Interest Rate Conundrums in the 21st Century*

SAMUEL G. HANSON
Harvard Business School

DAVID O. LUCCA
Federal Reserve Bank of New York

JONATHAN H. WRIGHT
Johns Hopkins University

February 27, 2018

Abstract

A large literature argues that long-term interest rates appear to react far more to high-frequency (e.g., daily or monthly) movements in short-term interest rates than predicted by the standard expectations hypothesis. We find that, since 2000, such high-frequency "excess sensitivity" remains evident in U.S. data and has grown even stronger. By contrast, the positive association between low-frequency changes (e.g., at a 6-month or 12-month horizon) in short- and long-term interest rates, which was quite strong before 2000, has weakened substantially in recent years. As a result, "conundrums"—defined as 6- or 12-month periods in which short rates and long rates move in opposite directions—have become far more common since 2000. We argue that the puzzling combination of high-frequency excess sensitivity and low-frequency decoupling between short- and long-term rates can be understood using a model in which (i) shocks to short-term interest rates lead to a rise in term premia on long-term bonds and (ii) arbitrage capital moves slowly over time. We discuss the implications of our findings for interest rate predictability, the transmission of monetary policy, and the validity of high-frequency event-study approaches for assessing the impact of monetary policy.

*We thank Richard Crump, Thomas Eisenbach, Robin Greenwood, Jeremy Stein, and Adi Sunderam for useful feedback. Hanson gratefully acknowledges funding from the Division of Research at Harvard Business School. All errors are our sole responsibility. The views expressed here are the authors' and are not representative of the views of the Federal Reserve Bank of New York or of the Federal Reserve System.
Emails: shanson@hbs.edu; david.lucca@ny.frb.org; wrightj@jhu.edu.

1 Introduction

The sensitivity of long-term interest rates to movements in short-term interest rates is a central feature of the term structure and plays a key role in the transmission of monetary policy to the real economy. Despite its importance, there is considerable disagreement about the extent to which long-term rates respond to short rates. For example, this basic disagreement features prominently in the ongoing debate about the extent to which easy monetary policy contributed to the 2003–2007 boom in house prices (Bernanke, 2010; Taylor, 2010).

In this paper, we document an important and previously unrecognized fact about the term structure of interest rates: the positive association between low- frequency changes (say, at a 6-month or 12-month horizon) in short- and long- term interest rates was quite strong before 2000, but has weakened substantially in recent years. In contrast, the association between high- frequency changes (say, at a daily or 1-month horizon) in short- and long-term interest rates has strengthened.[1]

A stark example of such low-frequency decoupling is the period after June 2004 when the Federal Reserve raised its short-term policy rate, but longer-term yields and forward rates actually declined. This was famously described by then Fed Chairman Alan Greenspan as a "conundrum," and has been discussed in many papers, including Backus and Wright (2007). One simple way to summarize our key finding is that we show that this 2004 episode was by no means unique, and that "conundrums"—defined as 6- or 12-month periods where short and long-term rates move in opposite directions—have become far more commonplace since 2000. For instance, from 1971 to 1999, 1- and 10-year nominal yields moved in the same direction in 83% of all 6-month periods. By contrast, since 2000, the corresponding figure has only been 61%.

[1]Throughout the paper, we take the short-rate to be the 1-year nominal Treasury yield rather than the overnight federal funds rate targeted by the Federal Reserve.

We construct a simple model to explain these facts. In this model, risk-averse investors can either invest in short- or long-term default-free nominal bonds. Monetary policy pins down the rate on short-term nominal bonds (i.e., short-term bonds are available in perfectly elastic supply), but long-term nominal bonds are available in a supply that varies stochastically over time. Since shocks to the supply of long-term bonds must be absorbed by risk-averse investors in equilibrium, supply shocks affect term premia on long-term bonds as in Vayanos and Vila (2009) and Greenwood and Vayanos (2014).

In the pre-2000 period, there is a large persistent component of short-term nominal interest rates. A natural interpretation is that this persistent component reflects shocks to trend inflation as in Stock and Watson (2007). The existence of this highly persistent component, in turn, readily explains the sensitivity of long-term interest rates to short-term interest rates both at low and high frequencies. In the post-2000 period, the volatility of the persistent component of short-term nominal interest rates dropped sharply. From a textbook expectations hypothesis perspective, this should have reduced the sensitivity of long rates to short rates. In the data however, this occurs, but only at low frequencies and we see even greater sensitivity at high frequencies.

We develop a simple model that explains how such frequency-specific decoupling can arise. Our explanation rests on two key ingredients: term premium shocks and slow moving capital. The first key assumption is term premium shocks. In the post-2000 period, we assume that shocks to the net supply of long- term bonds are positively correlated with shocks to short-term interest rates. This implies that increases in short-term rates are associated with increases in the term premium component of long-term rates, generating excess sensitivity at high frequencies. We discuss several distinct amplification mechanisms that have been highlighted in the recent literature that can rationalize this reduced-form assumption, including investors who have a tendency to "reach for yield" in response to declines in short rates, mortgage convexity hedging flows, and over-extrapolative investors.

The second key assumption is that capital is slow-moving as in Duffie (2010): the previously mentioned supply shocks encounter a short-run demand curve that is a good deal steeper than the long-run demand curve. This slow-moving capital dynamic implies that excess sensitivity is greatest when measured at short horizons, enabling our model to match the key stylized facts we have emphasized.

Thus our model explains the shift that occurred in 2000 as arising due to a combination of a decline in the volatility of the persistent component of short rates (say, because long-run inflation expectations have become better anchored) and an increase in the importance of the kinds of supply-and-demand-based amplification mechanisms noted above (say, because the mortgage market has become a larger part of the overall market for long-term bonds).

Our stylized fact has implications for the transmission of monetary policy. Stein (2013) points out that the excess sensitivity of long-term yields—whereby shocks to short rates move term premia in the same direction—should strengthen the real effects of monetary policy relative to a world where the expectation hypothesis holds. He refers to this as the "recruitment channel" of monetary policy transmission. Although we also find excess sensitivity at high-frequencies, our findings suggest that the recruitment channel may not be nearly as strong as Stein (2013) suggests because the resulting shifts in term premia tend to be more transitory—and, thus, likely to have only limited effects on aggregate demand—than one would infer from high-frequency changes. While there may nonetheless be a meaningful recruitment channel—i.e., monetary policy may have an important impact on financial conditions, our results caution against the practice of inferring the strength of this channel from high-frequency changes.

More generally, our stylized fact has implications for how one should interpret event-study evidence based on high-frequency changes. Economic news comes out in a lumpy manner, and the change in long-term interest rates around a particular piece of news is often used as a convenient measure of the long-run impact of that shock. However, if the impact on long-term rates tends to quickly wear off, then the shock's short-term and long-term impact will be quite different; and the event study approach will necessarily capture the short-term impact. As an example, in the week ending March 3, 2017, several Federal Reserve officials made comments that were widely interpreted as signaling a high likelihood of a rate hike at the upcoming FOMC meeting. Not surprisingly, short-term interest rates rose on the news. But in that week, 10-year yields and even 30-year forward rates rose over 15 basis points. It seems extremely unlikely that this news raised expected short rates decades into the future by 15 basis points. The interpretation that would be most consistent with the stylized fact documented in this paper is that, in addition to raising short rates, the news gave a *temporary* boost to term premia. In that case, the event-study methodology would give a misleading estimate of the longer-run impact of this news on long-term bond yields.

Related literature Gürkaynak et al. (2005) argue that the high sensitivity of daily changes in long-term *nominal* rates to daily changes in short-term *nominal* rates may reflect the fact that long-run inflation expectations are largely unmoored and are being continuously revised in light of incoming news. In other words, Gürkaynak et al. (2005) argue that the strong sensitivity of long-term rates to short-term rates may work through a simple expectations hypothesis channel: the puzzle can be resolved once one allows for very persistent shocks to inflation expectations. However, as shown by Beechey and Wright (2009), Abrahams et al. (2016) and Hanson and Stein (2015), in the post-2000 period, the strong sensitivity of long-term nominal rates primarily reflect the sensitivity of long-term *real* rates to short-term nominal rates and not the sensitivity of long-term break-even inflation. To the extent that one doubts that expected future real rates at distant horizons fluctuate meaningfully from day to day, this finding casts doubt on an expectations hypothesis explanation for the puzzling degree of sensitivity, at least in the post-2000 sample.

Hanson and Stein (2015) argue that excess sensitivity works through term premia on long-term bonds: shocks to short rates temporarily move term premia in the same direction. Specifically, they argue that the excess sensitivity of long-interest rates is due to shifts in the demand for long-term bonds from "yield-oriented" investors who target a certain level of yields in their investment portfolios and extend the maturity of their investments as rates decline. However, a variety of distinct supply and demand mechanisms may contribute to excess sensitivity. For instance, Hanson (2014) and Malkhozov et al. (2016) argue that negative shocks to short-term rates induce mortgage refinancing waves that lead to temporary decline in the duration of outstanding fixed-rate mortgages —i.e., a temporary reduction in the effective supply of long-term bonds. Relatedly, Domanski et al. (2017) and Shin (2017) point to an amplification mechanism where insurers and pension funds need to match the duration of their assets and liabilities and consequently demand more long duration assets following a reduction in the level of interest rates. Our model shows that these kinds of supply and demand mechanisms imply that long-term interest rates tend to "over-react" to movements in short rates at high frequencies, but that this over-reaction is far less pronounced at the lower frequencies that should be of greatest interest to monetary policymakers.

The plan for the remainder of the paper is as follows. In Section 2, we document our basic stylized fact. In Section 3, we construct a simple reduced-form time-series model of interest rates that can match this stylized fact: the key takeaway here is that lagged changes in the level of the term structure are negatively associated with future changes in the slope. Section 4 develops the simple model that we use to explain and interpret our key stylized fact. Section 5 discusses implications for trading strategies and affine term structure models. Finally, Section 6 concludes.

2 Main finding: Yield decoupling at low frequencies

This section presents our main empirical finding: the association between low-frequency (6 months or longer) changes in short- and long-term interest rates was quite strong before 2000, but has weakened substantially in recent years; by contrast, the association between high-frequency (1 month or shorter) changes in short- and long- term interest rates has actually strengthened in recent years. We begin by documenting this basic fact for the U.S. We then repeat this same exercise for Canada, Germany, and the U.K. and show that the patterns in those countries are broadly in line with what we find in the U.S.

2.1 Yield decoupling in the U.S.

We obtain historical data on the nominal and real Treasury yield curve from Gürkaynak et al. (2007); Gürkaynak et al. (2010). We focus on continuously compounded 10-year zero-coupon yields and 10-year instantaneous forward rates. We also decompose nominal yields into real yields and inflation compensation, defined as the difference between nominal and real yields derived from Treasury Inflation-Protected Securities (TIPS). Our sample begins in 1971, which is when reliable data on 10-year nominal yields first become available, and ends in 2017. As discussed in more detail below, we focus on the pre- and post-2000 samples. For real yields and inflation compensation, we only study the post-2000 sample, since data on TIPS is only available beginning in 1999. All data are measured as of the end of the relevant period—e.g., the last trading day of each month.

In standard monetary models, the central bank sets overnight nominal interest rates, and other rates move in response to expectations of the path of overnight rates. A large literature argues that central banks in the U.S. and abroad have increasingly relied on communication—explicit signaling about the future path of overnight rates—as an active policy instrument (Gurkaynak et al., 2004; Lucca and Trebbi, 2009). To capture news about the near-term path of monetary policy that would not impact the current overnight rate, we follow the recent literature (Campbell et al., 2012; Gertler and Karadi, 2015; Hanson and Stein, 2015) and take the short rate to be the 1-year nominal Treasury rate. Using 1-year rates also limits any potential distortions stemming from the 2008–2015 period when overnight nominal rates were stuck at the zero lower bound in the U.S. By contrast, 1-year nominal yields continued to fluctuate over the 2008–2015 period (Swanson and Williams, 2014).

To illustrate our key stylized fact, we begin by regressing changes in 10-year yields or forward rates on changes in 1-year nominal yields. Specifically, we estimate regressions of the form:

$$y_{t+h}^{(10)} - y_t^{(10)} = \alpha_h + \beta_h (y_{t+h}^{(1)} - y_t^{(1)}) + \varepsilon_{t,t+h} \tag{2.1}$$

and

$$f_{t+h}^{(10)} - f_t^{(10)} = \alpha_h + \beta_h (y_{t+h}^{(1)} - y_t^{(1)}) + \varepsilon_{t,t+h}, \tag{2.2}$$

where $y_t^{(n)}$ is the n-year zero-coupon rate on day t and $f_t^{(n)}$ is the n-year-ahead instantaneous forward rate. Panel A in Table 1 reports estimated slope coefficients in (2.1) for zero-coupon nominal yields, real yields, and inflation compensation using daily data and using monthly data for $h = 1, 3, 6, 12$—i.e., we report coefficients for daily, monthly, quarterly, semi-annual, and annual changes in yields. The results are shown for the pre-2000 and post-2000 subsamples separately. We base this sample split on a number of break-date tests that we will discuss shortly. Panel B reports the corresponding slope coefficients in (2.2) using changes in instantaneous forwards as the dependent variable.

At a daily frequency, there has been a meaningful increase in the regression coefficients between the pre-2000 and post-2000 subsamples. The increasing sensitivity of long-term rates at high frequencies is our first novel finding. Specifically, the coefficient on the 10-year zero yield from Panel A has increased from $\beta_h = 0.56$ in the pre-2000 subsample to $\beta_h = 0.85$ in the post-2000 subsample. Similarly, from Panel B, the coefficient for daily changes in 10-year forward rates is $\beta_h = 0.39$ in the pre-2000 subsample and $\beta_h = 0.48$ in the post-2000 subsample.

The large magnitude of β_h at a daily frequency is a long-standing puzzle in the macroeconomics and finance literature. Short-term nominal interest rates, are pinned down by current monetary policy and the near-term expect path of policy. However, shocks to monetary policy and macroeconomic data are empirically short-lived. Because of the transitory nature of these shocks, long-term rate should not be highly responsive to changes in short-term rates if the expectations hypothesis holds. However, Gürkaynak et al. (2005) note that the strong sensitivity of long-term nominal rates could be consistent with the expecations hypothesis if long-run inflation expectations are unanchored and are continuously being updated in light of incoming news. But Beechey and Wright (2009) and Hanson and Stein (2015) find that much of the sensitivity of long-term nominal rates is due to the response of long-term real rates rather than inflation compensation, suggesting that movements in term premia are an important part of the explanation. Consistent with these studies, we find that the majority of the response of 10-year nominal instantaneous forwards can be accounted for by the response of instantaneous real forwards in the post-2000 sample (Panel B).

Looking down the rows, Table 1 shows a second new fact that has not previously been documented. The regression coefficients at lower frequencies are much lower in the post-2000 sample than in the pre-2000 sample. For example, the coefficient for yearly changes in 10-year yields is $\beta_h = 0.56$ in the pre-2000 sample but only $\beta_h = 0.18$ in the post-2000 sample. Similarly, for 10-year forward rates, the coefficient at a yearly frequency is $\beta_h = 0.39$ in the pre-2000 sample but $\beta_h = -0.19$ in the post-2000 sample. More generally, in the post-2000 sample, the coefficient β_h is a steeply declining function of the horizon h over which yield changes are calculated. By contrast, β_h is a relatively constant function of horizon h in the pre-2000 sample. In terms of a decomposition between real yields and inflation compensation, we see that the majority of the decline in β_h as a function of h during the post-2000 sample is accounted for by the real yield component of the 10-year zero rate, and is equally split between real and inflation compensation components of the 10-year forward.

In summary, Panels A and B of Table 1 show that, prior to 2000, there was strong tendency for short- and long-term interest rates to rise and fall together at both high- and low-frequencies. While the high-frequency relationship has persisted and has even grown stronger since 2000, at lower frequencies, there is little relationship between movements in short- and long-term rates in the recent period. In other words, the excess sensitivity puzzle appears to be only present at high-frequencies but not at lower-frequencies in recent years. Put differently, events such as "Greenspan's conundrum" (as discussed by Backus and Wright, 2007, for example)—the period following June 2004 when the Federal Reserve raised its short-term policy rate and longer-term yields and forward rates declined—have grown increasingly common since 2000. Indeed, since 2000, 1- and 10-year yields have moved in the same direction in 61% of all 6-month periods. By contrast, from 1971 to 1999, the corresponding figure was 83%, and the difference is statistically significant ($t = 3.91$).

We use two other approaches to document our key stylized fact. The first is to estimate equations (2.1) and (2.2) using ten-year rolling windows. The estimated slope coefficients for yearly changes ($h = 12$ with monthly data) are shown in Figure 2 for 10-year yields and forward rates. The coefficient declines substantially in more recent windows. The second approach is to test for a structural break in equations (2.1) and (2.2), allowing for a break date that is not known *a priori*. We use the test of Andrews (1993) who conducts a Chow test (Chow, 1960) at all possible break dates, and then takes the maximum of these test statistics. Figure 3 plots the Wald test statistic for each possible break date in equations (2.1) and (2.2). The strongest evidence for a break is in 2000 or 2001 in both equations (2.1) and (2.2). Figure 3 shows the Andrews (1993) critical values for a null of no structural break and the break is clearly statistically significant.[2]

Before turning to the international evidence, it is important to note that the break around 2000 in Figure 3 is not a result of the period from December 2008 to December 2015 where the federal funds rate was stuck at the zero lower bound. Even if the full sample period is ended in 2008, we still detect a break around 2000. More specific to the results in Table 1, if the post-2000 sample ends in December 2008, we find a daily $\beta_h = 0.77$ and a yearly $\beta_h = 0.20$, which are essentially indistinguishable from the numbers in Table 1.

2.2 International evidence

Our primary focus is on the U.S., but to better understand plausible economic mechanisms it is useful to consider whether this low-frequency yield decoupling is observed in other large, highly-developed economies. In this subsection, we briefly explore international evidence on low-frequency yield-curve decoupling for the U.K., Germany, and Canada. We obtain yield

[2]In this paper, we are comparing pre-2000 and post-2000 data based on the estimated slope coefficients in equations (2.1) and (2.2) for yearly changes. There may be breaks in the slope coefficients for higher frequency changes at other dates, and for example, Thornton (forthcoming) argues that there is a break in the relationship between monthly changes in ten-year yields and monthly changes in the federal funds rate somewhat earlier in the sample.

curve data from each country's central bank website.

Table 2 reports estimates of equation (2.1) for the U.K. in Panel A as well as Germany and Canada in Panel B. For the U.K., the estimates are broken out into real yields and inflation compensation. Monthly data for Germany is available beginning in 1972 and daily data for Germany is available starting in 2000. For the U.K. and Canada, data is available at a both a daily and month frequency beginning in 1985 or 1986.

The evidence for the U.K. is remarkably similar to the U.S. evidence in Table 1. Before 2000, the daily coefficient ($\beta_h = .44$) and the yearly coefficient ($\beta_h = .38$) are very similar in the U.K. After 2000, the daily sensitivity increases ($\beta_h = .80$), and the yearly sensitivity drops ($\beta_h = .29$). Because we have data on real yields prior to 2000 in the U.K., we can decompose the change in β_h into its real and inflation compensation components. As shown in Table 2, the inflation compensation component of β_h is quite stable across sample periods and frequency h. Thus, most of the changes in β_h are accounted for by changes in the real yield components.

Moving to the bottom panel, we again observe very similar patterns to those in the U.S. and U.K. for Germany and Canada. In the pre-2000 sample, β_h is stable across frequencies. Post-2000 we observe a higher sensitivity at higher frequency and a decay in β_h at lower frequency.

3 Lead-lag relationships in the yield curve

In this section, we pinpoint the time-series properties of the term structure of interest rates that are needed to account for the low-frequency decoupling between short- and long-term rates documented above. In examining term structure dynamics, it is useful and customary to study the dynamics of principal components, especially level, slope, and curvature (Litterman and Scheinkman, 1991). Defining level as the 1-year zero coupon yield ($L_t \equiv y_t^{(1)}$), the slope as the 10-year yield less the 1-year yield ($S_t \equiv y_t^{(10)} - y_t^{(1)}$), and the curvature as the 5.5-year yield less the average of the 1- and 10-year yields ($C_t = y_t^{(5.5)} - (y_t^{(1)} + y_t^{(10)})/2$), the puzzle described in this paper can be restated as the observation that, since 2000, changes in level and slope have become *negatively* associated at low frequencies, but not at high frequencies. Specifically, the slope coefficient in equation (2.1) can be rewritten as:

$$\beta_h = 1 + \frac{\Sigma_{j=-h+1}^{h-1}(h-|j|)Corr(\Delta L_t, \Delta S_{t+j})}{\Sigma_{j=-h+1}^{h-1}(h-|j|)Corr(\Delta L_t, \Delta L_{t+j})}\sqrt{\frac{Var(\Delta S_t)}{Var(\Delta L_t)}}, \tag{3.1}$$

and so the decline in β_h at low, but not high, frequencies must logically mean that (i) there was a shift in the autocorrelation of ΔL_t, (ii) the cross-correlation between changes in level and future changes in slope declined, or (iii) the cross-correlation between changes in slope and future changes in levels has fallen. It turns out that both of these cross-correlations have declined, although (ii) plays a more important role than (iii) in driving the decline in β_h at low frequencies. This can be seen from Figure 4, which plots the cross correlation between the slope (difference between 10- and 1-year Treasury yields) and the level (1-year yield) of the yield curve. The contemporaneous correlation is less negative post-2000, meaning that yields are more likely to move in lockstep at a 1-month frequency. In contrast, cross-correlations between changes in level and slope are consistently negative post-2000, resulting in lower values of the β_h coefficient at lower frequencies.

3.1 Predicting level, slope, and curvature

As another way of looking at the evolution of yield curve dynamics, we consider predictive regressions for the level, slope, and curvature of the yield curve. Most term structure models are Markovian with respect to the filtration given by current state variables, meaning that the conditional mean of future yields depends only on today's state variables. However, our key finding—that the correlation between changes in short- and long-term rates has declined at low-frequencies, even though the two remain highly correlated at a daily frequency—suggests that it may be useful to include additional lags when forecasting yields. We therefore consider the following system of predictive regressions:

$$L_{t+1} = \delta_{0L} + \delta_{1L}L_t + \delta_{2L}S_t + \delta_{3L}C_t$$
$$+\delta_{4L}(L_t - L_{t-h}) + \delta_{5L}(S_t - S_{t-h}) + \delta_{6L}(C_t - C_{t-h}) + \varepsilon_{L,t+1} \tag{3.2}$$
$$S_{t+1} = \delta_{0S} + \delta_{1S}L_t + \delta_{2S}S_t + \delta_{3S}C_t$$
$$+\delta_{4S}(L_t - L_{t-h}) + \delta_{5S}(S_t - S_{t-h}) + \delta_{6S}(C_t - C_{t-h}) + \varepsilon_{S,t+1} \tag{3.3}$$
$$C_{t+1} = \delta_{0C} + \delta_{1C}L_t + \delta_{2C}S_t + \delta_{3C}C_t$$
$$+\delta_{4C}(L_t - L_{t-h}) + \delta_{5C}(S_t - S_{t-h}) + \delta_{6C}(C_t - C_{t-h}) + \varepsilon_{C,t+1}, \tag{3.4}$$

where L_t, S_t and C_t denote the level, slope, and curvature at the end of month t, respectively, as defined above. All of these predictive regressions include lags of level, slope, and curvature, as is standard, but also lagged changes in level, slope, and curvature. Several other authors have considered the use of lags in term structure models, including Duffee (2013), Feunou and Fontaine (2014), Cochrane and Piazzesi (2005), Joslin et al. (2013) and Monfort and Pegoraro (2013), although none of these have discussed the relationship to yield curve conundrums.

Table 3 reports estimates of various restricted forms of equations (3.2), (3.3), and (3.4) for $h = 12$ and for both the pre-2000 and post-2000 subsamples. All specifications include the first lag of level, slope, and curvature. We include specifications omitting all lagged changes ($\delta_4 = \delta_5 = \delta_6 = 0$), omitting lagged changes in slope and curvature ($\delta_5 = \delta_6 = 0$), omitting lagged changes in curvature ($\delta_6 = 0$) and including all predictors. Based on the AIC or BIC, the model with one lag of level, slope, and curvature and lagged changes in level is selected in the post-2000 subsample, while no lagged changes are needed in the pre-2000 subsample. In the post-2000 subsample, the lagged change in level is a highly significant negative predictor of the slope. Increases in the level of yields predict subsequent yield curve flattening, much more so than in the past.

The model in equations (3.2)-(3.4) can match the key stylized fact we documented above. These equations can be written jointly as a restricted vector autoregression in $\mathbf{y}_t = (L_t, S_t, C_t)'$ of the form:

$$\mathbf{y}_{t+1} = \mu + \mathbf{A}_1 \mathbf{y}_t + \mathbf{A}_2 \mathbf{y}_{t-h} + \varepsilon_{t+1}. \tag{3.5}$$

Let $\Gamma_{ij}(k)$ denote the ijth element of the autocovariance of $\mathbf{y}_t$ at a lag of k models—i.e., the the ijth element of $\mathbf{\Gamma}(k) = E[(\mathbf{y}_t - E[\mathbf{y}_t])(\mathbf{y}_{t-k} - E[\mathbf{y}_{t-k}])']$. Given the estimated parameters from equations (3.2) and (3.3), we can work out $\Gamma_{ij}(k)$ and hence obtain the model-implied values of β_h in equation (2.1) as:

$$\begin{aligned} \beta_h &= \frac{Cov(L_t - L_{t-h}, S_t - S_{t-h}) + Var(L_t - L_{t-h})}{Var(L_t - L_{t-h})} \\ &= 1 + \frac{2\Gamma_{12}(0) - \Gamma_{12}(h) - \Gamma_{12}(-h)}{2(\Gamma_{11}(0) - \Gamma_{11}(h))}. \end{aligned} \tag{3.6}$$

At an annual frequency, Table 1 reported estimated values of β_h in equation (2.1) of 0.56 and 0.18 for the pre-2000 and post-2000 subsamples, respectively. The last row of Table 3 includes the model-implied values of β_h from equations (3.2)-(3.4) for $h = 12$. We can see that, as a matter of statistical description, the model in equations (3.2)-(3.4) can match the basic stylized fact that we document in this paper, as long as the VAR includes lagged changes in level. If the VAR did not include lagged changes in level ($\delta_{4.} = \delta_{5.} = \delta_{6.} = 0$), the model-implied values of β_h would be 0.51 in both the pre-2000 and post-2000 subsamples, and so would be nowhere near what we observe in the data.

3.2 Specific episodes

The fact that lagged changes in the level of the yield curve are increasingly useful for predicting future changes in slope also helps to explain some important bond market episodes in recent years. We estimated equation (3.3) over the post-2000 sample with $h = 12$, and then re-estimated the equation restricting the coefficients on lagged changes to zero ($\delta_{4S} = \delta_{5S} = \delta_{6S} = 0$). We then simulated the path of the 10-year yield for one year starting in May 2004 holding the level and curvature at their actual values, and using the residuals from the unrestricted estimation of equation (3.3), but setting the parameters to their estimated values in the *restricted* regression. The top panel of Figure 4 plots the actual value of 1-year and 10-year yields over this period, and the values of the 10-year yield under this alternative scenario. This was the original "conundrum" period, as 10-year yields fell while short rates rose. However, had the slope not responded to lagged changes in the term structure, we can see that 10-year yields would actually have risen.

The bottom panel of Figure 4 performs a exercise, simulating the path of the 10-year yield for one year starting in December 2007 under the alternative scenario that the slope does not respond to lagged changes. This was a period when short rates fell sharply, but when 10-year yields actually rose, in something of a conundrum in reverse. We can see that had the slope not responded to lagged changes, 10-year yields would have fallen sharply.

4 Model

In this section, we construct a simple model that is useful for explaining the key stylized facts we have documented in this paper: short-term and long-term yields no longer move strongly together at low frequencies, but at high frequencies they move together even more strongly than in the past.

In our model, time is discrete and infinite. A set of risk-averse investors can either hold long-term, perpetual nominal bonds or short-term nominal bonds, both of which are default-free. Short-term nominal bonds are available in perfectly elastic supply and the nominal interest rate on short-term bonds from t to $t+1$, denoted i_t, follows an exogenous stochastic process. Nominal long-term bonds are available in a given net supply that must be absorbed by the investors in our model. In the model, this net supply s_t varies over time and also follows an exogenous stochastic process.

As in Vayanos and Vila (2009) and Greenwood and Vayanos (2014), shifts in the net supply of long-term bonds impact the term premium on long bonds. The first key assumption in our model is that shocks to the net supply of long-term bonds are positively correlated with shocks to short-term interest rates. This assumption, which we discuss in detail below, implies that increases in short-term rates are associated with increases in term premia, generating "excess sensitivity" of long rates to movements in short rates beyond the sensitivity implied by the expectations hypothesis.

The second key assumption, following Duffie (2010), is that capital is slow-moving, so these supply shocks walk down a short-run demand curve that is a good deal steeper than the long-run demand curve. This slow-moving capital dynamic implies that excess sensitivity is greatest when measured at short horizons, enabling our model to match the key stylized fact we have emphasized throughout. Formally, our model is a close cousin of the model in Greenwood et al. (2016), who incorporate slow-moving capital effects into a model of the term structure.

As we detail below, our model can match our key stylized fact—i.e., that the fact that β_h has fallen post-2000 for large h at the same time it has risen for small h—if (i) shocks to short-term nominal rates have become slightly less persistent and (ii) the kinds of supply-and-demand-based amplification mechanisms that we emphasize have grown in importance. We argue that (i) is justified since there is strong evidence that shocks to the persistent component of nominal inflation have become far less volatile since the mid-1990s (Stock and Watson (2007)). Similarly, we argue that (ii) is plausible since many of these amplification mechanisms appear to have become more powerful in recent years.

4.1 Model setting

4.1.1 Long-term nominal bonds

The long-term nominal bond is a perpetuity that pays a nominal coupon of K each period. Let $P_{L,t}$ denote the nominal price of this long-term bond at time t, so the nominal return on long-term bonds from t to $t+1$ is:

$$1 + R_{L,t+1} = \frac{P_{L,t+1} + K}{P_{L,t}}.$$

To generate a tractable linear model, we use a Campbell and Shiller (1988) log-linear approximation to the return on this perpetuity. Specifically, defining $\theta \equiv 1/(1+K) < 1$, the 1-period log return on the perpetuity is

$$r_{L,t+1} \equiv \ln\left(1 + R^L_{t+1}\right) \approx \overbrace{\frac{1}{1-\theta}}^{D} y_t - \overbrace{\frac{\theta}{1-\theta}}^{D-1} y_{t+1}, \tag{4.1}$$

where y_t is the log yield-to-maturity at time t and $D = 1/(1-\theta) = (K+1)/K$ is the Macaulay duration when the bond is trading at par.[3]

Let i_t denote the interest rate on short-term nominal bonds from t to $t+1$ and let $rx_{t+1} \equiv r_{L,t+1} - i_t$ denote the excess return on long-term nominal bonds over short-term nominal bonds from t to $t+1$. Then, iterating equation (4.1) forward and taking expectations,

[3]This log-linear approximation for default-free coupon-bearing bonds appears in Chapter 10 of Campbell et al. (1996).

the yield on long-term nominal bonds is given by:

$$y_t = (1-\theta)\sum_{j=0}^{\infty}\theta^j E_t\left[i_{t+j} + rx_{t+j+1}\right]. \tag{4.2}$$

Naturally, the long-term yield is the sum of (i) an expectations hypothesis component $(1-\theta)\sum_{j=0}^{\infty}\theta^j E_t\left[i_{t+j}\right]$ that reflects expected future nominal short rates and (ii) a term premium component $(1-\theta)\sum_{j=0}^{\infty}\theta^j E_t\left[rx_{t+j+1}\right]$ that reflects expected future excess returns on long-term nominal bonds over short-term bonds.

4.1.2 Market participants

There are two types of risk-averse investors in the model, each with identical risk tolerance τ, who differ solely in the frequency with which they can rebalance their bond portfolios.

The first group of investors are fast-moving investors who are free to adjust their holdings of the long-bond and the riskless short-term bond each period. Fast-moving investors are present in mass q and we denote their demand for long-term bonds at time t by b_t. Fast-moving investors have mean-variance preferences over 1-period portfolio log returns. Thus, their demand for long-term bonds at time t is given by

$$b_t = \tau\frac{E_t\left[rx_{t+1}\right]}{Var_t\left[rx_{t+1}\right]}, \tag{4.3}$$

where

$$rx_{t+1} \equiv r_{L,t+1} - i_t = \frac{1}{1-\theta}y_t - \frac{\theta}{1-\theta}y_{t+1} - i_t$$

is the excess return on long-term bonds from t to $t+1$.

The second group of investors is a set of slow-moving investors who can only adjust their holdings of long-term and short-term bonds every k periods. Slow-moving investors are present in mass $1-q$. A fraction $1/k$ of these slow-moving investors is active each period and can reallocate their portfolios. However, they must then maintain this same portfolio allocation for the next k periods. As in Duffie (2010), this is a reduced form way to model the frictions that limit the speed of capital flows. Since they only rebalance their portfolios every k periods, slow-moving investors have mean-variance preferences over their k-period *cumulative* portfolio excess return. Thus, the demand for long-term bonds from the subset of slow-moving investors who are active at time t is given by

$$d_t = \tau \frac{E_t\left[rx_{t\to t+k}\right]}{Var_t[rx_{t\to t+k}]},$$

where

$$rx_{t\to t+k} = \textstyle\sum_{j=1}^{k} rx_{t+j} = \sum_{j=0}^{k-1}\left(y_{t+j} - i_{t+j}\right) - \frac{\theta}{1-\theta}\left(y_{t+k} - y_t\right)$$

is the k-period cumulative excess return on long-term bonds from t to $t+k$.

4.1.3 Risk factors

Investors in long-term bonds face two different types of risk. First, they are exposed to *interest rate risk:* they will suffer a capital loss on their long-term bond holdings if short-term rates unexpectedly rise. Second, they are exposed to *supply risk*: there are shocks to the supply of long-term bonds that impact long-term bond yields, holding fixed the expected future path of short-term interest rates—i.e., these supply shocks impact the term premium on long-term bonds.

We make the following concrete assumptions about the evolution these two risk factors:

Short-term nominal interest rates: Short-term nominal bonds are available in perfectly elastic supply. At time t, investors learn that short-term bonds will earn a log riskless return of i_t in nominal terms between time t and $t+1$. One can think of the short-term nominal rate as being determined outside the model either by monetary policy or by a stochastic short-term storage technology that is available in perfectly elastic supply.

Crucially, we assume that the short-term nominal interest rate is the sum of a highly persistent component $i_{P,t}$ and a more transient component $i_{T,t}$:

$$i_t = i_{P,t} + i_{T,t}. \tag{4.4}$$

We assume that the persistent component $i_{P,t}$ follows an exogenous AR(1) process

$$i_{P,t+1} = \bar{\imath} + \rho_P \left(i_{P,t} - \bar{\imath}\right) + \varepsilon_{P,t+1}, \tag{4.5}$$

where $0 < \rho_P < 1$ and $Var_t\left[\varepsilon_{P,t+1}\right] = \sigma_p^2$. Similarly, we assume that the transient component $i_{T,t}$ follows an exogenous AR(1) process

$$i_{T,t+1} = \rho_T i_{T,t} + \varepsilon_{T,t+1}, \tag{4.6}$$

where $0 < \rho_T \leq \rho_P < 1$ and $Var_t\left[\varepsilon_{T,t+1}\right] = \sigma_T^2$.

It is natural to posit that σ_P—the volatility of shocks to the persistent component of short rates—have declined since the late 1990s. A natural explanation is that σ_P has declined because long-run inflation expectations have become far more anchored. Indeed, there is strong evidence that shocks to the persistent component of nominal inflation have become far less volatile since the mid-1990s (Stock and Watson (2007)).

As we show below, if σ_P is large, then long-term nominal rates will be highly sensitive to changes in short-term nominal rates due to standard expectations hypothesis logic. Precisely in this vein, Gürkaynak et al. (2005) argued that the strong sensitivity of long-term nominal rates arises because long-run inflation expectations are not well-anchored and are being continuously revised in light of incoming macroeconomic news. In other words, there is no "excess sensitivity" beyond the standard sensitivity that works through a textbook expectations hypothesis channel. The otherwise puzzling sensitivity of long-term nominal rates can be resolved by recognizing that the expected persistence of nominal short rates is far higher than one might think because long-run inflation expectations are largely unmoored.

While this is a plausible explanation for the strong sensitivity observed in the 1970s, 1980s, and early 1990s, this strikes us as something of a stretch in recent years since long-run inflation expectations have been so well anchored. And, as shown by Hanson and Stein (2015), in the post-2000 period, the strong sensitivity of long-term nominal rates to movements in short-term nominal rates primarily reflects the sensitivity of long-term real rates to short-term nominal rates. Thus, to match the strong sensitivity of long-term nominal rates in recent years—a sensitivity that is only pronounced at short-horizons—it is natural to evoke shocks to bond supply that impact the term premium on long-term bonds.

Supply: We assume that the long-term nominal bond is available in an exogenous, time-varying *net supply* s_t that must be held in equilibrium by fast investors and slow-moving investors. This net supply equals the *gross supply* of long bonds *minus the demand* for long-term bonds from any unmodeled agents who participate in the bond market. Formally, we assume that s_t follows an AR(1) process:

$$s_{t+1} = \overline{s} + \rho_s (s_t - \overline{s}) + \varepsilon_{s,t+1} + C\varepsilon_{P,t+1} + C\varepsilon_{T,t+1}, \tag{4.7}$$

where $0 < \rho_s < 1$, $C > 0$, and $Var_t [\varepsilon_{s,t+1}] = \sigma_s^2$.[4] We have made no assumptions about the correlation structure among the three underlying shocks $\varepsilon_{P,t+1}$, $\varepsilon_{T,t+1}$, and $\varepsilon_{s,t+1}$. Indeed, the model can be solved for any arbitrary correlation structure among these shocks. However, in our numerical calibrations below, we will assume for simplicity that the three underlying shocks are mutually orthogonal.

In the simple case where $\sigma_s^2 = 0$ so shocks to bond supply are entirely driven by shocks

[4]One could generalize this process to allow shocks to the persistent component of short rates to have a different impact on bond supply than shocks to the transient compoent.

to short-term interest rate, one can show that

$$\begin{aligned} s_t &= \bar{s} + C\left[(i_{P,t} - \bar{\imath}) - (\rho_P - \rho_s)\sum_{j=0}^{\infty} \rho_s^j \left(i_{P,t-j-1} - \bar{\imath}\right)\right] \\ &\quad + C\left[i_{T,t} - (\rho_T - \rho_s)\sum_{j=0}^{\infty} \rho_s^j i_{T,t-j-1}\right]. \end{aligned} \tag{4.8}$$

Thus, when $\rho_s < \rho_P$ and $\rho_s < \rho_T$, bond supply depends on the differences between the current level of each component of the short rate and a geometrically-decaying, weighted-average of past values of that component.

4.1.4 Shocks to supply and the short rate

Crucially, we assume that C in equation (4.7) is positive, so that shocks to short rates are positively associated with shocks to the net supply of the long-term bonds. Our assumption that $C > 0$ can be seen as a reduced-form way of capturing a variety of distinct supply-and-demand mechanisms that help explain why negative shocks to short-term rates are associated with declines in the term premium on long-term bonds. While the precise mix of these mechanisms and their combined strength may vary over time and across countries, there is a growing consensus in the literature that mechanisms of this sort play an increasingly important role in fixed income market.

Specific mechanisms of this sort include:

The "mortgage convexity" channel. According to the mortgage convexity channel (Hanson, 2014; Malkhozov et al., 2016), negative shocks to short-term interest rates induce mortgage refinancing waves that lead to temporary declines in the duration of outstanding fixed-rate mortgages—i.e., a temporary reduction in the effective gross supply of long-term bonds. As a result, declines in short-term interest rates are associated with temporary declines in the term premium on long-term bonds. The mortgage refinancing channel is only relevant in countries like the U.S. where fixed-rate mortgages with an embedded prepayment option are an important source of mortgage financing. Importantly for us, Hanson (2014) shows that the strength of this channel grew during the 1990s as mortgage-backed securities became a larger component of the U.S. fixed-income market. In our reduced-form framework, this is equivalent to the statement than C has risen over time.

Asset and liability management by insurers and pensions. Domanski et al. (2017) and Shin (2017) point to a related convexity-based amplification mechanism that stems from the desire of insurers and pensions to match the duration of their assets and liabilities. They argue that, as interest rates decline, the duration of insurers' and pensions' liabilities tends to increase more than the duration of their long-term bond holdings. As a result, the demand for long-term bonds from insurers and pensions rises following a decline in the level of interest rates. This means that the net supply of long-term bonds that must be held by other investors declines when short-rates are low, thereby depressing term premia. This mechanism is arguably quite important in European fixed-income markets where insurers and pensions play an especially important role. And, this dynamic may have grown stronger in recent years as regulators have pushed insurers to more prudently manage their interest rate exposures.

The "reaching-for-yield" channel. According to the reaching for yield channel (Hanson and Stein, 2015), negative shocks to short-term interest rates boost the demand for long-term bonds from "yield-oriented investors." The idea is that, for either frictional or behavioral reasons, these yield-oriented investors care about the *current yield* on their portfolios over and above their expected portfolio return. Because expected mean reversion in short rates means that the yield curve is steeper when short rates are low, yield-oriented investors' demand for long-term bonds is greater when short rates are low. Holding fixed the *gross supply* of long-term bonds, this means that the *net supply* of long-term bonds that must be held by other investors is lower when short rates are low. As a result, term premia on long-term bonds are low when short rates are low.

More generally, low short rates may increase investors' effective risk appetites, thereby depressing term premia, through a variety of channels (Maddaloni and Peydró, 2011; Becker and Ivashina, 2015; Di Maggio and Kacperczyk, 2017; Drechsler et al., 2014) and Lian, Ma, and Wang (2017). Importantly for us, Lian, Ma, and Wang (2017) provide experimental evidence that there is a non-linear relationship between reaching-for-yield behavior and the level of interest rates. Specifically, building on Kahneman and Tverky's (1979) Prospect Theory, they argue that reaching for yield becomes more pronounced as interest rates fall further below some reference level that investors are accustomed to based on past experience. Relatedly, they argue that, because people tend to think in proportions and as opposed to in differences—a well-documented psychological phenomenon known as Weber's law—small return differentials loom larger in investors' minds when the level of short rates is lower (see Bordalo, Gennaioli, and Shleifer (2013)). For instance, since $2/1 > 6/5$, a 1% pick-up in yield from buying long-term bonds instead short-term bonds seems like "a better deal" to many investors when short rates are 1% than when they are 5%. Thus, their experimental evidence suggests that the reaching-for-yield channel may have grown stronger in recent years as interest rates have reached historically low levels. In the language of our reduced-form model, this again suggests that C has increased.

A behavioral over-extrapolation mechanism. According to this channel hinted at by Cieslak and Povala (2015) and Piazzesi et al. (2015), there is a set of biased investors who over-estimate the persistence of short-term interest rates. As a result, negative shocks to short rates lead this set of biased investors to demand more long-term bonds relative to rational investors who properly estimate the persistence of short rates. Again, this means that the net supply of long-term bonds that must be held by unbiased investors declines when short-rates are low, leading to a decline in term premium.

In the simplest telling, there is little reason to expect that some primitive extrapolative tendency on the part of investors should have increased since 2000. However, in a more complicated telling, the amount of over-extrapolation might have risen if some investors are using an "outdated" model that features a larger fraction of persistent short rate shocks than there has been in recent years. In other words, some investors might still be using a model of short rates that was more appropriate to the 1970s through the 1990s where there were larger shocks to trend inflation. This more complicated version of the over-extrapolation story is again consistent with the idea that C has risen since 2000.

Transitory flight-to-quality flows. A final possibility is that there can be bad macro-financial news that leads investors to expect the Federal Reserve to cut short-term rates and that, at the same time, induces transitory increases in the demand for long-term nominal bonds—i.e., flight to quality flows. As documented by many authors including Campbell et al. (2017) and Campbell et al. (2015), the correlation between stock and bond returns flipped from being positive to negative around 1998. And the increased frequency of flight-to-quality flows associated with episodes of financial instability—i.e., an increase in C—is a potential interpretation of the change in the stock-bond correlation.

4.2 Equilibrium yields

At time t, there is a mass q of fast-moving investors, each with demand b_t, and a mass $(1-q)\,k^{-1}$ of active slow-moving investors who rebalance their portfolios, each with demand

d_t. These investors must accommodate the *active supply*, which is the total supply s_t of long-term bonds less any supply held off the market by inactive slow-moving investors who do not rebalance the portfolios, $(1-q)\,k^{-1}\sum_{j=1}^{k-1} d_{t-j}$. Thus, the market-clearing condition for long-term bonds at time t is

$$\overbrace{qb_t}^{\text{Fast demand}} + \overbrace{(1-q)k^{-1}d_t}^{\text{Active slow demand}} = \overbrace{s_t}^{\text{Total bond supply}} - \overbrace{(1-q)(k^{-1}\sum_{j=1}^{k-1} d_{t-j})}^{\text{Inactive slow holdings}}. \tag{4.9}$$

We conjecture that equilibrium yields y_t and the demands of active slow-moving investors d_t are linear functions of a state vector, $\mathbf{x}_t$, that includes the steady-state deviations of both components of short-term nominal interest rates, the net supply of bonds, and holdings of bonds by inactive slow-moving investors. Formally, we conjecture that the yield on long-term bonds is

$$y_t = \alpha_0 + \alpha_1'\mathbf{x}_t \tag{4.10}$$

and that slow-moving investors' demand for long-term bonds is

$$d_t = \delta_0 + \delta_1'\mathbf{x}_t, \tag{4.11}$$

where the $(k+2)\times 1$ dimensional state vector, $\mathbf{x}_t$, is given by

$$\mathbf{x}_t = [i_{P,t} - \bar{\imath}, i_{T,t}, s_t - \bar{s}, d_{t-1} - \delta_0, \cdots, d_{t-(k-1)} - \delta_0]'. \tag{4.12}$$

These assumptions imply that the state vector follows a VAR(1) process

$$\mathbf{x}_{t+1} = \mathbf{\Gamma}\mathbf{x}_t + \epsilon_{t+1}, \tag{4.13}$$

where the transition matrix $\mathbf{\Gamma}$ depends on the parameters δ_1 that govern slow-moving investors' demand for long-term bonds.

It can be shown that equilibrium yields take the form:

$$y_t = \overbrace{\left[\bar{\imath} + \frac{1-\theta}{1-\rho_P\theta}\left(i_{P,t}-\bar{\imath}\right) + \frac{1-\theta}{1-\rho_T\theta} i_{T,t}\right]}^{\text{Expected future short-term nominal rates}} \tag{4.14}$$

$$+ \overbrace{\left[(q\tau)^{-1} V^{(1)} \left(\bar{s} - (1-q)\,\delta_0\right)\right]}^{\text{Unconditional term premia}} \tag{4.15}$$

$$+ \overbrace{\left[(q\tau)^{-1} V^{(1)} \left(\begin{array}{c} \frac{1-\theta}{1-\theta\rho_s}\left(s_t - \bar{s}\right) \\ -\left(1-\theta\right)\left(1-q\right) k^{-1} \sum_{i=0}^{\infty} \theta^i E_t[\sum_{j=0}^{k-1}(d_{t+i-j}-\delta_0)] \end{array}\right)\right]}^{\text{Conditional term premia}},$$

where $V^{(1)} = Var_t\left[rx_{t+1}\right]$ is the equilibrium variance of 1-period excess returns on long-term nominal bonds. Solving the model involves numerically finding a solution to a system of $2k$ polynomial equations in $2k$ unknowns.

4.3 Matching the main finding

Consider the slope coefficient from a regression of $y_{t+h} - y_t$ on $i_{t+h} - i_t$ in the model:

$$\beta_h = \frac{Cov\left[y_{t+h} - y_t, i_{t+h} - i_t\right]}{Var\left[i_{t+h} - i_t\right]}. \tag{4.16}$$

This is the model counterpart of the empirical regression coefficient in equation (2.1). Our main empirical findings are that β_h has declined in the post-2000 sample at low frequencies (high h) but has actually risen at high frequencies (low h). In this section, we argue that our model can match these surprising patterns if two underlying parameters shifted between the pre-2000 period and the post-2000 period:

1. σ_P **has fallen**: This means that shocks to the persistent component of short-term nominal rates have become less volatile in the post-2000 period.

2. C **has risen**: This means that the kinds of supply-and-demand-based amplification mechanisms that we emphasize have grown in importance.

4.3.1 Model-implied regression coefficients

Special case without slow-moving capital To help build intuition about the behavior of β_h in the model, we first consider the special case in which there is no slow-moving capital (i.e., if either $q = 1$ or $k = 1$). For simplicity, we also assume that the three underlying shocks—i.e., $\varepsilon_{P,t+1}$, $\varepsilon_{T,t+1}$, and $\varepsilon_{s,t+1}$—are mutually orthogonal. In this special case, the model can be solved using pencil and paper and the equilibrium yield on long-term bonds is

$$y_t = \overbrace{\left[\bar{\imath} + \frac{1-\theta}{1-\rho_P\theta}\left(i_{P,t} - \bar{\imath}\right) + \frac{1-\theta}{1-\rho_T\theta} i_{T,t}\right]}^{\text{Expected future short-term nominal rates}} + \overbrace{\tau^{-1}V^{(1)}\left[\bar{s} + \frac{1-\theta}{1-\rho_s\theta}\left(s_t - \bar{s}\right)\right]}^{\text{Term premium}}, \tag{4.17}$$

where $V^{(1)}$ is the smaller root of the following quadratic equation:

$$\begin{aligned} 0 \;=\; & \left[\left(\tau^{-1}\frac{\theta}{1-\rho_s\theta}\sigma_s\right)^2 + \left(\tau^{-1}\frac{\theta}{1-\rho_s\theta}C\sigma_P\right)^2 + \left(\tau^{-1}\frac{\theta}{1-\rho_s\theta}C\sigma_T\right)^2\right] \times \left(V^{(1)}\right)^2 \qquad (4.18) \\ & + \left[2\left(\frac{\theta}{1-\rho_P\theta}\sigma_P\right)\left(\tau^{-1}\frac{\theta}{1-\rho_s\theta}C\sigma_P\right) + 2\left(\frac{\theta}{1-\rho_T\theta}\sigma_T\right)\left(\tau^{-1}\frac{\theta}{1-\rho_s\theta}C\sigma_T\right) - 1\right] \times V^{(1)} \\ & + \left[\left(\frac{\theta}{1-\rho_P\theta}\sigma_P\right)^2 + \left(\frac{\theta}{1-\rho_T\theta}\sigma_T\right)^2\right]. \end{aligned}$$

In this case, the model-implied regression coefficient is

$$\begin{aligned} \beta_h \;=\; & \frac{\begin{array}{c}\frac{1-\theta}{1-\rho_P\theta} \times Var\left[\Delta_h i_{P,t+h}\right] + \frac{1-\theta}{1-\rho_T\theta} \times Var\left[\Delta_h i_{T,t+h}\right] \\ +\tau^{-1}V^{(1)}\frac{1-\theta}{1-\rho_s\theta} \times Cov\left[\Delta_h i_{P,t+h}, \Delta_h s_{t+h}\right] + \tau^{-1}V^{(1)}\frac{1-\theta}{1-\rho_s\theta} \times Cov\left[\Delta_h i_{T,t+h}, \Delta_h s_{t+h}\right]\end{array}}{Var\left[\Delta_h i_{P,t+h}\right] + Var\left[\Delta_h i_{T,t+h}\right]} \qquad (4.19) \\ =\; & \frac{\begin{array}{c}\frac{1-\theta}{1-\rho_P\theta} \times 2\frac{1-\rho_P^h}{1-\rho_P^2}\sigma_P^2 + \frac{1-\theta}{1-\rho_T\theta} \times 2\frac{1-\rho_T^h}{1-\rho_T^2}\sigma_T^2 \\ +\tau^{-1}V^{(1)}\frac{1-\theta}{1-\rho_s\theta} \times \frac{2-\rho_s^h-\rho_P^h}{1-\rho_s\rho_P}C\sigma_P^2 + \tau^{-1}V^{(1)}\frac{1-\theta}{1-\rho_s\theta} \times \frac{2-\rho_s^h-\rho_T^h}{1-\rho_s\rho_T}C\sigma_T^2\end{array}}{2\frac{1-\rho_P^h}{1-\rho_P^2}\sigma_P^2 + 2\frac{1-\rho_T^h}{1-\rho_T^2}\sigma_T^2}. \end{aligned}$$

Suppose that $C \geq 0$ and $\rho_s \leq \rho_T \leq \rho_P$. Inspecting equation (4.19), it is easy to see that:

- **When $C = 0$, the level of β_h is increasing in σ_p for all h.** An increase in σ_P raises the fraction of total short-rate variation at all horizons that is due to movements in the more persistent component (i.e., raises $Var\left[\Delta_h i_{P,t+h}\right] / \left(Var\left[\Delta_h i_{P,t+h}\right] + Var\left[\Delta_h i_{T,t+h}\right]\right)$)

for all h). Since shocks to the permanent component of short rates have larger impact on long-term yields than shocks to the transitory component via a straightforward expectations hypothesis channel (i.e., $(1-\theta)/(1-\rho_P\theta) > (1-\theta)/(1-\rho_T\theta)$), an increase in σ_P raises the level of β_h at all horizons.

Thus, if σ_P declined between the pre-2000 and post-2000 periods as we argued above, this would lead β_h to decline for large h.

We next consider the way β_h behaves a function of h. Again, using equation (4.19), it is easy to show that:

- **β_h is a constant that is independent of h when (i) $C = 0$ or (ii) $C > 0$ and $\rho_s = \rho_T = \rho_P$.** In case (i), there is no excess sensitivity. In case (ii), there is excess sensitivity but it is the same irrespective of the horizon h. Assuming that $\sigma_s^2 = 0$, this is because $\Delta_h s_{t+h} = C\Delta_h i_{P,t+h} + C\Delta_h i_{T,t+h}$ and $Var[\Delta_h i_{P,t+h}] / (Var[\Delta_h i_{P,t+h}] + Var[\Delta_h i_{T,t+h}]) = \sigma_P^2 / (\sigma_P^2 + \sigma_T^2)$, which is independent of h.

- **β_h is an increasing function of h when $C = 0$ and $\rho_T < \rho_P$.** In this case, there is no excess sensitivity. However, β_h rises with h since (i) movements in the more persistent component of short rates are associated with larger movement in long-term yields (i.e., $(1-\theta)/(1-\rho_P\theta) > (1-\theta)/(1-\rho_T\theta)$) and (ii) because the persistent component of short rates dominates changes in short rates at longer horizons (i.e., $Var[\Delta_h i_{P,t+h}] / (Var[\Delta_h i_{P,t+h}] + Var[\Delta_h i_{T,t+h}])$ rises with h).

- **β_h is a decreasing function of h when $C > 0$ and $\rho_s < \rho_T = \rho_P$.** In this case, long-term interest rates exhibit excess sensitivity to movements in short-term rates that declines with the horizon h. Intuitively, if the supply shocks induced by innovations to short rates are more transient than the underlying shocks to short rates, then term premia will react more in the short-term than in the long-term. Thus, there will be greater excess sensitivity in the short term than over the longer term.

Thus, assuming $\rho_s < \rho_T$, if C rose significantly between the pre-2000 and post-2000 periods as we argued above, we would expect β_h to be a relatively constant (or even slightly increasing) function h in the pre-2000 period and a decreasing function of h in the post-2000 period.

General case with slow-moving capital Given a solution to the general model with slow-moving capital, we can also work out the model-implied regression coefficients β_h. Since the state vector $\mathbf{x}_t$ follows a VAR(1) in the model, if we let $\mathbf{V} = Var\left[\mathbf{x}_t\right]$, we have $vec(\mathbf{V}) = (I - \mathbf{\Gamma} \otimes \mathbf{\Gamma})^{-1} vec(\mathbf{\Sigma})$. We also have $Cov\left[\mathbf{x}_{t+j}, \mathbf{x}_t'\right] = \mathbf{\Gamma}^j \mathbf{V}$ and $Cov[\mathbf{x}_t, \mathbf{x}_{t+j}'] = \mathbf{V}\left(\mathbf{\Gamma}'\right)^j$, so $Var\left[\mathbf{x}_{t+h} - \mathbf{x}_t\right] = 2\mathbf{V} - \mathbf{\Gamma}^h \mathbf{V} - \mathbf{V}(\mathbf{\Gamma}')^h$. Thus, letting $\mathbf{e}_i$ denote the vector with ones in the first and second positions and zeros everywhere else, we have[5]:

$$\beta_h = \frac{Cov[\alpha_1' \left(\mathbf{x}_{t+h} - \mathbf{x}_t\right), \left(\mathbf{x}_{t+h} - \mathbf{x}_t\right)' \mathbf{e}_i]}{Var\left[\left(\mathbf{x}_{t+h} - \mathbf{x}_t\right)' \mathbf{e}_i\right]} = \frac{\alpha_1'(2\mathbf{V} - \mathbf{\Gamma}^h \mathbf{V} - \mathbf{V}(\mathbf{\Gamma}')^h)\mathbf{e}_i}{\mathbf{e}_i(2\mathbf{V} - \mathbf{\Gamma}^h \mathbf{V} - \mathbf{V}(\mathbf{\Gamma}')^h)\mathbf{e}_i}. \tag{4.20}$$

In the general model, the assumption that $C > 0$ is necessary to match the decline in β_h as a function of h. Assuming that $C > 0$ so there is excess sensitivity, two features of our general model can help match the finding that β_h is a declining function of h. First, if $\rho_s < \rho_T = \rho_P$, so supply shocks are less persistent than short-rate shocks, then, as shown above, β_h will be a declining function of h even in the absence of slow-moving capital. For instance, if shocks to net supply are driven by the mortgage convexity channel, then it is natural to suppose that the resulting supply shocks are highly transient in nature since

[5] This approach to computing β_h assumes that $\rho_P < 1$, so $\mathbf{x}_t$ is stationary. However, the model can also be solved in the limiting case where $\rho_P = 1$. Although $\mathbf{x}_t$ is not stationary in that limiting case and $Var\left[\mathbf{x}_t\right]$ is undefined, $Var\left[\mathbf{x}_{t+h} - \mathbf{x}_t\right]$ is still well-defined and can be easily computed.

Hanson (2014) shows that even persistent declines in rates only induce short-lived mortgage refinancing waves.

Second, when there is slow-moving capital (i.e., if $k > 1$ and $q < 1$), the short-run demand curve for long-term bonds is steeper than the long-run demand curve. As a result, so long as $C > 0$, β_h will decline with h even if $\rho_s = \rho_T = \rho_P$. For instance, the simplest versions of the "reaching-for-yield" channel (Hanson and Stein, 2015) would suggest that $\rho_s = \rho_T = \rho_P$. Thus, in order for this channel to explain our results, we would need to appeal to slow-moving capital effects (Duffie, 2010; Greenwood et al., 2016), which imply that these induced supply shocks move the market along a short-run demand curve that is a good deal steeper than the long-run demand curve.

In our framework, whether a meaningful "recruitment channel" of monetary policy exists at business-cycle frequencies, hinges on the relationship between ρ_s and ρ_T. When $C > 0$ and when ρ_s is well below ρ_T as under the mortgage-convexity view, then Stein's (2013) "recruitment channel" would not add significantly to the potency of monetary policy over the business cycle: the induced shocks to term premia would be too fleeting to have a meaningful impact on firm or household behavior. By contrast, when $C > 0$ and $\rho_s \approx \rho_T$ as under the reaching-for-yield view, then Stein's (2013) "recruitment channel" would add to the potency of monetary policy over the business cycle. Nonetheless, in the presence of slow-moving-capital ($q < 1$ and $k > 1$), one would tend to meaningfully overstate the influence of monetary policy on term premia, and perhaps other financial conditions, by focusing on the short-run market reaction to Fed announcements. In such a world, the Fed should care about the way that monetary policy impacts financial conditions at business-cycle frequencies, but the Fed should focus less on the immediate market response to its announcements since much of the latter is may be mean reverting.

In summary, for β_h to be a steeply declining function of h as in the post-2000 data, we need (i) $C > 0$ and (ii) either $\rho_s < \rho_T$ or slow-moving investors (i.e., $k > 1$ and $q < 1$). In practice, we believe that both $\rho_s < \rho_T$ and slow-moving capital likely play some role in explaining why β_h is a declining function of h in the recent data. Furthermore, in our numerical calibrations, we find that these mechanisms reinforce one another: it is easiest to match the decline in β_h using calibrations, such as our baseline calibration below, that feature both mechanisms.

4.3.2 Model calibration

We consider a illustrative calibration of the model in which each time period corresponds to a month. We assume that the following parameters were the same in both the pre-2000 and post-2000 periods:

- **Persistence:** $\rho_P = 0.995$, $\rho_T = 0.96$, and $\rho_s = 0.75$. This implies that shocks to the persistent component of short rates have a half-life of 11.5 years, shocks to the transient component of short rates have a half-life of 1.4 years, and shocks to the net supply of long-term bond have a half-life of just over 2 months.

- **Slow-moving capital:** $q = 25\%$ and $k = 12$. Thus, $1 - q = 75\%$ of the investors are slow-moving and only rebalance their bond portfolios every 12 months.

- **Volatility of the transient component of short rates:** $\sigma_T^2 = 0.15\%$. Thus, the standard deviation of the transient component of short rates is $\sqrt{\sigma_T^2 / (1 - \rho_T^2)} = 1.38\%$.

- **The transient component of short rates is uncorrelated with the permanent component:** For simplicity, we assume that $Cov_t [\varepsilon_{P,t+1}, \varepsilon_{T,t+1}] = 0$ which implies $Cov [i_{P,t}, i_{T,t}] = 0$.

- **No independent supply shocks:** $\sigma_s^2 = 0$. We make this assumption purely for simplicity. This means that the supply shocks induced by shocks to short rates are the only reason term premia vary in this calibration of the model.

- **Other parameters:** $\tau = 0.5$ and $\theta = 119/120$, so the duration of the perpetuity is $D = 1/(1-\theta) = 120$ months—i.e., 10 years.

We assume that two model parameters, C and σ_P, changed between the pre-2000 and the post-2000 periods. For the pre-2000 period, we assume that:

- **Large persistent component of short rates:** $\sigma_P^2 = 0.15\%$. Thus, the standard deviation of the persistent component of short rates is $3.88\% = \sqrt{\sigma_P^2/(1-\rho_P^2)}$ and the standard deviation of the short rate is $4.12\% = \sqrt{(1.38\%)^2 + (3.88\%)^2}$. This compares with a pre-2000 volatility of 1-year yields of 2.63% in the data.

- **No supply shocks induced by short rate shocks:** $C = 0$. Thus, we assume that there is no excess sensitivity in the early period.

By contrast, for the post-2000 period, we assume that:

- **Small persistent component of short rates:** $\sigma_P^2 = 0.015\%$. Thus, the standard deviation of the persistent component of short rates is 1.23% in the post-2000 period and the standard deviation of the short rate is $1.85\% = \sqrt{(1.38\%)^2 + (1.23\%)^2}$. This compares with a post-2000 volatility of 1-year yields of 1.85% in the data.

- **Important supply shocks induced by short rate shocks:** $C = 0.55 > 0$.

Figure 6 plots the model-implied coefficients β_h in equation (4.16) against horizon (h) in months for the pre-2000 calibration and the post-2000 calibration. In the pre-2000 calibration where σ_P is large and $C = 0$, the level of β_h is high and β_h rises gradually with h because the more persistent component of short rates dominates changes in short rates at longer horizons. By contrast, in the post-2000 calibration where σ_P is smaller and C is large, β_h declines steeply with h and, since σ_P is lower, eventually reaches a lower level than in the pre-2000 calibration for large h. As emphasized above, β_h declines steeply with h in the post-2000 calibration because short-rate rate shocks give rise to transient shocks to the supply of long-term bonds ($C > 0$ and $\rho_s < \rho_T$) that encounter of short-run demand curve that is far steeper than the long-run demand curve because of slow-moving capital ($q < 1$ and $k > 1$).

In Figure 7, we show the model-implied impulse response functions following a one-time +100 basis point shock to the transient component of short rates that lands in month $t = 13$. By definition, the long-term bond yield is the sum of an expectations hypothesis component and a term premium component $y_t = eh_t + tp_t$, where $eh_t = \bar{\imath} + [(1-\theta)\,/\,(1-\rho_P\theta)]\,(i_{P,t} - \bar{\imath}) + [(1-\theta)\,/\,(1-\rho_T\theta)]\,i_{T,t}$. Thus, the term spread (i.e., the slope factor) is $ts_t = y_t - i_t = tp_t + (eh_t - i_t)$, where $(eh_t - i_t) = -\theta\,[(1-\rho_P)\,/\,(1-\rho_P\theta)]\,(i_{P,t} - \bar{\imath}) - \theta\,[(1-\rho_T)\,/\,(1-\rho_T\theta)]\,i_{T,t}$. We show the impulse responses for short-term rates (i_t), long-term yields (y_t), the term spread ($y_t - i_t$), and the term premium (tp_t) in Figure 7. The initial shock to short rates at $t = 13$ leads to a rise in term premia. This means that, relative to the expectations hypothesis benchmark, long-term rates are excessively sensitive to movements in the short rate. Nonetheless, the rise in short-rates still causes the yield curve to flatten contemporaneously. This is because the normal flattening due to the expectations hypothesis outweighs the steepening due to the rise in term premia. But, the rise in term premia tends to wear off quickly, explaining our stylized fact. Indeed, the initial rise in short rates is associated with additional future flattening of the yield curve over the following months. And, consistent with the decomposition of β_h in equation (3.1), raising C or increasing the degree of slow-moving capital raises $Corr(\Delta i_t, \Delta ts_t)$ in the model, giving rise to high-frequency excess sensitivity. These same forces tend to lower $Corr(\Delta i_t, \Delta ts_{t+j})$ for $j > 0$, explaining low-frequency decoupling.

In addition to quantitatively matching the β_h coefficients in the pre-2000 and post-2000 periods, our calibrations are also capable of qualitatively matching the related empirical facts that we documented above. Specifically, let $L_t = i_t$ and $S_t = y_t - i_t$ denote the model-implied level and slope of the yield curve, respectively. If we then estimate

$$S_{t+1} = \delta_0 + \delta_1 L_t + \delta_2 S_t + \delta_3 (L_t - L_{t-h}) + \delta_4 (S_t - S_{t-h}) + \varepsilon_{S,t+1}$$

using data simulated from the model, we obtain $\delta_4 < 0$ for the post-2000 calibration and $\delta_4 = 0$ for the pre-2000 calibration. The intuition for the former result is that, all else equal, past increases in the level of interest rates are associated with higher levels of active bond supply when $C > 0$—i.e., the right-hand side of equation (4.9) is higher—and, hence, a higher current risk premium on long-term bonds. Since the risk premium on long-term bonds is $E_t[rx_{t+1}] = S_t - \theta(1-\theta)^{-1}(E_t[\Delta S_{t+1}] + E_t[\Delta L_{t+1}])$, a higher risk premium means that, all else equal, $E_t[\Delta S_{t+1}]$ is lower—i.e., we have $\delta_3 < 0$.

Finally, we are more likely to have $sign(y_{t+h} - y_t) \neq sign(i_{t+h} - i_t)$ for $h = 6$ or $h = 12$ months in data simulated from the post-2000 calibration than in data from the pre-2000 calibration. In other words, the calibrated model is consistent with the increasing prevalence of "interest rate conundrums" in the post-2000 period.

5 Trading strategies and affine term structure models

This section first describes trading strategies that exploit the predictability described in Section 3. We then show how standard Markovian affine term structure models have a hard time fitting the fact that the sensitivity of long to short rates, β_h, declines at lower frequencies in recent years.

5.1 Trading strategies

Level and slope-mimicking portfolios We can recast our findings as a finding about bond return predictability. To do so, we form bond portfolios that locally mimic changes in the level and slope factors—i.e., in response to small changes in yields. The 1-month return on n-year zero coupons bonds in month t is defined as $R_t^{(n)} = (P_t^{(n-1/12)} - P_{t-1}^{(n)})/P_{t-1}^{(n)}$. Because $L_t \equiv y_t^{(1)}$ and $S_t \equiv y_t^{(10)} - y_t^{(1)}$, we have $R_t^{(10)} \approx -10 \times (\Delta L_t + \Delta S_t)$ and $R_{t+1}^{(1)} \approx -1 \times \Delta L_t$ for a small changes in yields. We then follow Joslin et al. (2014) and construct factor-mimicking portfolios with a weight w_i on bonds with yield n_i. The excess returns on these portfolios are:

$$RX_t = \frac{\sum_i w_i (R_t^{(n_i)} - R_t^{(1/12)})}{|\sum_i w_i|},$$

where $R_t^{(1/12)}$ is the riskless return on 1-month bills.[6] The level-mimicking portfolio has a weight -1 on the 1-year bond and no weight on any other bonds. Thus, the level- mimicking portfolio has a monthly excess return of:

$$RX_t^{LEVEL} = -1 \times (R_t^{(1)} - R_t^{(1/12)}) \approx \Delta L_t + R_t^{(1/12)}.$$

The slope-mimicking portfolio has a weight 1 on the 1-year bond and -0.1 on the 10-year bond and has an excess return of:

$$RX_t^{SLOPE} = \frac{1}{0.9} \times (R_t^{(1)} - R_t^{(1/12)}) - \frac{0.1}{0.9} \times (R_t^{(10)} - R_t^{(1/12)}) \approx \frac{\Delta S_t}{0.9} - R_t^{(1/12)}.$$

Note that the slope-mimicking portfolio is approximately hedged against parallel shifts in the level of the yield curve and that RX_t^{SLOPE} corresponds to the excess returns on what

[6] Joslin et al. (2014) also consider excess returns on such factor-mimicking portfolios, although they were using principal components whereas we define level and slope from fixed points on the yield curve.

fixed-income practitioners would call a "steepener" trade—i.e., a trade that will profit if the yield curve steepens.

We then consider the following predictive regressions:

$$\begin{aligned} RX_{t+1}^{LEVEL} &= \delta_{0L} + \delta_{1L} L_t + \delta_{2L} S_t + \delta_{3L} C_t \\ &\quad + \delta_{4L}(L_t - L_{t-h}) + \delta_{5L}(S_t - S_{t-h}) + \delta_{6L}(C_t - C_{t-h}) + \varepsilon_{L,t+1} \end{aligned} \tag{5.1}$$

and

$$\begin{aligned} RX_{t+1}^{SLOPE} &= \delta_{0S} + \delta_{1S} L_t + \delta_{2S} S_t + \delta_{3S} C_t \\ &\quad + \delta_{4S}(L_t - L_{t-h}) + \delta_{5S}(S_t - S_{t-h}) + \delta_{6S}(C_t - C_{t-h}) + \varepsilon_{S,t+1}, \end{aligned} \tag{5.2}$$

where RX_t^{LEVEL} and RX_t^{SLOPE} denote the monthly excess returns on level- and slope-mimicking portfolios as defined above. We report the results from estimating these predictive regressions in Table 4.

Trading strategies The results in Table 4 are entirely consistent with those in Table 3. This is as expected because the dependent variables in equations (3.2) and (3.3) are the future level and slope respectively, whereas the dependent variables in equations (5.1) and (5.2) are approximately the future changes in level and slope. Specifically, in the post-2000 sample, lagged changes in level are highly significant predictors of excess returns on the slope-mimicking portfolio.[7] Our basic finding is that an increase in the level of rates has

[7] Many researchers have examined bond return predictability by considering the excess returns on 10-year or other long-term bonds bonds. By construction, the excess return on a 10-year bond is a linear combination of the excess returns on our level- and slope- mimicking portfolios: $R_t^{(10)} - R_t^{(1/12)} = -9 \times RX_t^{SLOPE} - 10 \times RX_t^{LEVEL}$. From Table 5, in the post-2000 sample, the excess return on the level-mimicking portfolio depends positively on $L_{t-1} - L_{t-h}$ but the excess return on the slope-mimicking portfolio depends negatively on $L_{t-1} - L_{t-h}$. The two effects partially cancel out when predicting 10 -year excess returns, although the net effect (not shown) is estimated to be positive. Thus, we prefer to work with level- and slope- mimicking portfolios as it makes the dependence on lagged changes in levels more stark, and these are clearly tradeable portfolios.

been followed by subsequent yield curve flattening. As another way of assessing this, we consider trading strategies in which every month, the investor decides to take either a long or short position in the slope-mimicking portfolio[8]. We assume that the investor takes a long (short) position in the slope-mimicking portfolio from month t to month $t+1$ if $L_t < L_{t-h}$ ($L_t > L_{t-h}$). Alternatively, we could assume that the investor takes a position in the slope-mimicking portfolio from month t to month $t+1$ that is proportional to $-(L_t - L_{t-h})$. And then we compute the annualized Sharpe ratios of these trading strategies for different choices of h, over the post-2000 sample. Table 5 shows implied annualized Sharpe ratios between 0.4 to 0.7.

5.2 Affine Term-Structure Models

Affine term-structure models (ATSMs) are a widely-used, reduced-form tools for understanding bond yields. A standard discrete-time affine term-structure model (Duffee, 2002; Duffie and Kan, 1996) starts from the assumption that there is a $m \times 1$ state vector $\mathbf{x}_t$ that follows a VAR(1):

$$\mathbf{x}_t = \mu + \mathbf{\Phi}\mathbf{x}_{t-1} + \mathbf{\Sigma}\varepsilon_t, \tag{5.3}$$

where the error term is Gaussian with mean zero and identity variance-covariance matrix. The short-term interest rate is $r_t = \delta_0 + \delta_1'\mathbf{x}_t$. Meanwhile, the pricing kernel is

$$M_{t+1} = \exp(-r_t - \lambda_t'\varepsilon_{t+1} - \frac{1}{2}\lambda_t'\lambda_t), \tag{5.4}$$

where $\lambda_t = \lambda_0 + \mathbf{\Lambda_1}\mathbf{x}_t$. After extensive, but well-known algebra, it follows that the price of an n-period zero-coupon bond, $P_t^{(n)}$, is

$$P_t^{(n)} = E_t\left[\Pi_{i=1}^n M_{i+1}\right] = \exp(a_{(n)} + \mathbf{b}_{(n)}'\mathbf{x}_t), \tag{5.5}$$

[8]There is a huge literature on momentum and contrarian strategies in equity markets. The literature on such strategies in the bond market is much smaller, but Durham (2013) considers returns on certain bond momentum strategies.

where $a_{(n)}$ is a scalar and $\mathbf{b}_{(n)}$ is an $m \times 1$ vector. Letting $\mu^{*}= \mu - \mathbf{\Sigma}\lambda_0$ and $\mathbf{\Phi}^{*}= \mathbf{\Phi} - \mathbf{\Sigma}\mathbf{\Lambda}_1$, $a_{(n)}$ and $\mathbf{b}_{(n)}$ satisfy the recursions:

$$a_{(n+1)} = -\delta_0 + a_{(n)} + \mathbf{b}'_{(n)}\mu^{*} + \frac{1}{2}\mathbf{b}'_{(n)}\mathbf{\Sigma}\mathbf{\Sigma}'\mathbf{b}_{(n)} \quad (5.6)$$

$$\mathbf{b}_{(n+1)} = \mathbf{\Phi}^{*\prime}\mathbf{b}_{(n)} - \delta_1, \quad (5.7)$$

starting from $a_{(1)} = -\delta_0$ and $\mathbf{b}_1 = -\delta_{(1)}$. The continuously compounded yield on an n-period zero-coupon bond, $y_t^{(n)}$, is in turn given by

$$y_t^{(n)} = -n^{-1}\log(P_t^{(n)}) = -n^{-1}a_{(n)} - n^{-1}\mathbf{b}'_{(n)}\mathbf{x}_t. \quad (5.8)$$

Let $\mathbf{\Gamma}(j) = E[(\mathbf{x}_{t+j} - E\left[\mathbf{x}_{t+j}\right])\left(\mathbf{x}_t - E\left[\mathbf{x}_t\right]\right)']$ be the autocovariance function of the state vector, which can be obtained from the equations $vec(\mathbf{\Gamma}(0)) = (\mathbf{I} - \mathbf{\Phi} \otimes \mathbf{\Phi})^{-1}vec(\mathbf{\Sigma}\mathbf{\Sigma}')$ and $\mathbf{\Gamma}(j) = \mathbf{\Phi}^j\mathbf{\Gamma}(0)$ for $j \geq 1$.The population coefficient in a regression of h-month changes in 120-month yields on h-month changes in 12-month yields is then

$$\beta_h = \frac{E[(y_{t+h}^{(120)} - y_t^{(120)})(y_{t+h}^{(12)} - y_t^{(12)})]}{E[(y_{t+h}^{(12)} - y_t^{(12)})^2]} = \frac{1}{10}\frac{\mathbf{b}'_{(120)}[2\mathbf{\Gamma}(0) - \mathbf{\Gamma}(h) - \mathbf{\Gamma}(h)']\mathbf{b}_{(12)}}{\mathbf{b}'_{(12)}[2\mathbf{\Gamma}(0) - \mathbf{\Gamma}(h) - \mathbf{\Gamma}(h)']\mathbf{b}_{(12)}}. \quad (5.9)$$

We can fit this model using the first K principal components of 1- to 10-year yields as the state variables in $\mathbf{x}_t$, and applying the estimation methodology of Adrian et al. (2013). We do this in the pre-2000 and post-2000 samples separately. We take the estimated parameters and work out the model-implied regression coefficients using equation (5.9). These coefficients are shown in Panel A of Table 6. As can be seen, the model fails to match the low-frequency decoupling between short- and long-term yields seen in the post-2000 data.

An alternative affine term-structure model augments the state vector $\mathbf{x}_t$ to include not just K principal components of yields, but also L additional lags of these principal components. These lags are treated as "unspanned factors." This means that if the first K elements of the state vector $\mathbf{x}_t$ are the current principal components, all but the first K elements of δ_1

are equal to zero, and the upper right $K \times KL$ block of $\mathbf{\Phi}^*$ is a matrix of zeros. This implies that the lags are important for expected future yields, but are not reflected in the yield curve today. It's a rather unusual model, but has been considered in Joslin et al. (2013), and is a non-Markovian model. Again, the parameters can be estimated, with the restriction that the lags are unspanned factors as described in Adrian et al. (2013). The model-implied regression coefficients can again be deduced using equation (5.9). These coefficients are shown in Panel B of Table 6. The augmented model is able to get reasonably close to matching the regression coefficients at both high- and low-frequencies and in both samples.

Our conclusion is that the affine term structure model needs lagged principal components to match the low-frequency decoupling of yields in recent years. A large number of static principal components does not do the job. These findings are consistent with the lead-lag relations between changes in level and slope that we documented in the previous section.

6 Conclusions

The excess sensitivity of changes in long rates to changes in short rates is a long-standing puzzle about the term structure of interest rates. In this paper, we have documented that this puzzle has disappeared since 2000 when looking at low-frequency changes. As a result, low-frequency decoupling between long and short rates—the phenomenon that former Federal Reserve Chairman Greenspan called a conundrum—has become increasingly common. At the same time, we find that high-frequency excess sensitively has actually become stronger since 2000.

From an expectations hypothesis perspective, the puzzle is not the weak relationship between short- and long-term rates observed recently at low frequencies. Instead, the puzzle is why this relationship was previously so strong at low frequencies and why it still remains so strong at high frequencies. We have proposed a simple model that can explain these stylized facts. In the model, before 2000 there was a sizeable persistent component in short rates due to uncertainty about long-run trend inflation. Since 2000, the persistent component has become much smaller. However, risk-averse investors demand a term premium to compensate them for the risk of holding long-term bonds, and that term premium is increasing in the net supply of long-term bonds that investors must hold. As a shorthand for a variety of amplification mechanisms—including "reaching for yield", mortgage convexity hedging flows, and asset and liability management by insurers—we assume that shocks to the short rate lead to an increase in the net supply of long-term bonds. As a result, shocks to short rates move term premia in the same direction, giving rise to excess sensitivity at high frequencies. Our model also incorporates slow-moving capital, which makes the demand curve for bonds steeper in the short run than in the long run. This allows the model to capture the surprising frequency-specific sensitivity of long rates to short rates that we observe in recent data.

Our findings have important implications for the transmission of monetary policy and for event-study methodology. The excess sensitivity of long-term yields reinforces the effects of monetary policy, but in recent years this channel of monetary policy transmission has been far more short-lived than one might conclude based on high-frequency evidence. That, in turn, makes it harder for monetary policy to influence aggregate demand, and requires the central bank to move monetary policy more aggressively to achieve the same effect. The event-study methodology studies high-frequency responses of asset prices in windows where the nature of the underlying shock is easy to identify. However, part of the high-frequency response of long rates to shocks to short rates represents term premium movements that appear to wear off systematically over time. Consequently, it is important to remember that the event-study methodology only measures high-frequency responses and that the effects may often be more muted at the lower frequencies that are typically of greatest concern to macroeconomists and policymakers.

Table 1: Regression of changes in long-term rates on short-term rates. This table reports the estimated slope coefficients from equations (2.1) and (2.2) for each reported sample. The dependent variable is the change in the 10-year yield or forward rate, either nominal, real or their difference (IC, or inflation compensation). The independent variable is the change in the 1-year nominal yield in all cases. Changes are considered with daily data, and with monthly data using monthly ($h = 1$), quarterly ($h = 3$), semi-annual ($h = 6$) and annual ($h = 12$) horizons. We report Newey-West standard errors in brackets, using a lag truncation parameter of $1.5 \times (h-1)$ (rounded to the nearest integer). Significance: $^{*}p < 0.1$, $^{**}\ p < 0.05$, $^{***}p < 0.01$.

Panel A: 10-year zero coupon yields and IC

	(1) Nominal	(2) Nominal	(3) Real	(4) IC
Daily	0.56***	0.86***	0.55***	0.31***
	[0.02]	[0.03]	[0.03]	[0.02]
Monthly	0.46***	0.64***	0.37***	0.26**
	[0.04]	[0.12]	[0.10]	[0.10]
Quarterly	0.48***	0.42***	0.21*	0.22*
	[0.04]	[0.07]	[0.11]	[0.13]
Semi-annual	0.50***	0.31***	0.20**	0.12
	[0.04]	[0.07]	[0.08]	[0.10]
Yearly	0.56***	0.20***	0.13**	0.07
	[0.05]	[0.04]	[0.06]	[0.05]
Sample	1971-1999	2000-2017	2000-2017	2000-2017

Panel B: 10-year instantaneous forward yields and IC

	(1) Nominal	(2) Nominal	(3) Real	(4) IC
Daily	0.39***	0.48***	0.31***	0.18***
	[0.03]	[0.04]	[0.03]	[0.03]
Monthly	0.29***	0.22	0.17**	0.06
	[0.04]	[0.14]	[0.08]	[0.09]
Quarterly	0.31***	0.03	0.08	-0.04
	[0.05]	[0.09]	[0.05]	[0.05]
Semi-annual	0.33***	-0.06	0.03	-0.09**
	[0.06]	[0.07]	[0.04]	[0.04]
Yearly	0.39***	-0.17***	-0.03	-0.14***
	[0.07]	[0.05]	[0.05]	[0.03]
Sample	1971-1999	2000-2017	2000-2017	2000-2017

Table 2: Regression of changes in long-term international rates on short-term rates. This table reports the estimated slope coefficients from equations (2.1) and (2.2) for UK, German and Canadian interest rates on each reported sample. Daily data for Germany are not available pre-1997. The dependent variable is the change in the 10-year yield or forward rate, either nominal, real or their difference (IC, or inflation compensation). The independent variable is the change in the 1-year nominal yield in all cases. Changes are considered with daily data, and with monthly data using monthly ($h = 1$), quarterly ($h = 3$), semi-annual ($h = 6$) and annual ($h = 12$) horizons. We report Newey-West standard errors in brackets, using a lag truncation parameter of $1.5 \times (h-1)$ (rounded to the nearest integer). Significance: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Panel A: UK 10-year zero-coupon yields and inflation compensation

	(1) Nominal	(2) Nominal	(3) Real	(4) Real	(5) IC	(6) IC
Daily	0.44***	0.86***	0.14***	0.63***	0.29***	0.23***
	[0.04]	[0.03]	[0.01]	[0.03]	[0.04]	[0.02]
Monthly	0.47***	0.55***	0.19***	0.12	0.28***	0.43***
	[0.06]	[0.14]	[0.05]	[0.26]	[0.09]	[0.13]
Quarterly	0.49***	0.43***	0.23***	0.04	0.26***	0.39***
	[0.08]	[0.10]	[0.04]	[0.18]	[0.10]	[0.10]
Semi-annual	0.45***	0.39***	0.22***	0.07	0.23**	0.32***
	[0.09]	[0.08]	[0.05]	[0.11]	[0.11]	[0.06]
Yearly	0.38***	0.29***	0.16**	0.05	0.22***	0.24***
	[0.06]	[0.06]	[0.06]	[0.08]	[0.08]	[0.03]
Sample	1985-1999	2000-2017	1985-1999	2000-2017	1985-1999	2000-2017

Panel B: German and Canadian 10-year nominal zero coupon yields

	(1) DE	(2) DE	(3) CAN	(4) CAN
Daily		0.65***	0.42***	0.71***
		[0.03]	[0.03]	[0.03]
Monthly	0.34***	0.50***	0.46***	0.51***
	[0.06]	[0.11]	[0.05]	[0.07]
Quarterly	0.41***	0.44***	0.51***	0.38***
	[0.04]	[0.07]	[0.05]	[0.05]
Semi-annual	0.41***	0.41***	0.50***	0.26***
	[0.04]	[0.08]	[0.07]	[0.05]
Yearly	0.43***	0.33***	0.43***	0.12**
	[0.04]	[0.10]	[0.08]	[0.06]
Sample	1972-1999	2000-2017	1986-1999	2000-2017

Table 3: Estimates of predictive equations for level, slope and curvature. This table reports the estimated slope coefficients in equations (3.2), (3.3) and (3.4) over August 1971-December 2000 and January 2001-December 2017 subsamples, with $h = 12$. White standard errors, are included. Dependent variables are monthly changes in level and slope. The table also shows AIC and BIC values (to be minimized) for each possible specification of the system of three equations. Lastly, the implied coefficient β_h in equation (2.1) corresponding to each possible specification of the system is reported. Significance: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

	Pre-2000				**Post-2000**			
Dependent Variable: Level								
L_t	0.972***	0.969***	0.952***	0.955***	0.961***	0.980***	0.983***	0.985***
	[0.028]	[0.030]	[0.028]	[0.027]	[0.012]	[0.009]	[0.009]	[0.010]
S_t	-0.077	-0.076	-0.107	-0.045	-0.070**	-0.012	-0.009	-0.001
	[0.130]	[0.146]	[0.143]	[0.118]	[0.021]	[0.020]	[0.020]	[0.019]
C_t	0.281	0.336	0.333	0.126	0.164	0.119	0.135	0.094
	[0.481]	[0.517]	[0.515]	[0.431]	[0.085]	[0.083]	[0.085]	[0.080]
$L_t - L_{t-12}$		0.022	0.067**	0.055		0.062***	0.045*	0.039*
		[0.023]	[0.025]	[0.031]		[0.018]	[0.018]	[0.019]
$S_t - S_{t-12}$			0.110*	0.010			-0.023	-0.044
			[0.051]	[0.106]			[0.016]	[0.025]
$C_t - C_{t-12}$				0.303				0.070
				[0.314]				[0.067]
Dependent Variable: Slope								
L_t	0.014	0.017	0.025	0.024	0.011	-0.014	-0.013	-0.012
	[0.018]	[0.019]	[0.018]	[0.018]	[0.013]	[0.011]	[0.011]	[0.010]
S_t	1.039***	1.037***	1.051***	1.023***	1.020***	0.941***	0.942***	0.946***
	[0.087]	[0.097]	[0.095]	[0.081]	[0.033]	[0.033]	[0.032]	[0.030]
C_t	-0.265	-0.319	-0.318	-0.220	-0.195	-0.135	-0.125	-0.140
	[0.320]	[0.343]	[0.342]	[0.297]	[0.111]	[0.111]	[0.116]	[0.120]
$L_t - L_{t-12}$		-0.022	-0.044**	-0.038*		-0.084***	-0.094***	-0.097***
		[0.014]	[0.015]	[0.019]		[0.016]	[0.025]	[0.026]
$S_t - S_{t-12}$			-0.053	-0.006			-0.015	-0.023
			[0.035]	[0.075]			[0.029]	[0.041]
$C_t - C_{t-12}$				-0.143				0.027
				[0.231]				[0.108]
Dependent Variable: Curvature								
L_t	0.009	0.011	0.012	0.011	-0.000	-0.007	-0.008	-0.008
	[0.006]	[0.007]	[0.006]	[0.006]	[0.005]	[0.005]	[0.005]	[0.005]
S_t	0.074*	0.075*	0.078*	0.064*	0.019	-0.001	-0.002	-0.002
	[0.030]	[0.033]	[0.033]	[0.029]	[0.013]	[0.015]	[0.015]	[0.014]
C_t	0.734***	0.706***	0.706***	0.751***	0.870***	0.885***	0.881***	0.881***
	[0.109]	[0.116]	[0.116]	[0.102]	[0.040]	[0.040]	[0.042]	[0.047]
$L_t - L_{t-12}$		-0.009	-0.012*	-0.009		-0.021**	-0.016	-0.016
		[0.005]	[0.005]	[0.007]		[0.006]	[0.011]	[0.012]
$S_t - S_{t-12}$			-0.008	0.013			0.006	0.007
			[0.012]	[0.025]			[0.013]	[0.018]
$C_t - C_{t-12}$				-0.066				-0.001
				[0.080]				[0.055]
AIC	-9965.7	-9608.4	-9609.5	-9605.8	-6191.2	-6236.7	-6235.5	-6230.7
BIC	-9919.3	-9550.9	-9540.5	-9525.3	-6151.4	-6187.0	-6175.8	-6161.1
Implied β_h	0.51	0.47	0.53	0.53	0.51	0.23	0.22	0.220

Table 4: Estimates of predictive equations for level and slope-mimicking portfolio excess returns This table reports the estimated slope coefficients in equations (5.1) and (5.2) over August 1971-December 2000 and January 2001-December 2017 subsamples. White standard errors, are included. Dependent variables are monthly excess returns on level- and slope-mimicking portfolios. Significance: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

	Pre-2000	Post-2000
Dependent Variable: xrt^{LEVEL}_{t+1}.		
L_t	-0.04	-0.02
	(-1.65)	(-1.14)
S_t	0.01	-1.11***
	(0.13)	(-12.72)
C_t	-0.22	0.353
	(-0.69)	(1.10)
$L_t - L_{t-12}$	0.02	0.06**
	(0.71)	(2.76)
$S_t - S_{t-12}$	0.07	0.03
	(0.82)	(0.32)
$C_t - C_{t-12}$	-0.06	0.17
	(-0.26)	(0.69)
Dependent Variable: xrt^{SLOPE}_{t+1}.		
L_t	-0.02	0.00
	(-1.96)	(0.33)
S_t	-0.00	-1.07***
	(-0.05)	(-26.22)
C_t	0.00	0.31*
	(0.06)	(2.15)
$L_t - L_{t-12}$	0.027	0.12***
	(1.64)	(4.04)
$S_t - S_{t-12}$	-0.03	0.03
	(-1.35)	(0.69)
$C_t - C_{t-12}$	0.01	-0.06
	(0.13)	(-0.47)

Table 5: Sharpe ratios for slope-mimicking portfolios This table reports the annualized Sharpe ratios since 2000 of the strategy of going long (short) the slope-mimicking portfolio if the level fell (rose) over the previous h months and also the strategy of taking a position in the slope-mimicking portfolio that is proportional to $-(L_t - L_{t-h})$, and holding the position from t to $t+1$. The position is rebalanced each month. Annualized Sharpe ratios are computed as the sample average monthly excess returns multiplied by $\sqrt{12}$ and divided by the standard deviation of those monthly excess returns.

Investment in slope-mimicking portfolio	$h = 1$	$h = 3$	$h = 6$	$h = 12$
$2 * 1(L_t - L_{t-h} < 0) - 1$	0.42	0.63	0.47	0.45
$-(L_t - L_{t-h})$	0.53	0.62	0.69	0.65

Table 6: Affine Term Structure Model-Implied coefficients in regression of monthly/yearly changes in 10-year yields on changes in 1-year yields This table reports the slope coefficients in equation (5.9) corresponding to the parameters in an affine term structure model estimated as proposed by Adrian et al. (2013) over August 1971-December 2000 and January 2001-December 2017 subsamples. The term structure model uses K principal components of yields as state variables in panel A, and adds L additional lags of these principal components (for a total of $K + LK$ state variables) in panel B. As memo items the results of the regressions using actual yields are included—these are simply transcribed from Table 1.

		Pre-2000		Post-2000	
		Monthly	Yearly	Monthly	Yearly
Panel A: ATSM with K principal components of yields as factors					
$K = 2$		0.42	0.49	0.83	0.74
$K = 3$		0.46	0.52	0.80	0.59
$K = 4$		0.47	0.52	0.74	0.55
$K = 5$		0.47	0.52	0.71	0.49
Panel B: ATSM with L lags as additional unspanned factors					
$K = 2$	$L = 1$	0.42	0.50	0.83	0.53
$K = 3$	$L = 1$	0.46	0.52	0.80	0.44
$K = 2$	$L = 2$	0.42	0.50	0.85	0.43
$K = 3$	$L = 2$	0.46	0.53	0.81	0.36
$K = 2$	$L = 3$	0.42	0.48	0.85	0.41
$K = 3$	$L = 3$	0.46	0.51	0.81	0.39
Memo: Estimates in data (from Table 1)		0.46	0.56	0.64	0.20

Figure 1: Treasury yield curves around selected episodes This figure plots the Treasury yield curve (dotted line) in the original 2004 "conundrum" episode, the 2008 "conundrum in reverse" episode and the "2017 conundrum". Diamonds display the target for the federal funds rate in each episode.

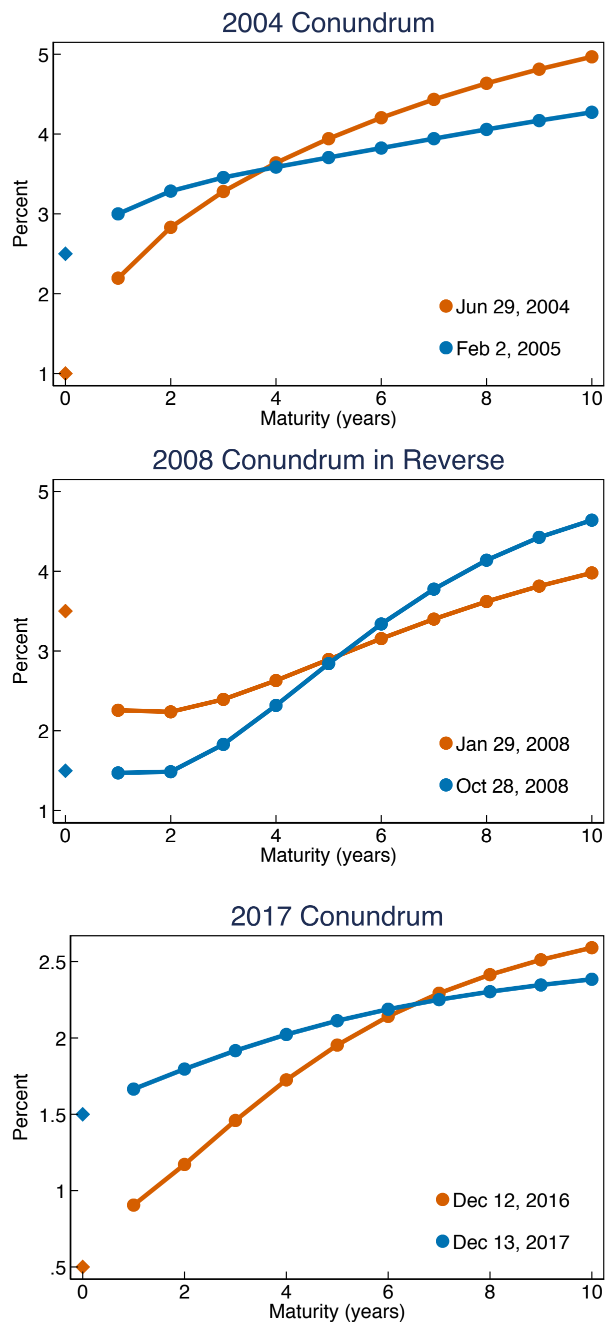


Figure 2: Rolling Regression Estimates of Equations (2.1) and (2.2) This figure plots rolling estimates of the slope coefficients in equations (2.1) and (2.2) with one-year changes (monthly data with h=12) using 10-year rolling windows for estimation. 95% confidence intervals are included (red dashed lines), formed using Newey-West standard errors with a lag truncation parameter of $1.5 \times (h-1)$. Results are plotted against the midpoint of the rolling window.

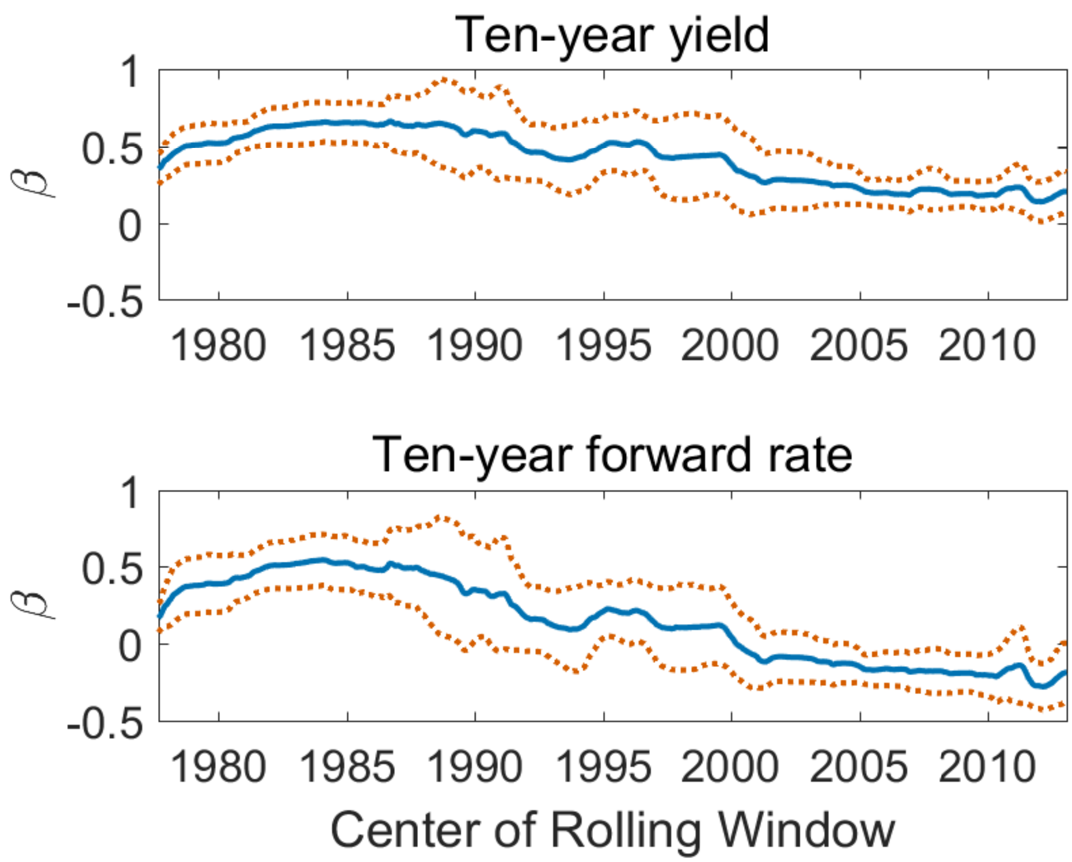


Figure 3: Break Tests for Equations (2.1) and (2.2) This figure plots the Wald test statistic for each possible break date in equations (2.1) and (2.2) with one-year changes (monthly data with h=12) from a fraction 15% of the way through the sample to 85% of the way through the sample. The horizontal red dashed lines denote 10%, 5% and 1% critical values for the maximum of these Wald statistics, from the critical values of Andrews (1993). The overlapping forecasts are accounted for using Newey-West standard errors with a lag truncation parameter of $1.5 \times (h-1)$.

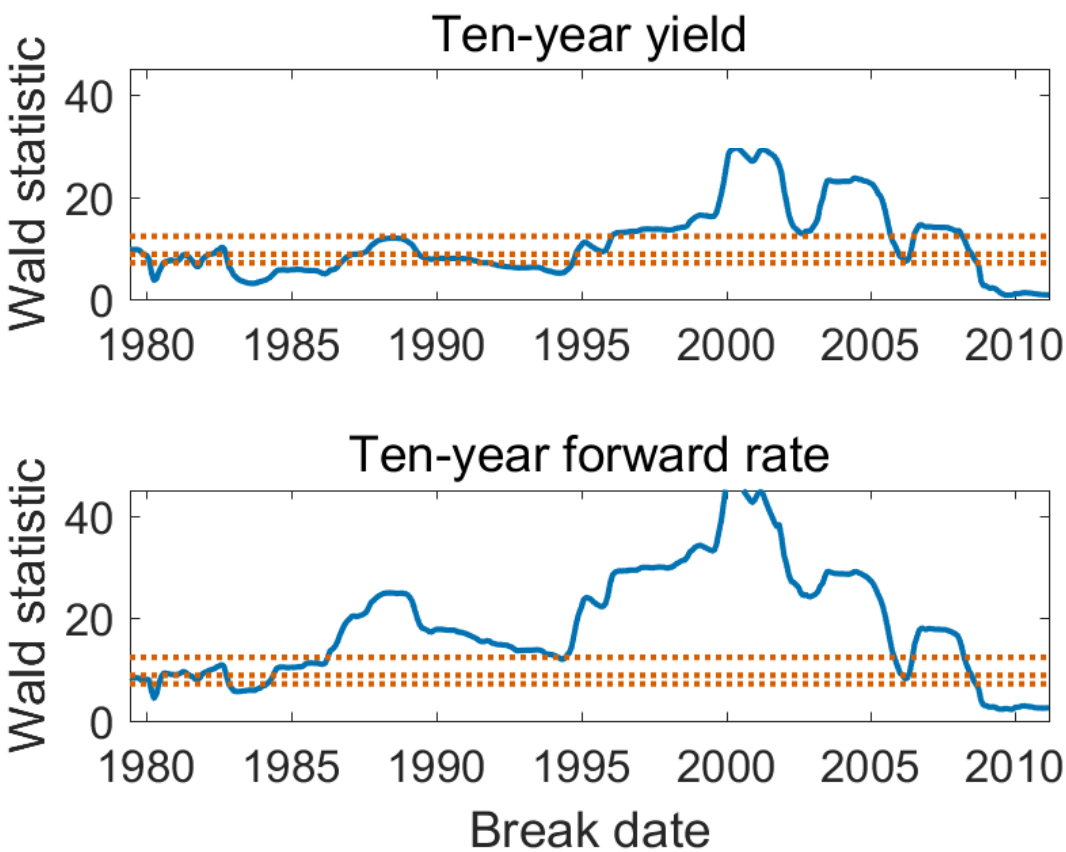


Figure 4: Cross-Correlation of Changes in Level and Slope. Level is the 1-year zero coupon Treasury yield; slope is the 10-year less the 1-year yield. The cross-correlation at lag 0 is the contemporaneous correlation of daily changes in level and slope. Correlations at negative lags (on the left of the figure) denote correlations between monthly changes in level and *future* changes in slope.

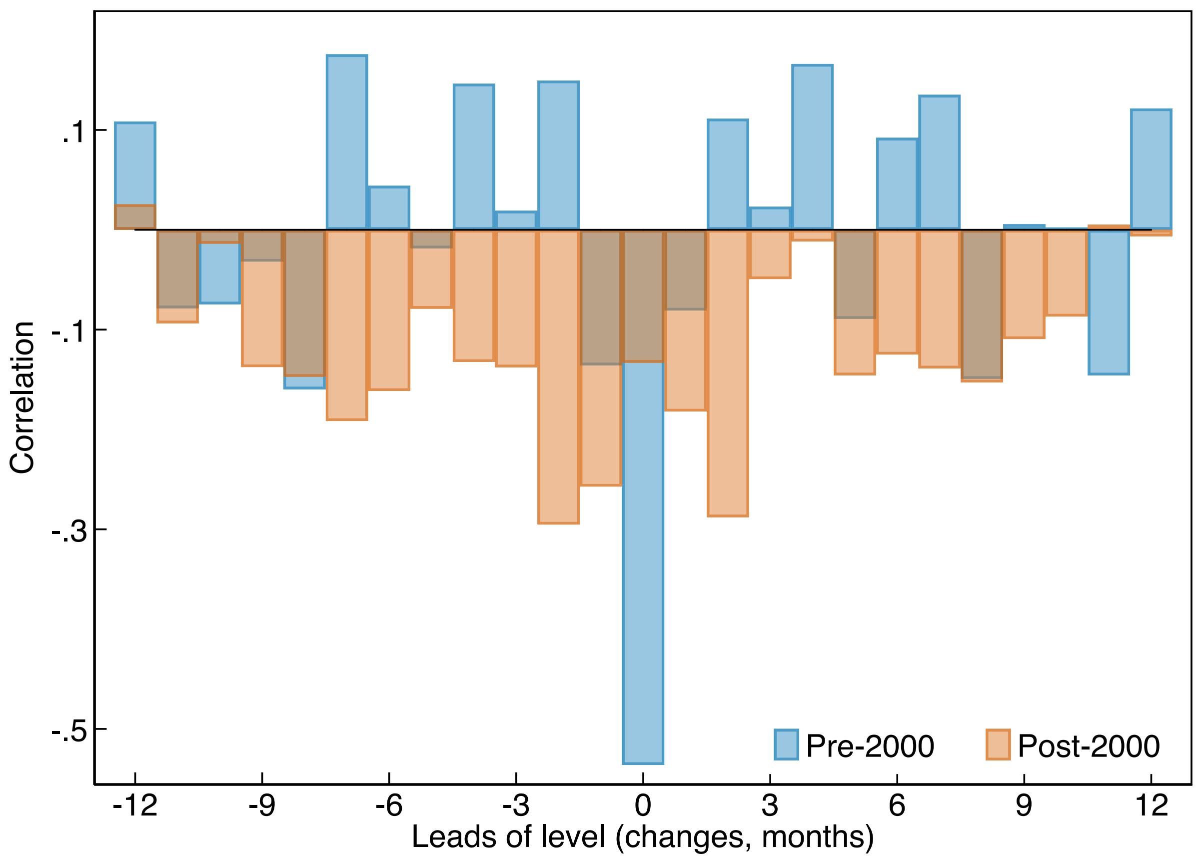


Figure 5: Counterfactual paths of ten-year yields in selected episodes This figure plots one- and ten-year yields in the original "conundrum" period and in 2008, along with counterfactual ten-year yields generated from restricting the slope to depend on lags of level and slope, but not also on lagged changes, as described in the text.

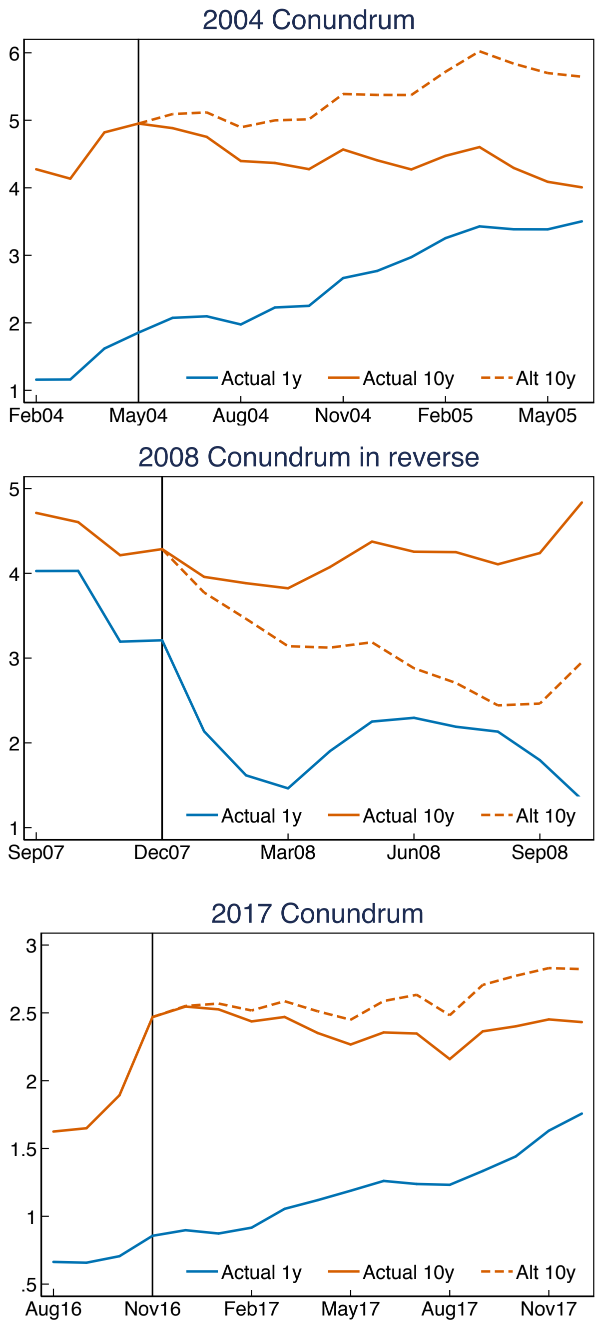


Figure 6: Model-implied coefficients β_h in equation (4.16) against horizon (h) in months. We show this for the baseline calibration and in an alternate calibration where persistence of short rates is greater, and there is lower feedback from short rates to supply. In both cases, β_h is a declining function of h. But the decline is much more pronounced in the baseline calibration.

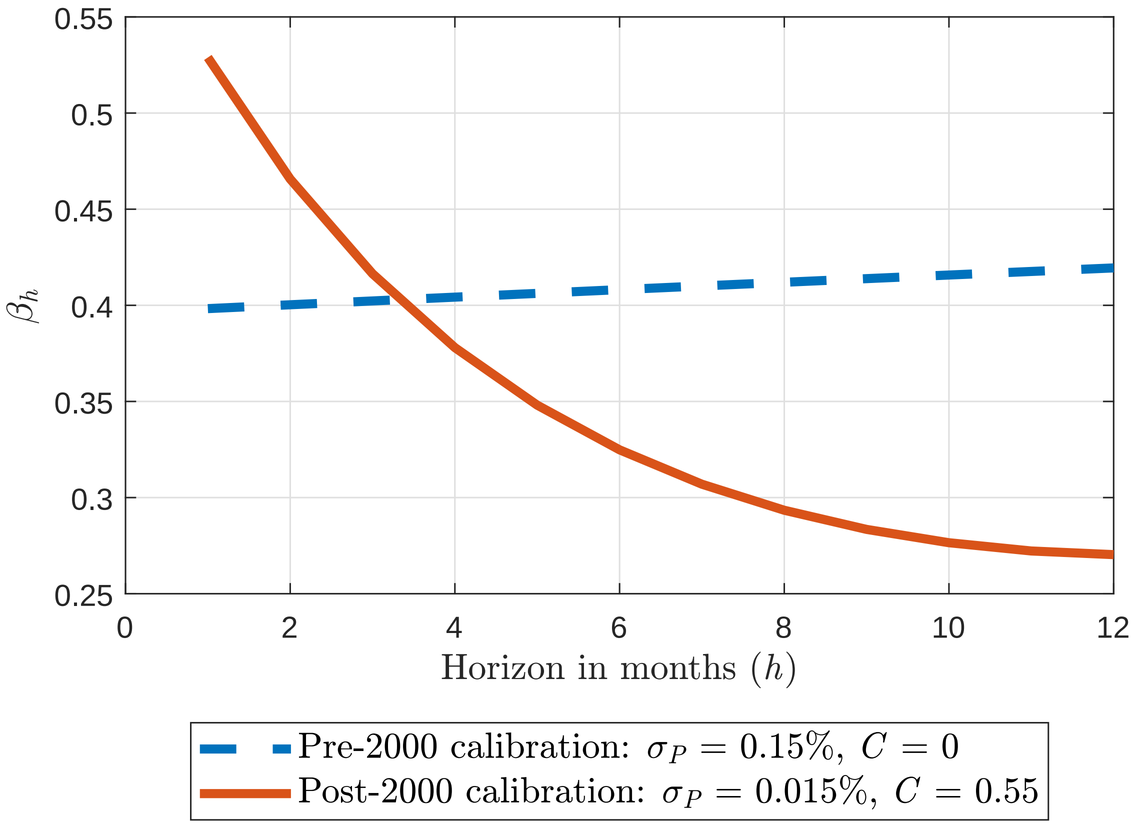
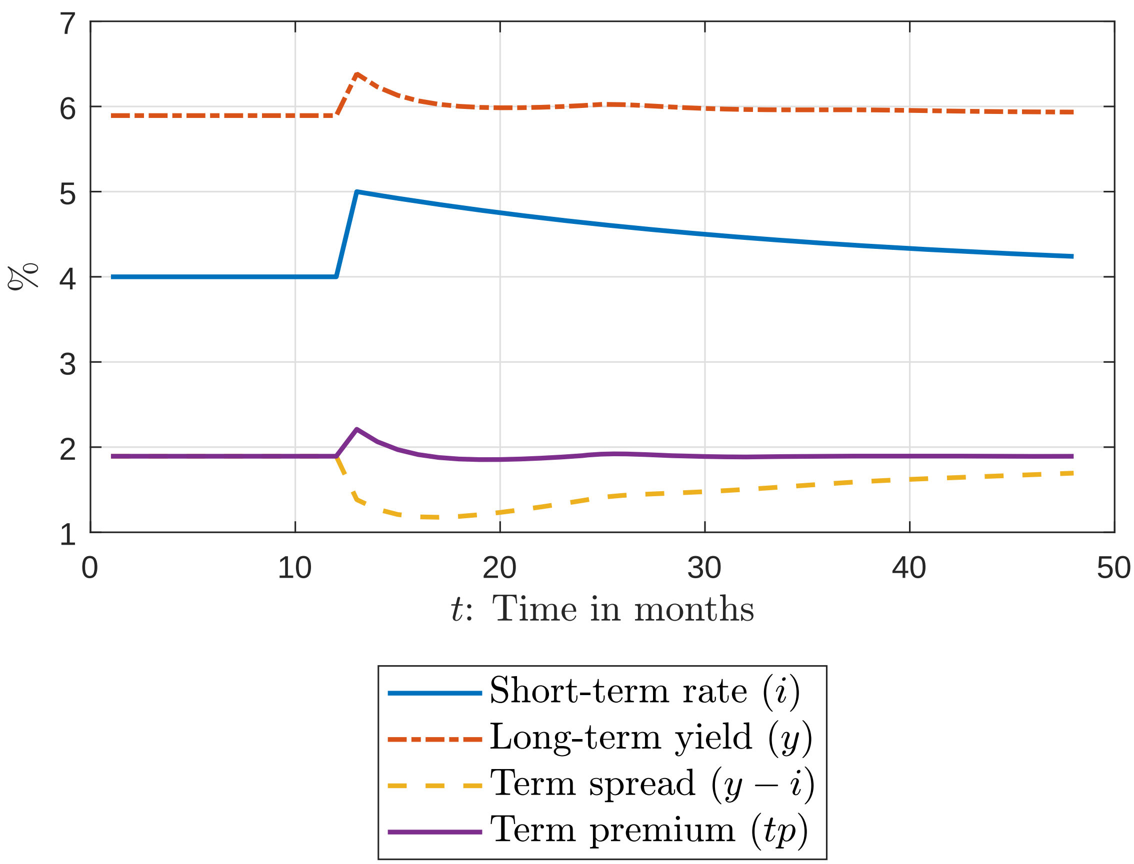


Figure 7: Model-implied impulse response functions over time For the baseline model calibration, we show the response of interest rates following a one-time shock to short-term interest rates.



References

Abrahams, M., T. Adrian, R. K. Crump, E. Moench, and R. Yu (2016): "Decomposing real and nominal yield curves," *Journal of Monetary Economics*, 84, 182–200.

Adrian, T., R. K. Crump, and E. Moench (2013): "Pricing the Term Structure with Linear Regressions," *Journal of Financial Economics*, 110, 110–138.

Andrews, D. W. (1993): "Tests for parameter instability and structural change with unknown change point," *Econometrica*, 61, 821–856.

Backus, D. and J. H. Wright (2007): "Cracking the conundrum," *Brookings Papers on Economic Activity*, 1, 293–316.

Becker, B. and V. Ivashina (2015): "Reaching for yield in the bond market," *Journal of Finance*, 70, 1863–1902.

Beechey, M. J. and J. H. Wright (2009): "The high-frequency impact of news on long-term yields and forward rates: Is it real?" *Journal of Monetary Economics*, 56, 535–544.

Bernanke, B. S. (2010): "Monetary policy and the housing bubble," Speech at the American Economic Assoication annual meeting, Atlanta Georgia.

Campbell, J. R., C. L. Evans, J. D. Fisher, and A. Justiniano (2012): "Macroeconomic effects of Federal Reserve forward guidance," *Brookings Papers on Economic Activity*, 2012, 1–80.

Campbell, J. Y., A. W. Lo, and A. C. MacKinlay (1996): *The Econometrics of Financial Markets*, Princeton, New Jersey: Princeton University Press.

Campbell, J. Y., C. Pflueger, and L. Viceira (2015): "Monetary Policy Drivers of Bond and Equity Risks," NBER Working Paper 20070.

Campbell, J. Y. and R. J. Shiller (1988): "Stock Prices, Earnings and Expected Dividends," *Journal of Finance*, 43, 661–676.

Campbell, J. Y., A. Sunderam, and L. Viceira (2017): "Inflation Bets or Deflation Hedges? The Changing Risks of Nominal Bonds," *Critical Finance Review*, 6, 263–301.

Chow, G. C. (1960): "Tests of Equality Between Sets of Coefficients in Two Linear Regressions," *Econometrica*, 28, 591–605.

Cieslak, A. and P. Povala (2015): "Expected returns in Treasury bonds," *Review of Financial Studies*, 28, 2859–2901.

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

Di Maggio, M. and M. T. Kacperczyk (2017): "The unintended consequences of the zero lower bound policy," *Journal of Financial Economics*, 123, 59–80.

Domanski, D., H. S. Shin, and V. Shushko (2017): "The hunt for duration: Not waving but drowning?" *IMF Economic Review*.

Drechsler, I., A. Savov, and P. Schnabl (2014): "A model of monetary policy and risk premia," Tech. rep., National Bureau of Economic Research.

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

Duffee, G. R. (2013): "Forecasting interest rates," in *Handbook of Economic Forecasting, Volume 2*, ed. by G. Elliott and A. Timmermann, Elsevier.

Duffie, D. (2010): "Asset price dynamics with slow-moving capital," *Journal of Finance*, 65, 1238–1268.

Duffie, D. and R. Kan (1996): "Yield Factor Models of Interest Rates," *Mathematical Finance*, 64, 379–406.

Durham, J. B. (2013): "Momentum and the Term Structure of Interest Rates," Federal Reserve Bank of New York Staff Reports, 657.

Feunou, B. and J.-S. Fontaine (2014): "Non-Markov Gaussian Term Structure Models: The Case of Inflation," *Review of Finance*, 18, 1953–2001.

Gertler, M. and P. Karadi (2015): "Monetary policy surprises, credit costs, and economic activity," *American Economic Journal: Macroeconomics*, 7, 44–76.

Greenwood, R., S. G. Hanson, and G. Y. Liao (2016): "Asset Price Dynamics in Partially Segmented Markets," Working Paper, Harvard University.

Greenwood, R. and D. Vayanos (2014): "Bond supply and excess bond returns," *Review of Financial Studies*, 27, 663–713.

Gürkaynak, R. S., B. Sack, and E. T. Swanson (2005): "The sensitivity of long-term interest rates to economic news: evidence and implications for macroeconomic models," *American Economic Review*, 95, 425–436.

Gürkaynak, R. S., B. Sack, and J. H. Wright (2007): "The U.S. Treasury yield curve: 1961 to the present," *Journal of Monetary Economics*, 54, 2291–2304.

Gürkaynak, R. S., B. Sack, and J. H. Wright (2010): "The TIPS yield curve and inflation compensation," *American Economic Journal: Macroeconomics*, 2, 70–92.

Gurkaynak, R. S., B. P. Sack, and E. T. Swanson (2004): "Do actions speak louder than words? The response of asset prices to monetary policy actions and statements," .

Hanson, S. G. (2014): "Mortgage convexity," *Journal of Financial Economics*, 113, 270–299.

Hanson, S. G. and J. C. Stein (2015): "Monetary policy and long-term real rates," *Journal of Financial Economics*, 115, 429–448.

Joslin, S., A. Le, and K. J. Singleton (2013): "Gaussian macro-finance term structure models with lags," *Journal of Financial Econometrics*, 11, 589–609.

Joslin, S., M. Preibsch, and K. J. Singleton (2014): "Risk premiums in dynamic term structure models with unspanned macro risks," *Journal of Finance*, 69, 1197–1233.

Litterman, R. and J. Scheinkman (1991): "Common factors affecting bond returns," *Journal of Fixed Income*, 1, 54–61.

Lucca, D. O. and F. Trebbi (2009): "Measuring central bank communication: an automated approach with application to FOMC statements," Tech. rep., National Bureau of Economic Research.

Maddaloni, A. and J.-L. Peydró (2011): "Bank risk-taking, securitization, supervision, and low interest rates: Evidence from the euro-area and the US lending standards," *Review of Financial Studies*, 24, 2121–2165.

Malkhozov, A., P. Mueller, A. Vedolin, and G. Venter (2016): "Mortgage risk and the yield curve," *Review of Financial Studies*.

Monfort, A. and F. Pegoraro (2013): "Switching VARMA term structure models," *Journal of Financial Econometrics*, 11, 589–609.

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

Shin, H. S. (2017): "How much should we read into shifts into long-dated yields?" U.S. Monetary Policy Forum, speech.

Stein, J. C. (2013): "Yield-Oriented Investors and the Monetary Transmission Mechanism," .

Swanson, E. T. and J. C. Williams (2014): "Measuring the effect of the zero lower bound on medium- and longer-term interest rates," *American Economic Review*, 104, 3154–3185.

Taylor, J. B. (2010): "The Fed and the crisis: A reply to Ben Bernanke," .

Thornton, D. L. (forthcoming): "Greenspan's conundrum and the Fed's ability to affect long-term yields," *Journal of Money, Credit and Banking*.

Vayanos, D. and J.-L. Vila (2009): "A preferred-habitat model of the term structure of interest rates," Tech. rep., National Bureau of Economic Research.