacKinlay (1997)).

We have the following properties:

1. When the MBS convexity effect is present ($C > 0$), long-term yields are excessively senstivity to movements in short rates and, therefore, are excessively volatile. Furthermore, these effects are more pronounced when the convexity effect is strong (C is larger) or when MBS supply shocks are more persistent. Formally, the yield coefficients on s_t are increasing in the strength of the MBS convexity effect ($\partial a_s^{(n)}/\partial C > 0$ for all n) and the persistence of the MBS supply shocks ($\partial a_s^{(n)}/\partial \rho_s > 0$ for all n).

2. As in any Vasicek-type model, shocks to the short-rate impact the slope of the yield curve via a pure expectations hypothesis effect. Specifically, $n^{-1}\sum_{j=0}^{n-1}\rho_r^j = n^{-1}(1-\rho_r^n)/(1-\rho_r)$ is decreasing in n.

3. $TP^{(n)}_{(uncond)} = n^{-1}\sum_{j=1}^n rp_0^{(j)}$ is increasing in n. This follows immediately since $rp_0^{(n)}$ is increasing in n.

4. $TP^{(n)}_{(cond)} = n^{-1}\sum_{j=1}^n \rho_s^{n-j} rp_s^{(j)}$ is either increasing in n or a hump-shaped function of n. Specifically, when ρ_s is small, $TP^{(n)}_{(cond)}$ is a hump-shaped function of n.

Forward rates are given by

$$
\begin{aligned}
f_t^{(n)} &= n y_t^{(n)} - (n-1) y_t^{(n-1)} \\
&= \underbrace{\left(\bar{r} + \rho_r^{n-1}(r_t - \bar{r})\right)}_{\text{EH}=E_t[r_{t+n-1}]} + \underbrace{rp_0^{(n)}}_{\text{RP (uncond)}} + \underbrace{\left(rp_s^{(n)} - (1-\rho_s)\left(\textstyle\sum_{j=1}^{n-1}\rho_s^{n-1-j} rp_s^{(j)}\right)\right) \cdot s_t}_{\text{TP (cond)}}.
\end{aligned}
$$

When duration shocks are highly transient (i.e., when ρ_s is low), a shock to duration has a humped-shaped effect on the yield curve and the forward curve. Specifically, $a_s^{(n)}$ and $f_s^{(n)} = n a_s^{(n)} - (n-1) a_s^{(n-1)}$ are both hump-shaped functions of n. This suggests that transient shocks to MBS duration may account for some of the predictive power of the tent-shaped linear combination of foreward rates identified by Cochrane and Piassezi (2005). However, any state variable which has a transient effect on bond risk premia should impact the curvature of the yield curve in a similar fashion. This a a simple matter of no arbitrage bond pricing logic, so there is a sense in which the Cochrane and Piassezi (2005) factor may work as a catch all for all factors that generate transient variation in risk premia. By contrast, when duration shocks are highly persistent (i.e., ρ_s is high), a shock to duration steepends the yield curve and forward curve. Specifically, $a_s^{(n)}$ and $f_s^{(n)} = na_s^{(n)} - (n-1)\, a_s^{(n-1)}$ are increasing functions of n.

B.4 Model Solution

To begin, we know that $[\mathbf{b}_0]_1 = 0$, $[\mathbf{b}_r]_1 = -1$, and $[\mathbf{b}_s]_1 = 0$. Making use of our conjecture, excess returns are

$$\mathbf{rx}_{t+1} = \left(\mathbf{I}^{(sub)} - \mathbf{I}\right)\mathbf{b}_0 + \left(\mathbf{I}^{(sub)} r_{t+1} - \mathbf{I} r_t\right)\mathbf{b}_r + \left(\mathbf{I}^{(sub)} s_{t+1} - \mathbf{I} s_t\right)\mathbf{b}_s - r_t\mathbf{1}$$

Using this expression for excess returns, we can then express $E_t\left[\mathbf{rx}_{t+1}\right]$ and $Var_t\left[\mathbf{rx}_{t+1}\right]$ in terms of the unknown model parameters. Specifically, we have

$$\begin{aligned} E_t\left[\mathbf{rx}_{t+1}\right] &= \left(\mathbf{I}^{(sub)} - \mathbf{I}\right)\mathbf{b}_0 + \left(\mathbf{I}^{(sub)} E_t[r_{t+1}] - \mathbf{I} r_t\right)\mathbf{b}_r + \left(\mathbf{I}^{(sub)} E_t[s_{t+1}] - \mathbf{I} s_t\right)\mathbf{b}_s - r_t\mathbf{1} \\ &= [(\mathbf{I}^{(sub)} - \mathbf{I})\mathbf{b}_0 + \mathbf{I}^{(sub)}\overline{r}\,(1-\rho_r)\,\mathbf{b}_r] + [\left(\mathbf{I}^{(sub)}\rho_r - \mathbf{I}\right)\mathbf{b}_r - \mathbf{1}]r_t + \left[\left(\mathbf{I}^{(sub)}\rho_s - \mathbf{I}\right)\mathbf{b}_s\right] s_t, \end{aligned}$$

and

$$Var_t\left[\mathbf{rx}_{t+1}\right] = Var\left[\mathbf{I}^{(sub)}\left(\mathbf{b}_r + C\mathbf{b}_s\right)\varepsilon_{r,t+1}\right] = \sigma_r^2 \mathbf{I}^{(sub)}\left(\mathbf{b}_r + C\mathbf{b}_s\right)\left(\mathbf{b}_r + C\mathbf{b}_s\right)' \mathbf{I}^{(sub)'}.$$

We then can solve for the unknown coefficients by plugging the above expressions into the equilibrium condition to obtain

$$\begin{aligned} & [(\mathbf{I}^{(sub)} - \mathbf{I})\mathbf{b}_0 + \mathbf{I}^{(sub)}\overline{r}\,(1-\rho_r)\,\mathbf{b}_r] + [\left(\mathbf{I}^{(sub)}\rho_r - \mathbf{I}\right)\mathbf{b}_r - \mathbf{1}]r_t + \left[\left(\mathbf{I}^{(sub)}\rho_s - \mathbf{I}\right)\mathbf{b}_s\right] s_t \\ = \; & \tau^{-1}\sigma_r^2\left[\mathbf{I}^{(sub)}\left(\mathbf{b}_r + C\mathbf{b}_s\right)\left(\mathbf{b}_r + C\mathbf{b}_s\right)' \mathbf{I}^{(sub)'}\right]\left(\mathbf{q}_0 + \mathbf{q}_1 s_t\right) \end{aligned}$$

Matching terms in the above equation, we obtain three conditions

$$\text{Constant} \; : \; \left(\mathbf{I}^{(sub)} - \mathbf{I}\right)\mathbf{b}_0 + \mathbf{I}^{(sub)}\overline{r}\,(1-\rho_r)\,\mathbf{b}_r = \mathbf{B}\mathbf{q}_0 \qquad (17)$$

$$r_t \; : \; \left(\mathbf{I}^{(sub)}\rho_r - \mathbf{I}\right)\mathbf{b}_r - \mathbf{1} = \mathbf{0} \qquad (18)$$

$$s_t \; : \; \left(\mathbf{I}^{(sub)}\rho_s - \mathbf{I}\right)\mathbf{b}_s = \mathbf{B}\mathbf{q}_1 \qquad (19)$$

where

$$\mathbf{B} = \tau^{-1} Var\left[\mathbf{I}^{(sub)}\left(\mathbf{b}_r + C\mathbf{b}_s\right)\varepsilon_{r,t+1}\right] = \tau^{-1}\sigma_r^2 \mathbf{I}^{(sub)}\left(\mathbf{b}_r + C\mathbf{b}_s\right)\left(\mathbf{b}_r + C\mathbf{b}_s\right)' \mathbf{I}^{(sub)'}. \qquad (20)$$

Note that the top-most row and left-most collumn of $\mathbf{B}$ will be zeros by construction. These three conditions are the vector analogs of functional conditions that appear in Vayanos and Vila (2009) and Greenwood and Vayanos (2013).

I now characterize the equilibrium coefficients on the short rate, the supply shock, and the constant terms in greater detail.

Coefficients on r_t The solution for $\mathbf{b}_r$ takes the standard form

$$\mathbf{b}_r^* = \left[\mathbf{I}^{(sub)}\rho_r - \mathbf{I}\right]^{-1}\mathbf{1}. \tag{21}$$

For $n > 1$, the condition for r_t implies that

$$b_r^{(n)} = b_r^{(n-1)}\rho_r - 1.$$

Since $b_r^{(1)} = -1$, we have

$$b_r^{(n)} = -\left(\sum_{j=0}^{n-1}\rho_r^j\right) = -\left(\frac{1-\rho_r^n}{1-\rho_r}\right),$$

and

$$a_r^{(n)} = n^{-1}\left(\sum_{j=0}^{n-1}\rho_r^j\right) = n^{-1}\left(\frac{1-\rho_r^n}{1-\rho_r}\right).$$

In other words, $a_r^{(n)}$ and $b_r^{(n)}$ take the same form as in a discrete-time analog of the Vasicek (1977) model (see, e.g., Campbell, Lo, and MacKinlay (1997)). Note that $b_r^{(n)}$ is a decreasing function of n. It is also easy to see that $a_r^{(n)}$ is a decreasing function of n for all $0 < \rho_r < 1$.

Coefficients on s_t We can write the condition for $\mathbf{b}_s^*$ as

$$\begin{aligned}
\left[\mathbf{I}^{(sub)}\rho_s - \mathbf{I}\right]\mathbf{b}_s^* &= \mathbf{B}^*\mathbf{q}_1 \\
&= \tau^{-1}\sigma_r^2\mathbf{I}^{(sub)}\left(\mathbf{b}_r^* + C\mathbf{b}_s^*\right)\left(\mathbf{b}_r^* + C\mathbf{b}_s^*\right)'\mathbf{I}^{(sub)\prime}\mathbf{q}_1 \\
&= \mathbf{I}^{(sub)}\left(\mathbf{b}_r^* + C\mathbf{b}_s^*\right)\left(\frac{\sigma_r^2}{\tau}\mathbf{b}_r^{*\prime}\mathbf{q}_1^{(sub)} + C\frac{\sigma_r^2}{\tau}\mathbf{b}_s^{*\prime}\mathbf{q}_1^{(sub)}\right) \\
&= \mathbf{I}^{(sub)}\left(\mathbf{b}_r^* + C\mathbf{b}_s^*\right)\left(\lambda_{r1}^* + C\lambda_{s1}^*\right)
\end{aligned}$$

where $\mathbf{q}_1^{(sub)} = \mathbf{I}^{(sub)\prime}\mathbf{q}_1$, $\lambda_{r1}^* = \sigma_r^2\mathbf{b}_r^{*\prime}\mathbf{q}_1^{(sub)}/\tau$, and $\lambda_{s1}^* = \sigma_r^2\mathbf{b}_s^{*\prime}\mathbf{q}_1^{(sub)}/\tau$.

So we require

$$\mathbf{b}_s^* = \left[\mathbf{I}^{(sub)}\left(\rho_s - C\left(\lambda_{r1}^* + C\lambda_{s1}^*\right)\right) - \mathbf{I}\right]^{-1}\mathbf{I}^{(sub)}\mathbf{b}_r\left(\lambda_{r1}^* + C\lambda_{s1}^*\right) \tag{22}$$

where

$$\lambda_{s1}^* = \frac{\sigma_r^2}{\tau}\mathbf{q}_1^{(sub)\prime}\left[\mathbf{I}^{(sub)}\left(\rho_s - C\left(\lambda_{r1}^* + C\lambda_{s1}^*\right)\right) - \mathbf{I}\right]^{-1}\mathbf{I}_N^{(sub)}\mathbf{b}_r\left(\lambda_{r1}^* + C\lambda_{s1}^*\right). \tag{23}$$

Below we use these two equations for solve numerically for $\mathbf{b}_s^*$.

To see the intuition, note that the condition for s_t for $n > 1$ implies that

$$b_s^{(n)} = \rho_s b_s^{(n-1)} - rp_s^{(n)}$$

where

$$rp_s^{(n)} \equiv \frac{\partial E_t[rx_{t+1}^{(n)}]}{\partial s_t} = \left[\mathbf{B}^*\mathbf{q}_1\right]_n$$

is the conditional one-period risk-premium on n-period bonds associated with a shift in s_t. Since $b_s^{(1)} = -rp_s^{(1)} = 0$, we have

$$b_s^{(n)} = -\left(\sum_{j=1}^{n}\rho_s^{n-j}rp_s^{(j)}\right),$$

and

$$a_s^{(n)} = n^{-1}\left(\sum_{j=1}^{n}\rho_s^{n-j} rp_s^{(j)}\right).$$

Intuitvely, $a_s^{(n)}$ reflects the expected future risk-premia associated with a currently increase in duration supply.

How do shifts in MBS duration, captured here by shifts in s_t, affect bond risk premia along the curve? I now show that $rp_s^{(n)} \equiv \partial E_t[rx_{t+1}^{(n)}]/\partial s_t$ is an increasing function of maturity n. To see this, note that $rp_s^{(1)} = 0$ and

$$rp_s^{(n)} = \left(b_r^{(n-1)} + Cb_s^{(n-1)}\right)\left(\lambda_{r1} + C\lambda_{s1}\right)$$

for $n > 1$. Thus, we have

$$\begin{aligned} -b_s^{(n)} &= -\rho_s b_s^{(n-1)} + rp_s^{(n)} \\ &= -b_s^{(n-1)}\rho_s + \left(b_r^{(n-1)} + Cb_s^{(n-1)}\right)\left(\lambda_{r1} + C\lambda_{s1}\right), \end{aligned}$$

which implies that

$$-b_s^{(n)} = \left(\sum_{j=2}^{n}\rho_s^{n-j}\cdot\left(-b_r^{(j-1)} - Cb_s^{(j-1)}\right)\left(-\lambda_{r1} - C\lambda_{s1}\right)\right).$$

Under the assumption that $\lambda_{r1} = \sigma_r^2 \mathbf{b}_r' \mathbf{q}_1^{(sub)}/\tau < 0$ and $\lambda_{s1} = \sigma_r^2 \mathbf{b}_s' \mathbf{q}_1^{(sub)}/\tau < 0$, it follows that $-b_s^{(n)}$ is an increasing function of n. Since $-b_r^{(n)}$ is also an increasing function of n, it follows that $rp_s^{(n)} = \left(b_r^{(n-1)} + Cb_s^{(n-1)}\right)\left(\lambda_{r1} + C\lambda_{s1}\right)$ is increasing in n.

How do shifts in MBS duration affect the shape of the yield and forward curves?

Proposition 1 *The persistence of duration shocks, ρ_s, plays a critical role in the way that the $a_s^{(n)}$ vary as a function of n. Specifically, while it is always they case that $na_s^{(n)} = \sum_{j=1}^{n}\rho_s^{n-j}\cdot rp_s^{(j)}$ is increasing in n, $a_s^{(n)}$ is a hump-shaped function of n when ρ_s is low or moderate and a monotically increasing function of n where ρ_s is large.*

Proof: To see this, we show that $a_s^{(n)} = -b_s^{(n)}/n$ is an increasing function when $\rho_s = 1$ and a hump-shaped function of n where $\rho_s = 0$. The result then follows by continuity of the model solution in ρ_s.

First, note that when $\rho_s = 1$, we have $a_s^{(n)} = n^{-1}\sum_{j=1}^{n} rp_s^{(j)}$. Since we showed that $rp_s^{(j)}$ increasing in j, it follows that $a_s^{(n)}$ is increasing in n where $\rho_s = 1$.

Next, when $\rho_s = 0$, we have $b_s^{(n)} = -rp_s^{(n)}$ and $a_s^{(n)} = rp_s^{(n)}/n$. Note that we always have $a_s^{(1)} = 0 < rp_s^{(2)}/2 = a_s^{(2)}$. Furthermore, we have

$$\begin{aligned} 0 &> a_s^{(n)} - a_s^{(n-1)} = \frac{rp_s^{(n)}}{n} - \frac{rp_s^{(n-1)}}{n-1} \\ &\Leftrightarrow rp_s^{(n)} - rp_s^{(n-1)} < n^{-1} rp_s^{(n)} \end{aligned}$$

Note that when $\rho_s = 0$, we have

$$\begin{aligned} rp_s^{(n)} &= \left(-b_r^{(n-1)} - Cb_s^{(n-1)}\right)\left(-\lambda_{r1} - C\lambda_{s1}\right) \\ &= \left(-b_r^{(n-1)} + Crp_s^{(n-1)}\right)\left(-\lambda_{r1} - C\lambda_{s1}\right). \end{aligned}$$

Thus, we have $rp_s^{(1)} = 0$, $rp_s^{(2)} = (-\lambda_{r1} - C\lambda_{s1})$, and

$$rp_s^{(n)} = (\lambda_{r1} + C\lambda_{s1}) \cdot \sum_{j=1}^{n-1} \left(-C\lambda_{r1} - C^2\lambda_{s1}\right)^{n-1-j} b_r^{(j)}.$$

Thus, $rp_s^{(n)} - rp_s^{(n-1)} < n^{-1} rp_s^{(n)}$ holds if

$$\begin{aligned} & (\lambda_{r1} + C\lambda_{s1})\, b_r^{(n-1)} - (\lambda_{r1} + C\lambda_{s1}) \left(1 + C\lambda_{r1} + C^2\lambda_{s1}\right) \sum_{j=1}^{n-2} \left(-C\lambda_{r1} - C^2\lambda_{s1}\right)^{n-2-j} b_r^{(j)} \\ < \; & \frac{(\lambda_{r1} + C\lambda_{s1}) \sum_{j=1}^{n-1} \left(-C\lambda_{r1} - C^2\lambda_{s1}\right)^{n-1-j} b_r^{(j)}}{n} \end{aligned}$$

or

$$b_r^{(n-1)} < \frac{\left[\left(n^{-1}\left(1 + C\lambda_{r1} + C^2\lambda_{s1}\right) - C\lambda_{r1} - C^2\lambda_{s1}\right)\right]}{\left(n^{-1} - 1\right)} \sum_{j=1}^{n-2} \left(-C\lambda_{r1} - C^2\lambda_{s1}\right)^{n-2-j} b_r^{(j)}.$$

For n large the right-hand side approaches

$$\left(C\lambda_{r1} + C^2\lambda_{s1}\right) \sum_{j=1}^{n-2} \left(-C\lambda_{r1} - C^2\lambda_{s1}\right)^{n-2-j} b_r^{(j)} > 0 > b_r^{(n-1)}.$$

Thus, when $\rho_s = 0$, we must have $a_s^{(n-1)} > a_s^{(n)}$ for n sufficiently large.

Intercepts Finally, $\mathbf{b}_0$ follows from

$$\begin{aligned} \mathbf{b}_0^* &= \left[\mathbf{I}^{(sub)} - \mathbf{I}\right]^{-1} \mathbf{I}^{(sub)} \left(\mathbf{b}_r^* + C\mathbf{b}_s^*\right) \left(\frac{\sigma_r^2}{\tau} \mathbf{b}_r^{*\prime} \mathbf{q}_0^{(sub)} + C \frac{\sigma_r^2}{\tau} \mathbf{b}_s^{*\prime} \mathbf{q}_0^{(sub)}\right) \\ & \quad - \left[\mathbf{I}^{(sub)} - \mathbf{I}\right]^{-1} \mathbf{I}^{(sub)} \overline{r} \left(1 - \rho_r\right) \mathbf{b}_r^*. \end{aligned} \tag{24}$$

We have $b_0^{(n)} = rp_0^{(n)} = 0$. For $n > 1$, the condition for intercepts implies that

$$b_0^{(n-1)} - b_0^{(n)} + \overline{r}\left(1 - \rho_r\right) \cdot b_r^{(n-1)} = rp_0^{(n)},$$

where $rp_0^{(n)} = [\mathbf{B}\mathbf{q}_0]_n$ is the unconditional risk premium for bonds of maturity n (i.e., the risk premium when $s_t = 0$ which is also the average risk premium since $E[s_t] = 0$). Using $b_r^{(n-1)} = -\left(1 - \rho_r^n\right) / \left(1 - \rho_r\right)$, this equals

$$b_0^{(n)} = b_0^{(n-1)} - \overline{r}\left(1 - \rho_r^n\right) - rp_0^{(n)}.$$

Thus, we have

$$\begin{aligned} b_0^{(1)} &= 0 \\ b_0^{(2)} &= -\overline{r}\left(1 - \rho_r\right) - rp_0^{(2)} \\ & \;\;\ldots \\ b_0^{(n)} &= -n\overline{r} + \overline{r} \frac{1 - \rho_r^n}{1 - \rho_r} - \textstyle\sum_{j=1}^{n} rp_0^{(j)} \end{aligned}$$

and

$$a_0^{(n)} = \overline{r} - \frac{1}{n}\frac{1-\rho_r^n}{1-\rho_r}\overline{r} + \frac{1}{n}\sum_{j=1}^{n} rp_0^{(j)}. \tag{25}$$

Yields, forwards, and bond risk premia Putting all of this together, bond risk premia are

$$\begin{aligned} E_t[rx_{t+1}^{(n)}] &= E_t[ny_t^{(n)} - (n-1)\, y_{t+1}^{(n-1)} - r_t] \\ &= \underbrace{rp_0^{(n)}}_{\text{RP (uncond)}} + \underbrace{rp_s^{(n)} \cdot s_t}_{\text{RP (cond)}}. \end{aligned}$$

In vector notation, we have

$$E_t\left[\mathbf{rx}_{t+1}\right] = \mathbf{Bq}_0 + \mathbf{Bq}_1 \cdot s_t = \mathbf{rp}_0 + \mathbf{rp}_s \cdot s_t.$$

where

$$\begin{aligned} \mathbf{rp}_0 &= \mathbf{I}^{(sub)}\left(\mathbf{b}_r + C\mathbf{b}_s\right)\left(\frac{\sigma_r^2}{\tau}\mathbf{b}_r'\mathbf{q}_0^{(sub)} + C\frac{\sigma_r^2}{\tau}\mathbf{b}_s'\mathbf{q}_0^{(sub)}\right) \\ &= \mathbf{I}^{(sub)}\left(\mathbf{b}_r + C\mathbf{b}_s\right)\left(\lambda_{r0} + C\lambda_{s0}\right) \end{aligned}$$

and

$$\begin{aligned} \mathbf{rp}_s &= \mathbf{I}^{(sub)}\left(\mathbf{b}_r + C\mathbf{b}_s\right)\left(\frac{\sigma_r^2}{\tau}\mathbf{b}_r'\mathbf{q}_1^{(sub)} + C\frac{\sigma_r^2}{\tau}\mathbf{b}_s'\mathbf{q}_1^{(sub)}\right) \\ &= \mathbf{I}^{(sub)}\left(\mathbf{b}_r + C\mathbf{b}_s\right)\left(\lambda_{r1} + C\lambda_{s1}\right). \end{aligned}$$

where $\mathbf{q}_0^{(sub)} = \mathbf{I}^{(sub)'}\mathbf{q}_0$ and $\mathbf{q}_1^{(sub)} = \mathbf{I}^{(sub)'}\mathbf{q}_1$.

Similarly, yields are

$$\begin{aligned} y_t^{(n)} &= a_0^{(n)} + a_r^{(n)} r_t + a_s^{(n)} \cdot s_t \\ &= \underbrace{\left(\overline{r} - n^{-1}\frac{1-\rho_r^n}{1-\rho_r}\overline{r}\right)}_{a_0^{(n)}} + \underbrace{\left(n^{-1}\frac{1-\rho_r^n}{1-\rho_r}\right)}_{a_r^{(n)}} \cdot r_t + \underbrace{\left(n^{-1}\sum_{j=1}^{n} \rho_s^{n-j} rp_s^{(j)}\right)}_{a_s^{(n)}} \cdot s_t \\ &= \underbrace{\left(\overline{r} + n^{-1}\frac{1-\rho_r^n}{1-\rho_r}\left(r_t - \overline{r}\right)\right)}_{EH^{(n)}(r_t) = n^{-1}\sum_{j=1}^{n} E_t[r_{t+j-1}]} + \underbrace{\left(n^{-1}\sum_{j=1}^{n} rp_0^{(j)}\right)}_{TP_{(uncond)}^{(n)}} + \underbrace{\left(n^{-1}\sum_{j=1}^{n} \rho_s^{n-j} rp_s^{(j)}\right) \cdot s_t}_{TP_{(cond)}^{(n)}}. \end{aligned}$$

Forward rates are given by

$$\begin{aligned} f_t^{(n)} &= ny_t^{(n)} - (n-1)\, y_t^{(n-1)} \\ &= \underbrace{\left(\overline{r} + \rho_r^{n-1}\left(r_t - \overline{r}\right)\right)}_{\text{EH}=E_t[r_{t+n-1}]} + \underbrace{rp_0^{(n)}}_{\text{RP (uncond)}} + \underbrace{\left(rp_s^{(n)} - (1-\rho_s)\left(\sum_{j=1}^{n-1} \rho_s^{n-1-j} rp_s^{(j)}\right)\right) \cdot s_t}_{\text{TP (cond)}}. \end{aligned}$$

B.5 Comparative Statics

As in Vayanos and Vila (2009) and Greenwood and Vayanos (2013), we need some natural regularity conditions for the model to be well behaved.

1. Suppose that $\lambda^*_{r1} < 0$
2. Need τ large enough for everything to be well behaved
3. Need $q_1' 1 = 0$ and $q_1' b_r < 0$.

Let $\boldsymbol{\alpha}_r = -\mathbf{b}_r$, $\gamma_{r1} = -\lambda_{r1}$, and $\gamma_{s1} = -\lambda_{s1}$, we have

$$\gamma^*_{s1} = \frac{\sigma_r^2}{\tau}\mathbf{q}_1^{(sub)\prime}\left[\mathbf{I} - \mathbf{I}^{(sub)}\left(\rho_s + C\gamma_{r1} + C^2\gamma^*_{s1}\right)\right]^{-1}\mathbf{I}^{(sub)}\boldsymbol{\alpha}_r\left(\gamma_{r1} + C\gamma^*_{s1}\right).$$

For any parameter x, we have

$$\begin{aligned}
&\frac{\partial \gamma^*_{s1}}{\partial x}\left(\begin{array}{c} 1 - \frac{\sigma_r^2}{\tau}\mathbf{q}_1^{(sub)\prime}\left[\mathbf{I} - \mathbf{I}^{(sub)}\left(\rho_s + C\gamma_{r1} + C^2\gamma^*_{s1}\right)\right]^{-1}\mathbf{I}^{(sub)}\boldsymbol{\alpha}_r C \\ -\frac{\sigma_r^2}{\tau}\mathbf{q}_1^{(sub)\prime}\left[\mathbf{I} - \mathbf{I}^{(sub)}\left(\rho_s + C\gamma_{r1} + C^2\gamma^*_{s1}\right)\right]^{-1} C^2\mathbf{I}^{(sub)}\left[\mathbf{I} - \mathbf{I}^{(sub)}\left(\rho_s + C\gamma_{r1} + C^2\gamma^*_{s1}\right)\right]^{-1}\mathbf{I}^{(sub)}\boldsymbol{\alpha}_r\left(\gamma_{r1} + C\gamma^*_{s1}\right)\end{array}\right) \\
&= \frac{\partial}{\partial x}\frac{\sigma_r^2}{\tau}\mathbf{q}_1^{(sub)\prime}\left[\mathbf{I} - \mathbf{I}^{(sub)}\left(\rho_s + C\gamma_{r1} + C^2\gamma^*_{s1}\right)\right]^{-1}\mathbf{I}^{(sub)}\boldsymbol{\alpha}_r\left(\gamma_{r1} + C\gamma^*_{s1}\right).
\end{aligned}$$

For τ sufficiently large enough and for $\gamma_{r1} > 0$, we have:

1. γ^*_{s1} is increasing in ρ_s

$$\frac{\partial \gamma^*_{s1}}{\partial \rho_s} = \frac{\frac{\sigma_r^2}{\tau}\mathbf{q}_1^{(sub)\prime}\left[\mathbf{I} - \mathbf{I}^{(sub)}\left(\rho_s + C\gamma_{r1} + C^2\gamma^*_{s1}\right)\right]^{-1}\mathbf{I}^{(sub)}\left[\mathbf{I} - \mathbf{I}^{(sub)}\left(\rho_s + C\gamma_{r1} + C^2\gamma^*_{s1}\right)\right]^{-1}\mathbf{I}^{(sub)}\boldsymbol{\alpha}_r\left(\gamma_{r1} + C\gamma^*_{s1}\right)}{\begin{array}{c} 1 - \frac{\sigma_r^2}{\tau}\mathbf{q}_1^{(sub)\prime}\left[\mathbf{I} - \mathbf{I}^{(sub)}\left(\rho_s + C\gamma_{r1} + C^2\gamma^*_{s1}\right)\right]^{-1}\mathbf{I}^{(sub)}\boldsymbol{\alpha}_r C \\ -\frac{\sigma_r^2}{\tau}\mathbf{q}_1^{(sub)\prime}\left[\mathbf{I} - \mathbf{I}^{(sub)}\left(\rho_s + C\gamma_{r1} + C^2\gamma^*_{s1}\right)\right]^{-1} C^2\mathbf{I}^{(sub)}\left[\mathbf{I} - \mathbf{I}^{(sub)}\left(\rho_s + C\gamma_{r1} + C^2\gamma^*_{s1}\right)\right]^{-1}\mathbf{I}^{(sub)}\boldsymbol{\alpha}_r\left(\gamma_{r1} + C\gamma^*_{s1}\right)\end{array}} > 0.$$

2. γ^*_{s1} is increasing in C

$$\frac{\partial \gamma^*_{s1}}{\partial C} = \frac{\begin{array}{c}\frac{\sigma_r^2}{\tau}\mathbf{q}_1^{(sub)\prime}\left[\mathbf{I}-\mathbf{I}^{(sub)}\left(\rho_s + C\gamma_{r1} + C^2\gamma^*_{s1}\right)\right]^{-1}\left(\gamma_{r1}+2C\gamma^*_{s1}\right)\mathbf{I}^{(sub)}\left[\mathbf{I}-\mathbf{I}^{(sub)}\left(\rho_s + C\gamma_{r1} + C^2\gamma^*_{s1}\right)\right]^{-1}\mathbf{I}^{(sub)}\boldsymbol{\alpha}_r\left(\gamma_{r1}+C\gamma^*_{s1}\right)\\ +\frac{\sigma_r^2}{\tau}\mathbf{q}_1^{(sub)\prime}\left[\mathbf{I}-\mathbf{I}^{(sub)}\left(\rho_s + C\gamma_{r1} + C^2\gamma^*_{s1}\right)\right]^{-1}\mathbf{I}^{(sub)}\boldsymbol{\alpha}_r\gamma^*_{s1}\end{array}}{\begin{array}{c}1-\frac{\sigma_r^2}{\tau}\mathbf{q}_1^{(sub)\prime}\left[\mathbf{I}-\mathbf{I}^{(sub)}\left(\rho_s + C\gamma_{r1} + C^2\gamma^*_{s1}\right)\right]^{-1}\mathbf{I}^{(sub)}\boldsymbol{\alpha}_r C\\ -\frac{\sigma_r^2}{\tau}\mathbf{q}_1^{(sub)\prime}\left[\mathbf{I}-\mathbf{I}^{(sub)}\left(\rho_s + C\gamma_{r1} + C^2\gamma^*_{s1}\right)\right]^{-1}C^2\mathbf{I}^{(sub)}\left[\mathbf{I}-\mathbf{I}^{(sub)}\left(\rho_s + C\gamma_{r1} + C^2\gamma^*_{s1}\right)\right]^{-1}\mathbf{I}^{(sub)}\boldsymbol{\alpha}_r\left(\gamma_{r1}+C\gamma^*_{s1}\right)\end{array}} > 0.$$

3. γ^*_{s1} is increasing in γ_{r1}

$$\frac{\partial \gamma^*_{s1}}{\partial \gamma_{r1}} = \frac{\begin{array}{c}\frac{\sigma_r^2}{\tau}\mathbf{q}_1^{(sub)\prime}\left[\mathbf{I}-\mathbf{I}^{(sub)}\left(\rho_s + C\gamma_{r1} + C^2\gamma^*_{s1}\right)\right]^{-1}C\mathbf{I}^{(sub)}\left[\mathbf{I}-\mathbf{I}^{(sub)}\left(\rho_s + C\gamma_{r1} + C^2\gamma^*_{s1}\right)\right]^{-1}\mathbf{I}^{(sub)}\boldsymbol{\alpha}_r\left(\gamma_{r1}+C\gamma^*_{s1}\right)\\ +\frac{\sigma_r^2}{\tau}\mathbf{q}_1^{(sub)\prime}\left[\mathbf{I}-\mathbf{I}^{(sub)}\left(\rho_s + C\gamma_{r1} + C^2\gamma^*_{s1}\right)\right]^{-1}\mathbf{I}^{(sub)}\boldsymbol{\alpha}_r\end{array}}{\begin{array}{c}1-\frac{\sigma_r^2}{\tau}\mathbf{q}_1^{(sub)\prime}\left[\mathbf{I}-\mathbf{I}^{(sub)}\left(\rho_s + C\gamma_{r1} + C^2\gamma^*_{s1}\right)\right]^{-1}\mathbf{I}^{(sub)}\boldsymbol{\alpha}_r C\\ -\frac{\sigma_r^2}{\tau}\mathbf{q}_1^{(sub)\prime}\left[\mathbf{I}-\mathbf{I}^{(sub)}\left(\rho_s + C\gamma_{r1} + C^2\gamma^*_{s1}\right)\right]^{-1}C^2\mathbf{I}^{(sub)}\left[\mathbf{I}-\mathbf{I}^{(sub)}\left(\rho_s + C\gamma_{r1} + C^2\gamma^*_{s1}\right)\right]^{-1}\mathbf{I}^{(sub)}\boldsymbol{\alpha}_r\left(\gamma_{r1}+C\gamma^*_{s1}\right)\end{array}} > 0.$$

Next let $\boldsymbol{\alpha}_s = -\mathbf{b}_s$, $\gamma_{r1} = -\lambda_{r1}$, and $\gamma_{s1} = -\lambda_{s1}$, so that

$$\boldsymbol{\alpha}_s = \left[\mathbf{I}-\mathbf{I}^{(sub)}\left(\rho_s + C\gamma_{r1} + C^2\gamma^*_{s1}\right)\right]^{-1}\mathbf{I}^{(sub)}\boldsymbol{\alpha}_r\left(\gamma_{r1}+C\gamma^*_{s1}\right)$$

Thus:

1. The coefficients $\boldsymbol{\alpha}_s$ are larger when C is larger.

$$\begin{aligned}\frac{\partial \boldsymbol{\alpha}_s}{\partial C} &= \left[\mathbf{I}-\mathbf{I}^{(sub)}\left(\rho_s + C\gamma_{r1} + C^2\gamma^*_{s1}\right)\right]^{-1}\mathbf{I}^{(sub)}\left(\gamma_{r1}+2C\gamma^*_{s1}+C^2\frac{\partial\gamma^*_{s1}}{\partial C}\right)\left[\mathbf{I}-\mathbf{I}^{(sub)}\left(\rho_s + C\gamma_{r1} + C^2\gamma^*_{s1}\right)\right]^{-1}\mathbf{I}^{(sub)}\boldsymbol{\alpha}_r\left(\gamma_{r1}+C\gamma^*_{s1}\right)\\ &\quad+\left[\mathbf{I}-\mathbf{I}^{(sub)}\left(\rho_s + C\gamma_{r1} + C^2\gamma^*_{s1}\right)\right]^{-1}\mathbf{I}^{(sub)}\boldsymbol{\alpha}_r\left(\gamma^*_{s1}+C\frac{\partial\gamma^*_{s1}}{\partial C}\right) > \mathbf{0}.\end{aligned}$$

A larger value of C (strong MBS convexity effect) increases the excess sensitivity of long-term yields and forward rates. A larger value of C also increases the excess volatility of long-term yields and forward rates.

2. The coefficients $\boldsymbol{\alpha}_s$ are larger when ρ_s is larger:

$$\begin{aligned}\frac{\partial \boldsymbol{\alpha}_s}{\partial \rho_s} &= \left[\mathbf{I} - \mathbf{I}^{(sub)}\left(\rho_s + C\gamma_{r1} + C^2\gamma_{s1}^*\right)\right]^{-1}\mathbf{I}^{(sub)}\left(1 + C^2\frac{\partial \gamma_{s1}^*}{\partial \rho_s}\right)\left[\mathbf{I} - \mathbf{I}^{(sub)}\left(\rho_s + C\gamma_{r1} + C^2\gamma_{s1}^*\right)\right]^{-1}\mathbf{I}^{(sub)}\boldsymbol{\alpha}_r\left(\gamma_{r1} + C\gamma_{s1}^*\right)\\ &\quad + \left[\mathbf{I} - \mathbf{I}^{(sub)}\left(\rho_s + C\gamma_{r1} + C^2\gamma_{s1}^*\right)\right]^{-1}\mathbf{I}^{(sub)}\boldsymbol{\alpha}_r\left(C\frac{\partial \gamma_{s1}^*}{\partial \rho_s}\right) > \mathbf{0}.\end{aligned}$$

When ρ_s is large, MBS shocks have more persistent effects on term premia and, therefore, a large impact on current yields.

3. Larger when $\boldsymbol{\alpha}_r$ is larger (i.e., when ρ_r is larger);

$$\frac{\partial \boldsymbol{\alpha}_s}{\partial \boldsymbol{\alpha}_r'} = \left[\mathbf{I} - \mathbf{I}^{(sub)}\left(\rho_s + C\gamma_{r1} + C^2\gamma_{s1}^*\right)\right]^{-1}\left(\gamma_{r1} + C\gamma_{s1}^*\right)\mathbf{I}^{(sub)}\mathbf{I} > \mathbf{0}$$

4. Larger when γ_{r1} is larger;

$$\begin{aligned}\frac{\partial \boldsymbol{\alpha}_s}{\partial \gamma_{r1}} &= \left[\mathbf{I} - \mathbf{I}^{(sub)}\left(\rho_s + C\gamma_{r1} + C^2\gamma_{s1}^*\right)\right]^{-1}\left(C + C^2\frac{\partial \gamma_{s1}^*}{\partial \gamma_{r1}}\right)\mathbf{I}^{(sub)}\left[\mathbf{I} - \mathbf{I}^{(sub)}\left(\rho_s + C\gamma_{r1} + C^2\gamma_{s1}^*\right)\right]^{-1}\mathbf{I}^{(sub)}\boldsymbol{\alpha}_r\left(\gamma_{r1} + C\gamma_{s1}^*\right)\\ &\quad + \left[\mathbf{I} - \mathbf{I}^{(sub)}\left(\rho_s + C\gamma_{r1} + C^2\gamma_{s1}^*\right)\right]^{-1}\mathbf{I}^{(sub)}\boldsymbol{\alpha}_r > \mathbf{0}\end{aligned}$$

B.6 Numerical Illustration

B.6.1 Solving the Model Numerically

To solve the model numerically, we need to find a vector $\mathbf{b}_s$ that solves equations (22) and (23). We can solve for the relevant fixed point numerically using a simple recursive approach:

1. Start with an initial guess for $\mathbf{b}_s$, say $\mathbf{b}_s^{(0)}$. For instance, start with the value corresponding to $C = 0$ which is
$$\mathbf{b}_s^{(0)} = \left[\mathbf{I}^{(sub)}\rho_s - \mathbf{I}\right]^{-1} \mathbf{I}^{(sub)}\mathbf{b}_r\lambda_{r1}.$$

2. Start the following loop and then iterate until convergence:

   (a) Update λ_{s1} as
$$\lambda_{s1}^{(j)} = \frac{\sigma_r^2}{\tau}\mathbf{q}_1^{(sub)\prime}\mathbf{b}_s^{(j)}.$$

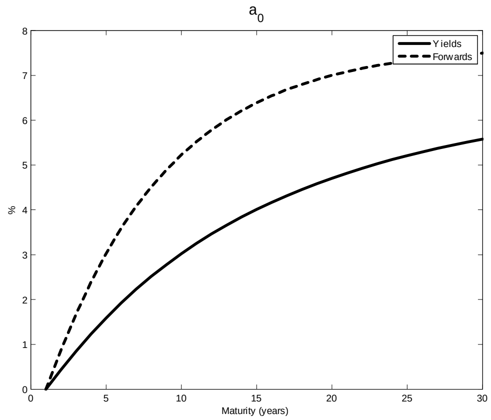
   (b) Update $\mathbf{b}_s^{(j+1)}$
$$\mathbf{b}_s^{(j+1)} = \left[\mathbf{I}^{(sub)}\left(\rho_s - C(\lambda_{r1} + C\lambda_{s1}^{(j)})\right) - \mathbf{I}\right]^{-1} \mathbf{I}^{(sub)}\mathbf{b}_r(\lambda_{r1} + C\lambda_{s1}^{(j)}).$$

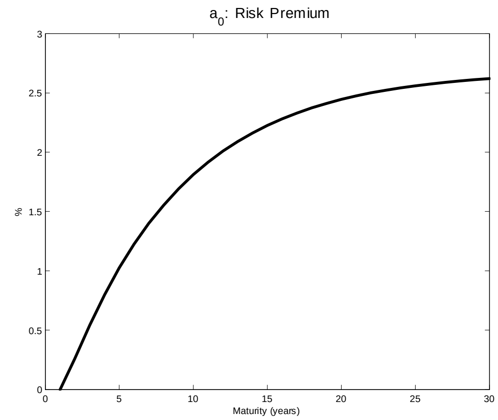
B.6.2 Model Solution

I solve the model numerically using the following parameters:

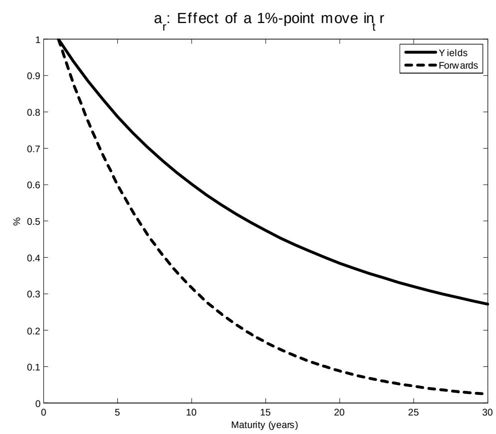
- $r = 5\%$;
- $\rho_r = 0.88$;
- $\rho_s = 0.15$;
- $\sigma_r = 1.5\%$;
- $C = 2$;
- $\tau = 4$;
- $N = 30$
- $q_{0n} = 0.2$ for $n = 1, ..., N$
- $q_{1n} = 9 \times (n - 16)$ for $n = 2, ..., N$ so that $\sum_{n=2}^{N} q_{1n} = 0$.

$\mathbf{a}_0$ coeffcients Below I plot the unconditional components of yields $a_0^{(n)}$ and forwards $f_0^{(n)} = na_0^{(n)} - (n-1)\, a_0^{(n-1)}$ versus maturity n. I also plot the unconditional component of bond risk premium $rp_0^{(n)}$ versus n.

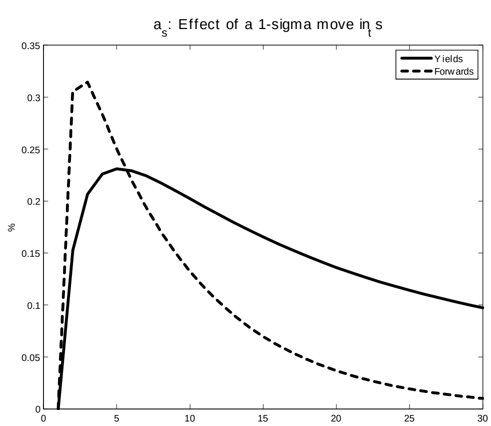


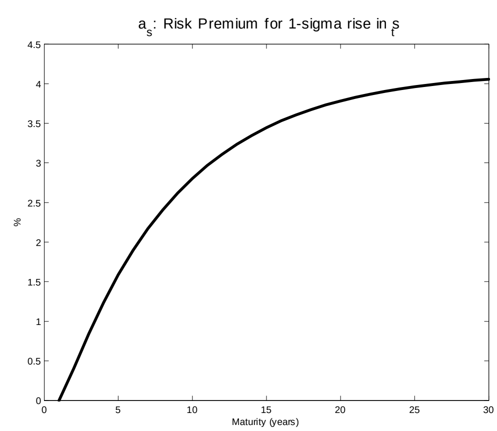


$\mathbf{a}_r$ **coeffcients** Below I plot the conditional components of yields $a_r^{(n)}$ and forwards $f_r^{(n)} = na_r^{(n)} - (n-1)\, a_r^{(n-1)}$ versus maturity n.

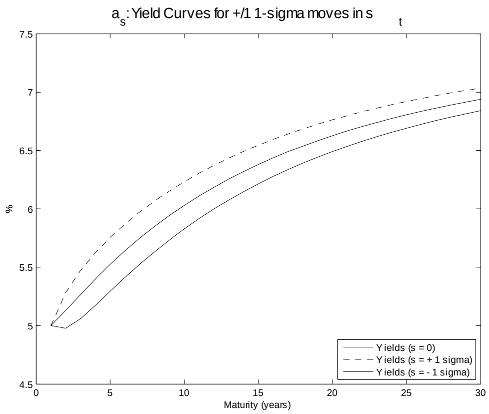


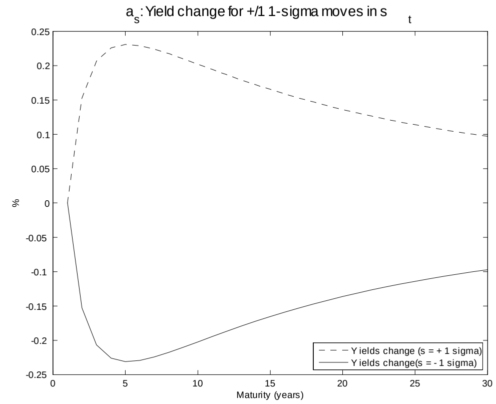
$\mathbf{a}_s$ **coeffcients** Below I plot the conditional components of yields $a_s^{(n)}$ and forwards $f_s^{(n)} = na_s^{(n)} - (n-1)\, a_s^{(n-1)}$ versus maturity n. I also plot the conditional component of bond risk premium $rp_s^{(n)}$ versus n. Since ρ_s is small, an increase in s_t has a hump-shaped effect on both the yield and forward rate curves.





Impact of a supply shock on yield curve Below I plot the yield curve when $r_t = \overline{r}$ following a $+1$ standard deviation and -1 standard deviation shock to MBS duration, s_t.

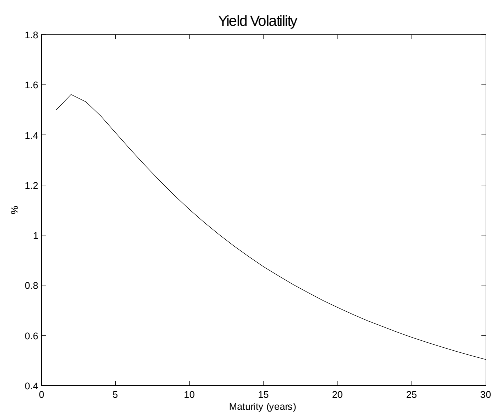




Yield Volatility We have

$$\begin{aligned} Var[\Delta y_t^{(n)}] &= Var\left[a_r^{(n)} \Delta r_t + a_s^{(n)} \Delta s_t\right] \\ &= \left[\left(a_r^{(n)}\right)^2 + \left(C a_s^{(n)}\right)^2 + 2C a_r^{(n)} a_s^{(n)}\right] \sigma_r^2 \end{aligned}$$

Below I plot $Var[\Delta y_t^{(n)}]$ versus n:



C Extended Model with Multiple Asset Classes

MBS investors can dynamically delta-hedge the varying interest rate exposure of MBS using Treasuries or interest rate swaps. They can also statically hedge their exposures by purchasing interest rate options. Regardless of the specific hedging instrument and whether MBS investors pursue a dynamic or a static hedging strategy, other investors must take the other side of these trades. It follows that, as a whole, fixed income investors must bear a time-varying amount of interest rate risk which impacts equilibrium term premia. Furthermore, in a simple model in which Treasuries and swaps are perfect substitutes, these different hedging activities would each impact duration risk premia in the same way, so one would not expect shifts in MBS duration to impact spreads between, say, swaps and Treasuries.

But different long-term fixed income assets with the same duration are not perfect substitutes.

Once we relax the assumption of perfect substitutability, hedging flows triggered by shifts in MBS duration may impact *yield spreads* between various *duration-matched* fixed income assets and give rise to the predictable variation in the excess returns on various fixed income assets over *duration-matched* government bonds. I consider a stylized example with three long-term fixed income asset classes: long-term government bonds (g), long-term swaps (s), and long-term high grade corporate bonds (c). Suppose fraction α of MBS duration Q_t is hedged via the government bond market and $(1-\alpha)$ is hedged via the interest rate swap market. Market clearing in the market for long-term government bonds (g), long-term swaps (s), and long-term corporate bonds (c) implies that

$$\begin{bmatrix} E_t[rx^g_{t+1}] \\ E_t[rx^s_{t+1}] \\ E_t[rx^c_{t+1}] \end{bmatrix} = \gamma \begin{bmatrix} \sigma^2_r + \sigma^2_g & \sigma^2_r & \sigma^2_r \\ \sigma^2_r & \sigma^2_r + \sigma^2_s & \sigma^2_r \\ \sigma^2_r & \sigma^2_r & \sigma^2_r + \sigma^2_c \end{bmatrix} \begin{bmatrix} q^g_0 + \alpha Q_t \\ q^s_0 + (1-\alpha) Q_t \\ q^c_0 \end{bmatrix}. \tag{26}$$

All of these excess returns are computed relative to riskless short-term government bonds.

The idea is that, while all long-term bonds are exposed to interest rate risk (σ^2_r), there are components of returns that are specific to government bonds (σ^2_g), swaps (σ^2_s), and corporate bonds (σ^2_c). (For simplicity, I assume that the government-, swap-, and corporate-specific factors are mutually orthogonal.) In the case of corporate bonds, this component might reflect changes in credit risk. And, for all three asset classes, we can think of these components as reflecting shifts in supply and demand for specific assets. For instance, in the case of government bonds, this might reflect idiosyncratic shifts in the demand for government bonds as in a "flight to quality" episode (e.g., Duffee (1996), Longstaff (2004), etc.) or idiosyncratic shifts in the supply of government bonds (e.g., Greenwood and Vayanos (2010) and Lou, Yan, and Zhang (2013)).

The expected excess returns on any long-term fixed income asset over short-term bonds is equal to a term premium earned by all long-term fixed income assets plus a risk premium that is specific to that asset. For instance, in the case of government bonds we have

$$\overbrace{E_t\left[rx^g_{t+1}\right]}^{\substack{\text{Excess return on LT govt.} \\ \text{bonds over ST bills}}} = \overbrace{\gamma\sigma^2_r\left(q^g_0 + q^s_0 + q^s_0 + Q_t\right)}^{\substack{\text{Term premium on all LT} \\ \text{fixed income assets}}} + \overbrace{\gamma\sigma^2_g\left(q^g_0 + \alpha Q_t\right)}^{\substack{\text{Govt-bond specific} \\ \text{risk premium}}} \tag{27}$$

Since both long-term government bonds and swaps are exposed to movements the general level of interest rates, the impact of MBS duration on the overall term premium is independent of the fraction of MBS hedged with government bonds and swaps—just as in the simpler model developed above.

However, MBS hedging flows now have an impact on the spread between long-term fixed income assets. For instance, the government-specific risk premium is high when MBS duration is high: to the extent that MBS investors hedge with government bonds, the government-specific risk premium must rise to induce arbitrageurs to hold more Treasuries. Since these hedging flows have no impact on the corporate-specific risk premium they results in tighter than normal spreads between long-term corporate and government bonds.

$$\overbrace{E_t\left[rx^c_{t+1} - rx^g_{t+1}\right]}^{\substack{\text{Excess return on corps. over} \\ \text{duration-matched govt. bonds}}} = \overbrace{\gamma\sigma^2_c q^c_0}^{\substack{\text{Corp-specific} \\ \text{risk premium}}} - \overbrace{\gamma\sigma^2_g\left(q^g_0 + \alpha Q_t\right)}^{\substack{\text{Govt-bond specific} \\ \text{risk premium}}}. \tag{28}$$

Naturally, the expected excess returns on swaps over duration-matched government bonds equals the swap-specific risk premium minus the government-bond specific premium.

$$\overbrace{E_t\left[rx^s_{t+1} - rx^g_{t+1}\right]}^{\substack{\text{Excess return on swaps over}\\ \text{duration-matched govt. bonds}}} = \overbrace{\gamma\sigma^2_s\left(q^s_0 + (1-\alpha)\,Q_t\right)}^{\substack{\text{Swap-specific}\\ \text{risk premium}}} - \overbrace{\gamma\sigma^2_g\left(q^g_0 + \alpha Q_t\right)}^{\substack{\text{Govt-bond specific}\\ \text{risk premium}}}. \tag{29}$$

If a sufficiently large volume of MBS hedging takes place in swap markets and the idiosyncratic movements in swaps are large relative to the idiosyncratic movements in Treasuries, we would expect and increase in MBS duration to raise the spread between swaps and Treasuries. Formally, if $\gamma\sigma^2_s(1-\alpha) > \gamma\sigma^2_g\alpha$, we have

$$\partial E_t\left[rx^c_{t+1} - rx^g_{t+1}\right]/\partial Q_t < 0 < \partial E_t\left[rx^S_{t+1} - rx^G_{t+1}\right]/\partial Q_t. \tag{30}$$

In summary, the analysis makes several predictions for corporate-Treasury spreads. Specifically, a high level of MBS duration should: (i) be associated with narrow current spreads between corporate bonds and duration-matched Treasuries; (ii) predict a future widening of spreads between corporate bonds and duration-matched Treasuries; and (iii) predict that the future returns on corporate bonds will underperform those on duration-matched Treasuries. Furthermore, since swaps are used to hedge MBS, each of these predictions should be reversed for swaps.

References

Amihud, Yakov, and Clifford M. Hurvich (2004), "Predictive Regressions: A Reduced-bias Estimation Method," Journal of Financial and Quantitative Analysis Vol. 39, pp. 813-841.

Baker, Malcolm and Jeremy Stein (2004), "Market Liquidity as a Sentiment Indicator," Journal of Financial Markets Vol. 7, pp. 271-299.

Baker, Malcolm, Ryan Taliaferro, and Jeffrey Wurgler (2006), "Predicting Returns with Managerial Decision Variables: Is There a Small-Sample Bias?" Journal of Finance, Vol. 61, No. 4, pp. 1711-1730.

Bates, Brandon (2010), "A Predictability Predicament," Harvard University mimeograph.

Cochrane, John H. (2008), "Comments on 'Bond Supply and Excess Bond Returns' by Robin Greenwood and Dimitri Vayanos," comments given at the conference "Beyond Liquidity: Modeling Frictions in Finance and Macroeconomics."

Hodrick, Robert J. & Prescott, Edward C (1997), "Postwar U.S. Business Cycles: An Empirical Investigation," Journal of Money, Credit and Banking,Vol. 29(1), pp. 1-16.

Kendall, M.G. (1954), "A note on bias in the estimation of autocorrelation", Biometrika, 41, pp. 403-404.

Kiefer, Nicholas M. & Vogelsang, Timothy J. (2005), "A New Asymptotic Theory For Heteroskedasticity-Autocorrelation Robust Tests," Econometric Theory, Vol. 21(06), pp 1130-1164.

Ludvigson, Sydney and Serena Ng (2010), "A Factor Analysis of Bond Risk Premia". Handbook of Empirical Economics and Finance, 2010, e.d. by Aman Uhla and David E. A. Giles, pp. 313-372. Chapman and Hall, Boca Raton, FL.

Newey, Whitney K., and Kenneth D. West (1987), "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," Econometrica, Vol. 55, 703-708.

Politis, Dimitris N. and Joseph P. Romano (1994), "The Stationary Bootstrap," Journal of the American Statistical Association Vol. 89, pp. 1303-1313.

Stambaugh, Robert F. (1999), "Predictive Regressions," Journal of Financial Economics 54, pp. 375-421.

Table A.1: Horseraces between duration and past changes in long-term forwards: Regressions of 12-month excess returns on 10-year bonds on the three measures of bond market duration and the change in 10-year forward rates over the prior 12 months, $f_t^{(10)} - f_{t-1}^{(10)}$:

$$rx_{t+1}^{(10)} = a + b \cdot DUR_t + c \cdot (f_t^{(10)} - f_{t-1}^{(10)}) + \mathbf{d}'\mathbf{x}_t + \varepsilon_{t+1}^{(10)}.$$

The regressions are estimated with monthly data, so each month I am forecasting the excess return over the following 12 months. To deal with the overlapping nature of returns, *t*-statistics are based on Newey-West (1987) standard errors allowing for serial correlation at up to 18 lags. I estimate these regressions with and without other forecasting variables. Specifically, we control for the term spread following Campbell and Shiller (1991) and the first five forward rates following Cochrane and Piazzesi (2005).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
DUR_t^{AGG}				2.688	6.985	2.688						
				[0.97]	[1.98]	[0.65]						
DUR_t^{MBS}							1.587	3.176	2.152			
							[2.04]	[3.57]	[2.42]			
$DUR\_CNTRB_t^{MBS}$										5.892	10.528	6.630
										[2.11]	[3.70]	[2.84]
$f_t^{(10)} - f_{t-1}^{(10)}$	5.611	5.051	5.283	5.373	4.232	4.986	4.995	3.480	4.359	4.607	2.952	3.973
	[5.40]	[4.11]	[4.09]	[4.62]	[2.50]	[2.96]	[4.21]	[2.27]	[2.93]	[3.59]	[1.88]	[2.64]
$y_t^{(10)} - y_t^{(1)}$		1.561			2.117			2.507			2.409	
		[2.05]			[2.31]			[2.92]			[2.93]	
$f_t^{(1)}$			-23.534			-21.878			-20.616			-18.786
			[-3.74]			[-3.57]			[-3.44]			[-3.25]
$f_t^{(2)}$			153.072			140.850			125.247			112.944
			[4.63]			[4.30]			[4.38]			[4.10]
$f_t^{(3)}$			-416.853			-384.613			-339.606			-307.903
			[-5.43]			[-4.83]			[-5.48]			[-5.30]
$f_t^{(4)}$			492.612			455.080			403.319			367.805
			[5.99]			[4.99]			[6.17]			[6.18]
$f_t^{(5)}$			-206.095			-190.179			-169.503			-154.660
			[-6.30]			[-4.92]			[-6.33]			[-6.41]
Constant	6.071	3.507	12.214	-6.162	-29.198	-0.966	0.640	-8.912	5.673	-0.479	-9.589	2.143
	[6.02]	[2.42]	[3.55]	[-0.50]	[-1.72]	[-0.05]	[0.21]	[-2.03]	[1.10]	[-0.14]	[-2.15]	[0.38]
Observations	268	268	268	268	268	268	268	268	268	268	268	268
R-squared	0.29	0.36	0.51	0.30	0.40	0.51	0.33	0.46	0.54	0.33	0.46	0.54

Table A.2: Additional controls: Regressions of 12-month excess returns on 10-year bonds on measures of MBS duration, the term spread, and additional controls:

$$rx_{t+1}^{(10)} = a + b \cdot DUR_t + c \cdot (y_t^{(10)} - y_t^{(1)}) + d \cdot x_t + \varepsilon_{t+1}^{(10)}.$$

The regressions are estimated with monthly data, so each month I am forecasting the excess return over the following 12 months. To deal with the overlapping nature of returns, *t*-statistics are based on Newey-West (1987) standard errors allowing for serial correlation at up to 18 lags. Row (1) presents the baseline specification which only controls for the term spread. Row (2) adds the 8 macroeconomic factors that Ludvigson and Ng (2010) extract from 131 macroeconomic time series. Row (3) controls from the implied interest rate volatility on 3-month into 10-year swaptions. Row (4) does the same for 2-year into 10-year swaption-implied volatility. Row (5) controls for corporate spreads (the option adjusted spread on the Barclays Corporate index). Row (6) controls for the 10-year swap spread from Bloomberg. Column (7) controls for equity option implied volatility using the *VIX*. Finally, column (8) controls for a linear time trend.

		DUR^{MBS}				DUR_CNTRB^{MBS}			
		b	$[t]$	N	R^2	b	$[t]$	N	R^2
(1)	Baseline	4.437	[6.48]	268	0.37	14.629	[5.78]	268	0.40
(2)	Ludvigson and Ng Macro Factors	3.854	[2.89]	228	0.36	14.338	[3.56]	228	0.43
(3)	Rate Vol (3-mo into 10-yr swaptions)	4.568	[7.71]	205	0.46	12.714	[6.30]	205	0.43
(4)	Rate Vol (2-yr into 10-yr swaptions)	4.559	[7.92]	205	0.46	12.725	[6.61]	205	0.44
(5)	Corporates spreads	4.662	[6.02]	263	0.38	14.659	[5.88]	263	0.41
(6)	Swap spreads	4.412	[6.56]	268	0.37	14.644	[5.91]	268	0.40
(7)	Past Stock Returns	4.681	[6.15]	268	0.39	14.786	[5.54]	268	0.41
(8)	Implied Equity Vol (VIX)	5.207	[7.23]	256	0.41	15.016	[5.61]	256	0.40
(9)	Time Trend	4.835	[6.28]	268	0.40	14.806	[6.05]	268	0.41

Table A.3: Forecasting 3-month excess nominal bond returns, 1989-Present: Regressions of 3-month excess returns on 10-year Treasuries on the effective duration of the Aggregate Index, the effective duration of the MBS Index, and the effective duration contribution of the MBS index

$$rx_{t+1}^{(10)} = a + b \cdot DUR_t + \mathbf{c}'\mathbf{x}_t + \varepsilon_{t+1}^{(n)}.$$

The regressions are estimated with monthly data, so each month I am forecasting the excess return over the following 3 months. To deal with the overlapping nature of returns, t-statistics are based on Newey-West (1987) standard errors allowing for serial correlation at up to 6 lags. I estimate these regressions with and without other forecasting variables identified in the literature on bond risk premia. Specifically, I control for the term spread following Campbell and Shiller (1991) and the first five forward rates following Cochrane and Piazzesi (2005).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
DUR_t^{AGG}			2.763	3.720	2.524						
			[1.81]	[2.44]	[1.40]						
DUR_t^{MBS}						1.285	1.707	1.290			
						[2.81]	[3.50]	[2.24]			
$DUR\_CNTRB_t^{MBS}$									4.645	5.517	4.200
									[3.19]	[3.91]	[2.78]
$y_t^{(10)} - y_t^{(1)}$	0.571			0.777			0.951			0.855	
	[1.79]			[2.29]			[2.73]			[2.49]	
$f_t^{(1)}$		-14.062			-11.771			-10.816			-8.789
		[-3.38]			[-2.66]			[-2.42]			[-1.90]
$f_t^{(2)}$		90.042			75.084			66.382			54.168
		[3.98]			[3.10]			[2.70]			[2.10]
$f_t^{(3)}$		-218.206			-184.006			-162.645			-134.287
		[-4.18]			[-3.26]			[-2.86]			[-2.23]
$f_t^{(4)}$		227.673			193.465			172.603			143.518
		[-4.03]			[3.15]			[2.81]			[2.19]
$f_t^{(5)}$		-85.188			-72.507			-65.596			-54.406
		[-3.66]			[-2.87]			[-2.61]			[-2.03]
Constant	0.517	-0.452	-11.144	-16.742	-12.233	-2.881	-5.813	-3.170	-3.594	-5.901	-5.034
	[0.80]	[-0.19]	[-1.62]	[-2.37]	[-1.31]	[-1.82]	[-2.97]	[-1.02]	[-2.12]	[-3.12]	[-1.45]
Observations	277	277	277	277	277	277	277	277	277	277	277
R-squared	0.02	0.12	0.03	0.07	0.14	0.06	0.12	0.16	0.08	0.13	0.17

Table A.4: Forecasting 1-month excess nominal bond returns, 1989-Present: Regressions of 1-month excess returns on 10-year Treasuries on the effective duration of the Aggregate Index, the effective duration of the MBS Index, and the effective duration contribution of the MBS index

$$rx_{t+1}^{(10)} = a + b \cdot DUR_t + \mathbf{c}'\mathbf{x}_t + \varepsilon_{t+1}^{(n)}.$$

t-statistics are based on heteroskedasticity robust standard errors I estimate these regressions with and without other forecasting variables identified in the literature on bond risk premia. Specifically, I control for the term spread following Campbell and Shiller (1991) and the first five forward rates following Cochrane and Piazzesi (2005).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
DUR_t^{AGG}			1.234	1.609	1.519						
			[1.66]	[2.15]	[1.94]						
DUR_t^{MBS}						0.511	0.678	0.630			
						[2.33]	[3.00]	[2.44]			
$DUR\_CNTRB_t^{MBS}$									1.914	2.263	2.123
									[2.68]	[3.11]	[2.90]
$y_t^{(10)} - y_t^{(1)}$	0.226			0.313			0.377			0.342	
	[1.71]			[2.39]			[2.83]			[2.56]	
$f_t^{(1)}$		-2.272			-0.922			-0.704			0.391
		[-0.99]			[-0.38]			[-0.30]			[0.17]
$f_t^{(2)}$		18.005			9.215			6.576			-0.115
		[1.27]			[0.62]			[0.45]			[-0.01]
$f_t^{(3)}$		-48.332			-28.411			-21.593			-5.963
		[-1.49]			[-0.83]			[-0.64]			[-0.18]
$f_t^{(4)}$		52.219			32.494			25.845			9.752
		[1.60]			[0.94]			[0.76]			[0.28]
$f_t^{(5)}$		-19.420			-12.171			-10.084			-3.894
		[-1.56]			[-0.93]			[-0.78]			[-0.29]
Constant	0.119	-0.987	-5.141	-7.347	-8.091	-1.232	-2.394	-2.323	-1.588	-2.513	-3.304
	[0.53]	[-0.65]	[-1.50]	[-2.11]	[-1.98]	[-1.57]	[-2.75]	[-1.37]	[-1.95]	[-2.80]	[-1.77]
Observations	279	279	279	279	279	279	279	279	279	279	279
R-squared	0.01	0.04	0.01	0.03	0.06	0.03	0.06	0.07	0.04	0.06	0.08

Table A.5. Time-Series Robustness Checks: Regressions of 12-month excess returns on 10-year nominal Treasuries on the effective duration of the Aggregate Index, the effective duration of the MBS Index, and the effective duration contribution of the MBS index:

$$rx_{t+1}^{\$(10)} = a + b \cdot DUR_t + \mathbf{c}'\mathbf{x}_t + \varepsilon_{t+1}^{\$(n)}.$$

The regressions are estimated with monthly data, so each month I am forecasting the excess return over the following 12 months. I estimate these regressions with and without other forecasting variables. t-statistics based on Newey-West (1987) allowing for serial correlation at up to 18 lags are shown in brackets. I report p-values based on standard asymptotics and the fixed-b asymptotics developed by Kiefer and Vogelsang (2005). We next report bootstrapped p-values using the stationary moving-blocks bootstrap of Politis and Romano (1994) using a average block length of 48 months and 10,000 bootstrap replications. I also report t-statistics based on Hansen-Hodrick (1980) standard errors using 12 lags as well as parametric standard errors which assume that the scores for 12-month returns follow an ARMA(1,12) process. I then study the impact of Stambaugh (1999) bias on the baseline results. I report a bootstrap bias-adjusted estimate using the approach of Baker and Stein (2004) and Baker, Taliaferro, and Wurgler (2006). For the univariate specifications, I also report the bias-adjusted estimates using the methods in Amihud and Hurvich (2004). Last, I decompose the bias as $BIAS(b_{OLS}) = \phi \times BIAS(\rho)$ following Amhiud and Hurvich (2004).

	DUR_t^{AGG}			DUR_t^{MBS}			$DUR\_CNTRB_t^{MBS}$		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
b_{OLS}	6.725	11.482	7.749	2.905	4.437	3.929	11.253	14.629	12.143
Controls	None	Term spread	Forwards	None	Term spread	Forwards	None	Term spread	Forwards
HAC standard errors									
[t] Newey-West	[2.84]	[5.44]	[2.73]	[2.76]	[6.48]	[4.59]	[3.70]	[5.78]	[4.44]
Standard p-value	0.005	0.000	0.007	0.006	0.000	0.000	0.000	0.000	0.000
Fixed-b p-value	0.011	0.000	0.015	0.014	0.000	0.000	0.001	0.000	0.000
Bootstrapped p-value	0.002	0.000	0.072	0.024	0.000	0.002	0.010	0.000	0.003
[t] Hansen-Hodrick	[3.23]	[7.64]	[2.89]	[2.66]	[8.16]	[5.48]	[3.71]	[6.32]	[4.56]
[t] ARMA(1,12)	[2.04]	[4.11]	[2.27]	[2.27]	[4.98]	[4.27]	[2.75]	[4.93]	[4.27]
Stambaugh bias									
[b] bootstrap bias-adjusted	6.619	11.361	7.766	2.872	4.402	3.922	11.182	14.565	12.212
$BIAS(b_{OLS})$ bootstrap	0.106	0.121	-0.017	0.033	0.036	0.007	0.071	0.064	-0.069
[b] Amihud bias-adjusted	6.609			2.873			11.188		
$BIAS(b_{OLS})$ Amihud	0.116			0.032			0.066		
ϕ	-8.433			-2.322			-4.863		
$BIAS\ (\rho)$	-0.014			-0.014			-0.014		

Table A.6: Forecasting 10-year excess nominal bond returns, Barclays vs. Bank of America measures, 1991-present: Regressions of 12-month excess returns on 10-year nominal Treasuries on the effective duration of the MBS Index and the effective duration contribution of the MBS index:

$$rx_{t+1}^{(10)} = a + b \cdot DUR_t + \mathbf{c}'\mathbf{x}_t + \varepsilon_{t+1}^{(n)}.$$

The regressions are estimated with monthly data, so each month we are forecasting the excess return over the following 12 months. To deal with the overlapping nature of returns *t*-statistics are based on Newey-West (1987) standard errors allowing for serial correlation at up to 18 lags. We estimate these regressions with and without other forecasting variables identified in the literature on bond risk premia. Specifically, we control for the term spread following Campbell and Shiller (1991) and the first five forward rates following Cochrane and Piazzesi (2005).

	Effective Duration from Barclays / Lehman						Effective Duration from Bank of America / Merrill Lynch					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
DUR_t^{MBS}	3.691	4.745	4.002				1.156	6.03	5.368			
	[3.77]	[7.12]	[4.47]				[1.26]	[6.47]	[4.60]			
$DUR\_CNTRB_t^{MBS}$				11.648	14.408	12.025				4.961	18.424	15.108
				[3.78]	[5.79]	[4.29]				[1.35]	[4.54]	[3.57]
$y_t^{(10)} - y_t^{(1)}$		2.943			2.876			4.874			4.590	
		[4.24]			[4.00]			[6.21]			[5.26]	
$f_t^{(1)}$			-22.854			-18.268			-32.239			-25.642
			[-2.95]			[2.53]			[-5.49]			[-4.77]
$f_t^{(2)}$			127.282			99.059			172.254			137.310
			[3.39]			[2.76]			[5.24]			[4.22]
$f_t^{(3)}$			-312.841			-247.020			-408.114			-332.623
			[-3.81]			[3.06]			[-5.24]			[-4.20]
$f_t^{(4)}$			340.494			273.582			428.029			355.822
			[3.93]			[3.14]			[5.11]			[4.09]
$f_t^{(5)}$			-132.414			-106.813			-160.274			-134.003
			[-3.81]			[3.01]			[-4.76]			[-3.73]
Constant	-6.738	-15.156	-8.401	-7.192	-15.020	-13.161	0.806	-26.103	-18.752	-0.664	-24.436	-22.249
	[-1.71]	[-4.69]	[-1.54]	[-1.85]	[-4.02]	[-2.14]	[0.20]	[-5.35]	[-2.90]	[-0.13]	[-3.98]	[-2.47]
Observations	243	243	243	243	243	243	243	243	243	243	243	243
R-squared	0.17	0.38	0.44	0.19	0.39	0.47	0.02	0.35	0.44	0.03	0.36	0.46