Special Repo Rates and the Cross-Section of Bond Prices*

Stefania D'Amico† and N. Aaron Pancost‡

This version: May 14, 2018

Abstract

We estimate a dynamic no-arbitrage term structure model that jointly prices the cross-section of Treasury bonds and special repo rates. We show that special repo rates on on-the-run Treasuries can explain almost 80% of the on-the-run premium, but only after incorporating a time-varying risk premium on the special spreads of both on- and off-the-run bonds. We show that the repo risk premium is priced in the cross-section of off-the-run bonds with very low special spreads.

*We thank Darrell Duffie, Jean-Sébastien Fontaine, John Griffin, Arvind Krishnamurthy, Gregor Matvos, Daniel Neuhann, Krista Schwarz (discussant), and seminar participants at the McCombs Finance Brownbag, the 2017 European Winter Meeting of the Econometric Society, and the fifth International Conference on Sovereign Bonds at the Bank of Canada for helpful comments and conversations. All remaining errors are our own. The views expressed here do not reflect official positions of the Federal Reserve.

†Federal Reserve Bank of Chicago. Contact: sdamico@frbchi.org

‡University of Texas at Austin McCombs School of Business. Contact: aaron.pancost@mccombs.utexas.edu

1 Introduction

Most studies in the literature have been modeling and estimating the pricing of U.S. Treasury cash securities and repo contracts separately, even though seminal work by Duffie (1996) and related empirical evidence (e.g., Jordan and Jordan 1997; D'Amico, Fan and Kitsul 2017) have suggested that the pricing in the cash and repo markets are tightly linked. In particular, the approach used so far in dynamic term-structure models (DTSMs) implicitly ignores the possibility that investors might be discounting the stream of future cash-flows of certain Treasury securities at a specific rate, lower than the generic short rate, determined by the value of those securities as collateral in the repo market.[1] In other words, a rational investor, in pricing a Treasury security, would not ignore the current and future expected profits that could be obtained by lending that specific security in the repo market.

This omission, in turn, might have generated price anomalies, such as time-varying on-the-run premiums, which have been investigated for decades (Krishnamurthy 2002) and that perhaps would be less anomalous if the collateral value of Treasury securities was accounted for in pricing those assets.[2] Most likely, this omission in the term-structure literature is due to both the lack of data on special collateral repo rates (i.e., the rate investors are willing to pay to borrow a specific security in the repo market) and the complexity of pricing each Treasury security individually within a DTSM. Using the technology developed in Pancost (2017) to price individual Treasury securities and our proprietary dataset on individual Treasury special repo rates, we estimate the joint term-structure of U.S. Treasury cash and repo rates, derive a risk premium associated to the special collateral value of Treasuries, and study whether this improves our understanding of some of the price anomalies observed in the Treasury cash market.

Specifically, we quantitatively link the on-the-run premium to observed special repo rates in a dynamic, no-arbitrage term structure model. Duffie (1996) shows in a simple static setting how a security's special repo rate that is below the generic short rate of interest (i.e., the general collateral repo rate) implies a higher price for that security in the cash market. However, he does not examine whether observed on-the-run premiums are consistent with special repo rates within

[1] This is nicely illustrated by the formulas in Buraschi and Menini (2002).

[2] Other price anomalies relevant to this study include the Treasury off-the-run note-bond spreads analyzed in Musto, Nini and Schwarz (2017) and Pancost (2017).

a dynamic setting and in the data. Our dynamic model nests Proposition 1 of Duffie (1996) in a setting that allows us to measure the risk premium on special repo rates—one of the contributions of this paper.

We find that special repo rates and the on-the-run premium in the cash market are largely consistent with one another only after the value of the Treasury collateral is explicitly priced as a risk factor in the model, generating time-varying risk premia. A special repo rate below the short rate of interest is equivalent to a convenience yield, or dividend, that accrues to the asset owner. We show that this dividend on on-the-run Treasury bonds varies substantially over time; in other words, it is risky. However, on average, the level of this dividend is not large enough to completely justify the size and persistence of the on-the-run premium in the data. Moreover, in our sample, the on-the-run premium is falling over time. Our model matches the size, persistence, and variability of the on-the-run premium by allowing the risk premium on the special repo risk factors to vary with the state of the economy.

Krishnamurthy (2002) notes that special spreads on the on-the-run 30-year Treasury bond are high, and that this is consistent with a price premium on the on-the-run bond. He quantifies the 30-year on-the-run premium and estimates the profits from trading on it, and finds that they are small. We go one step further by showing the substantial joint time variation in both the special spread and the on-the-run premium, and by quantifying the premium on that risk that makes the two consistent with each other.

The usual practice in the literature on the term structure of interest rates (see Pancost 2017 for a survey) is to exclude on-the-run bonds from the analysis. Gürkaynak, Sack and Wright (2007) exclude not just on-the-run bonds, but also the first off-the-run (i.e., those bonds that were on-the-run just before the latest bond was issued). Many empirical studies of the term structure of interest rates, including D'Amico, Kim and Wei (2010), Hamilton and Wu (2012), Kim and Orphanides (2012), and Bauer and Rudebusch (2014), use the estimated smooth yield curves from Gürkaynak, Sack and Wright (2007) as if they were data. Other studies, including Ang and Piazzesi (2003), Diebold, Rudebusch and Aruoba (2006), Rudebusch and Wu (2008), Christensen, Diebold and Rudebusch (2011), Joslin, Singleton and Zhu (2011), and Creal and Wu (2016), use the bootstrap method of Fama and Bliss (1987) applied to only a small number of bonds, which generally do not include on-the-run bonds. All of these studies explicitly ignore on-the-run bond prices. Unlike

these papers, we consider the prices of on-the-run bonds directly and seek to explain why they are higher than other bonds with similar cash-flows using CUSIP-level special repo rates.

The goal of our paper is to measure the time-varying risk premium that can make prices in the repo and cash markets consistent with one another. This is distinct from the question of why an on-the-run premium exists in the first place. Duffie (1996) conjectures that it may be difficult to find someone willing to trade an off-the-run bond, making the on-the-run bond more liquid and increasing its price. Vayanos and Weill (2008) formalize this intuition in a model of search frictions in which more-liquid securities can trade at a premium even when their promised cash-flows are identical to less-liquid securities. While their model can explain how an on-the-run premium (and high repo spreads) might arise in equilibrium, it does not allow either to vary over time. We conjecture that a dynamic version of their model, perhaps with time-varying issue sizes—for example driven by Treasury re-openings or Federal Reserve purchases of outstanding issues—might generate these effects. This is an exciting area for future research.

The rest of the paper is organized as follows. Section 2 sets up the model. Section 3 describes the data and the estimation. Section 4 presents our empirical results across alternative specifications, which help assess the contribution of time-varying risk premiums on the special repo factors. Section 5 offers concluding remarks.

2 Model

We assume that the prices of Treasury bonds depend on a $k \times 1$ vector X_t that consists of both observable and unobservable (latent) factors, which evolve according to

$$X_{t+1} = \mu + \Phi X_t + \Sigma \varepsilon_{t+1} \tag{1}$$

where the vector of shocks ε_{t+1} are independent of X_t and normally distributed:

$$\varepsilon_{t+1} \sim \mathbb{N}\left(\vec{0}, I\right).$$

The short rate of interest is assumed to be an affine function of the factors:

$$\log(1 + R_t) = \delta_0 + \delta_1' X_t. \tag{2}$$

The stochastc discount factor is given by

$$\log \frac{M_{t+1}}{M_t} = -\delta_0 - \delta_1' X_t - \frac{1}{2}\lambda_t'\lambda_t - \lambda_t'\varepsilon_{t+1}, \tag{3}$$

where

$$\lambda_t \equiv \lambda + \Lambda X_t.$$

These assumptions are standard and lead to simple formulas for zero-coupon bond prices, see for example Ang and Piazzesi (2003).

2.1 Special Repo Rates

A repo contract can be thought of as a collateralized loan, where the repo seller borrows at the repo rate in exchange for a Treasury bond, and regains the bond when she repays the loan plus interest at maturity. Duffie (1996) shows that rates on special collateral can be below the general collateral rate, without creating an arbitrage opportunity, because the supply of special collateral is fixed.

A special repo rate that is below the general collateral rate is a dividend that is proportional to the collateral's current price. Let $1 + R_t$ denote the current (gross) general collateral rate, and $1 + r_t$ the special collateral rate on a particular bond with current price P_t. Assume no haircut for simplicity. At time t, the owner of the special collateral borrows P_t against the collateral, at r_t, and simultaneously lends an amount Δ at the general collateral rate R_t. At time $t + 1$, she earns

$$(1 + R_t)\Delta - (1 + r_t)P_t$$

so that if $\Delta = \frac{1+r_t}{1+R_t} P_t$, she has no gain or loss at $t+1$, and earns

$$P_t - \Delta P_t = \left(1 - \frac{1+r_t}{1+R_t}\right) P_t$$

at time t.

In what follows, it will be convenient to parameterize the log gross special spread as

$$y_t \equiv \log \frac{1+R_t}{1+r_t} \geq 0, \tag{4}$$

which implies that the price of a zero-coupon bond with n periods left to maturity on special and with special spread equal to y_t must have a price given by

$$\begin{aligned} P_t &= \left(1 - e^{-y_t}\right) P_t + E_t^* P_{t+1} \\ &= e^{y_t} E_t^* P_{t+1} \end{aligned}$$

where E_t^* is the risk-neutral expectation. Because special rates are always weakly less than general-collateral rates, the "special dividend" $e^{y_t} \geq 1$ or, equivalently, $y_t \geq 0$.

In order to ensure that y_t is nonnegative for all values of the state, we parameterize it as a quadratic form in the factors:

$$y_t = X_t' \Gamma X_t, \tag{5}$$

where Γ is a symmetric positive semidefinite matrix. This modeling device is commonly used in DTSMs accounting for the zero lower bound on the short rate (e.g., Ahn, Dittmar and Gallant 2002; Kim and Singleton 2012). Equation (5) by construction forces the special spread to be weakly greater than zero for all values of the state vector X_t. This leads to the following proposition for pricing zero-coupon bonds that are on special for their entire life. We show later how we allow for bonds to be on special for only a limited (deterministic) time, for example while they are on-the-run.

Proposition 1. *Consider a zero-coupon bond on special with n periods to maturity, where the repo*

spread is given by equation (5). Then the zero-coupon bond price satisfies

$$\log P_t^{(n)} = A_n + B_n' X_t + X_t' C_n X_t \tag{6}$$

where the A_n, B_n', and C_n loadings are given by

$$\begin{aligned} C_n &= \Gamma + \Phi^{*\prime} C_{n-1} D_{n-1} \Phi^* \\ B_n' &= -\delta_1' + \left(2\mu^{*\prime} C_{n-1} + B_{n-1}'\right) D_{n-1} \Phi^* \\ A_n &= -\delta_0 + A_{n-1} + \frac{1}{2} B_{n-1}' \Sigma G_{n-1} \Sigma' B_{n-1} \\ &\quad + \frac{1}{2} \log |G_{n-1}| + \left(\mu^{*\prime} C_{n-1} + B_{n-1}'\right) D_{n-1} \mu^* \end{aligned} \tag{7}$$

where

$$\begin{aligned} G_{n-1} &= \left[I - 2\Sigma' C_{n-1} \Sigma\right]^{-1} \\ D_{n-1} &= \Sigma G_{n-1} \Sigma^{-1}, \end{aligned}$$

$C_0 = \vec{0}_{k\times k}$, $B_0 = \vec{0}_{k\times 1}$, and $A_0 = 0$, and the risk-neutral parameters μ^* and Φ^* are given by

$$\begin{aligned} \mu^* &\equiv \mu - \Sigma\lambda \\ \Phi^* &\equiv \Phi - \Sigma\Lambda. \end{aligned} \tag{8}$$

Proof. See Appendix A. ∎

The loadings in equation (7) include the loadings in, for example, Ang and Piazzesi (2003) as a special case when $\Gamma = \vec{0}$, since in this case $C_n = \vec{0}$ for all n and therefore $G_n = D_n = I$ for all n. Further, these loadings are usually obtained in quadratic-Gaussian term-structure models (e.g., Kim 2004; Breach, D'Amico and Orphanides 2016).

In the data, special spreads typically accrue to coupon-bearing bonds, which are linear combinations of the zero-coupon bonds priced in Proposition 1. We price coupon bonds by summing

across their coupon payments: the price of bond i at time t is given by

$$\begin{aligned} P_t^i &= \sum_j c_j \exp\left\{A_{m_j} + B'_{m_j} X_t + X_t' C_{m_j} X_t\right\} \\ &\equiv \sum_j c_j P^Z\left(m_j, X_t\right) \\ &\equiv \vec{P}^Z\left\{\vec{c}\left(i\right), \vec{m}\left(i\right), X_t\right\} \end{aligned} \tag{9}$$

where c_j denotes the size of the jth coupon payment, m_j its time to maturity, and the last two lines define notation. The notation of equation (9) includes the repo special spread at time t through the C_{m_j} loadings (which all contain a Γ term). Treasury bonds in our sample pay the same coupon amount every six months; accounting for these coupons, and pricing accrued interest, implies that the price of bond i is given by

$$P_t^i = P^Z\left(\tau_{iM_{it}}, Z\right) + \frac{c_i}{2 \times 100}\left[\sum_{j=1}^{M_{it}} P^Z\left(\tau_{ij}, X_t\right) - \frac{\tau_{i1}}{\xi_{it}} P^Z\left(\tau_{i1}, X_t\right)\right] \tag{10}$$

where τ_{ij} is the time to maturity of coupon j for bond i (including the final coupon paid on the maturity date), c_i is the coupon rate, M_{it} is the number of remaining coupon payments for bond i at time t, and ξ_{it} is the time between the next and previous coupon payment for bond i at time t. The last term in equation (10) accounts for the accrued interest on bond i at time t, which is shared pro rata between the buyer and seller depending on the time remaining to the next coupon payment. Equation (10) describes how the coupon rate c and time to maturity τ of a given bond i in the data translate into the cash-flows $\vec{c}\left(i\right)$ and their maturities $\vec{m}\left(i\right)$ in equation (9).

Stacking all n_t bonds at time t yields the measurement equation

$$\vec{P}_t = \begin{bmatrix} \vec{P}^Z\left\{\vec{c}\left\{1\right\}, \vec{m}\left\{1\right\}, X_t\right\} \\ \vec{P}^Z\left\{\vec{c}\left\{2\right\}, \vec{m}\left\{2\right\}, X_t\right\} \\ \cdots \\ \vec{P}^Z\left\{\vec{c}\left\{n_t\right\}, \vec{m}\left\{n_t\right\}, X_t\right\} \end{bmatrix} + \vec{\eta}_t \tag{11}$$

which, along with equation (1), constitute the state-space system to be estimated. In practice, the

number of bonds n_t in each cross-section is so large relative to the number of factors X_t that the latter can be estimated on each cross-section individually without regard to the state equation (1); see Andreasen and Christensen (2015) for a proof.

3 Data and Estimation

We use daily data on prices of Treasury bonds from CRSP covering the period from January 2, 2009 to December 29, 2017. We are limited to this sample period by the availability of our special collateral repo rate data. However, considering that we are interested in understanding how the collateral value of Treasury securities affect their prices in the cash market, this is a very interesting period, which also makes the estimation of the term structure of Treasury yields quite difficult. In particular, in those years, reduced issuance by the Treasury, the sharp increase in Treasury holdings by the Federal Reserve (Fed), and new financial regulation have reportedly shrunk the availability of Treasury securities, making this high-quality collateral quite scarce in the repo market. This, in turn, might have caused special spreads to be positive also for off-the-run securities and increased off-the-runs fails to delivery to unusual levels (see for example D'Amico, Fan and Kitsul 2017).

Our data covers 2,252 trading days and 628 unique CUSIPs; we drop all bonds with remaining time to maturity of less than one year. We supplement this data with CUSIP-level data on repo rates from a major electronic broker-dealer trading platform; see D'Amico, Fan and Kitsul (2017) for a description of this data. Table 1 provides some descriptive statistics of our data.

[Table 1 about here.]

We highlight three features of our data from Table 1. First, on average there are over 200 bonds per cross-section, so that there is plenty of variation for identifying 3, 4, or 5 pricing factors; second, although off-the-run bonds on average have much lower special spreads (i.e., the difference between the general collateral and the special collateral repo rates) than on-the-run bonds, they are still "on special" relatively often (i.e., they have $y_t > 0$ in equation 5); and third, that the 10-year on-the-run bond has the highest special spread of any mautirty at issuance, and also the highest premium in the cash market. The fact that we observe many more bonds in the cross-section than the number of latent factors allows us to identify the latter without implementing a nonlinear Kalman filter, as

described in the previous section. The other two features of Table 1 warrant more discussion.

Although off-the-run bonds have much lower special spreads on average, they are still very likely to go "on special." The third column of Table 1 shows that on average across all maturities, almost 85% of all off-the-run bonds are trading special in our sample. Moreover, the percentage trading special is uniform across maturities at issuance; the maturity at issuance least likely to trade special, the 30-year, is still on special over 80% of the time. Including these low but non-zero special spreads is important for pricing on-the-run bonds, as we document in section 4.

Table 1 also highlights that the 10-year on-the-run bond behaves differently from other on-the-run securities: it has higher special spreads, and it has a much higher average price premium in our sample than the other maturities. The fifth column of Table 1 reports the average special spread for on-the-run bonds in our sample; this spread is over 35 bps for the 10-year on-the-run bond, which is large given that the average general-collateral repo rate in our sample is 30 bps. The difference implies that on average, holders of the 10-year on-the-run bond are *paid* about 5 bps to borrow cash using their special collateral. Other maturities trade special at lower spreads, about 20 bps on average. The 30-year on-the-run bond trades at a much lower average special spread of about 8 bps.

A higher average special spread is not the only difference between the 10-year on-the-run bond and other on-the-run maturities: in our sample the 10-year bond also features a large and positive on-the-run premium in the cash market. The sixth column of Table 1 reports the average price residual (actual minus fitted) for on-the-run bonds of the indicated maturity from a 3-factor model estimated using only off-the-run bonds and ignoring all special spreads. Gürkaynak, Sack and Wright (2007) omit on-the-run and first-off-the-run bonds from their estimation because these bonds usually trade at a premium; in our sample this premium is about 83 bps of par on average for the 10-year on-the-run bond.

Issuance maturities besides 10 years tend to have much lower cash premia, closer to 13 bps of par; in fact, the 30-year bond on average trades about 13 bps *below* what it "should" according to our model. This finding for the 30-year on-the-run bond is unique to our sample period: before 2009 the 30-year on-the-run bond featured a large and time-varying on-the-run premium, often as large as the 10-year note. However, because we do not have special spreads before 2009, we omit this time period from our analysis in this paper.

Figures 1 and 2 plot the dynamics of the price premia and special spreads for the 10-year and 30-year on-the-run bonds. The top panel of Figure 1 offers a direct comparison between the 10- and 30-year on-the-run cash premia and special spreads over time; results for maturities other than 30 years are similar. The top panel of Figure 1 shows that the 10-year on-the-run bond features a positive, large, and time-varying on-the-run premium in our sample, while the 30-year bond price premium hovers around zero and often becomes negative.

[Figure 1 about here.]

Because our model is most conveniently formulated in terms of coupon bond prices, rather than yields (see equations 9 and 10), in this paper we define the on-the-run premium in prices rather than in yields. This choice is for only for convenience and does not materially affect the results: the bottom panel of Figure 1, which plots three measures of the on-the-run premium for the 10-year bond in yields rather than prices, shows that the yield and price measures are highly correlated. Two of the lines plotted in the bottom panel of Figure 1 report the yield difference between the on-the-run bond, and the yield to maturity on a "synthetic" bond with the exact same cash-flows, estimated using either the six parameters estimated each day by Gürkaynak, Sack and Wright (2007) (blue line) or our own 3-factor model (red line). The Gürkaynak, Sack and Wright (2007) on-the-run premium, a standard measure in the literature (see for example Adrian, Fleming and Vogt 2017), has a correlation coefficient of 0.977 with our 3-factor model-implied measure.

While the two synthetic-bond on-the-run measures plotted in the bottom panel of Figure 1 are very similar, a simpler "2-bond" measure looks completely different. The black line plotted in the bottom panel of Figure 1 reports the on-the-run premium calculated as in Krishnamurthy (2002); that is, as the difference between the yield to maturity on the first off-the-run 10-year bond, and the on-the-run bond. This 2-bond measure is model-free, but as noted by Gürkaynak, Sack and Wright (2007), it will tend to understate the on-the-run premium when the yield curve is upward sloping, because the first off-the-run bond usually has a lower duration than the on-the-run bond. This bias can be so severe the the 2-bond measure even returns an on-the-run *discount*, albeit a small one, as in the bottom panel of Figure 1. In other time periods (not reported), the on-the-run bond usually has a lower yield than the first off-the-run (a positive on-the-run premium), but the premium is generally an order of magnitude smaller than the premium implied by either

synthetic-bond measure.

[Figure 2 about here.]

The top panel of Figure 1 shows that the on-the-run premium on the 10-year bond is large, and muchy larger than the premium for the 30-year bond; Figure 2 shows that the 10-year bond also has alarger special spread than the 30-year bond. In Figure 2 we plot the special repo rates on the 10- and 30-year on-the-run bond over time, as well as the overnight general-collateral rate. Consistent with its low or negative cash premium, the 30-year on-the-run bond has a special rate that sticks very closely to the general-collateral repo rate (its special spread is usually close to zero), although on occasion, especially towards the end of the sample, it does exhibit a high special spread (low or negative repo rate). On the other hand, the 10-year special rate is often substantially lower than the general collateral rate, often dropping as low as -3% (annualized). The goal of our model is to quantitatively link both the level and the riskiness of these special dividends to the cash premium plotted in the top panel of Figure 1.

Because the 10-year on-the-run bond displays the largest price premium in the cash market, and because that bond also has the largest special spreads on average, for the remainder of the paper we consider only the 10-year on-the-run bond as "on-the-run" when separating bonds based on exposure to on-the-run and off-the-run repo factors. In future work we hope to incorporate special spreads and on-the-run premia for other maturities as well.

The Gürkaynak, Sack and Wright (2007) yield curve, coming from a six-factor model, tends to fit the price data better than our 3-factor model. But unlike our model, it is not clear how to extend the Gürkaynak, Sack and Wright (2007) model to incorporate risky special spreads on individual bonds. For the remainder of the paper, we will estimate models that try to "explain" the 10-year price residuals plotted in blue in the top panel of Figure 1.

3.1 Identification

Our model has five factors: the first three are latent factors that govern the short-rate process and the model's fit to bonds without special spreads. Because these factors are unobservable, they must be invariant to translation and rotation; this means that not all elements of Φ^*, μ^*, δ_0, and δ_1 are identifiable. We follow Pancost (2017) in setting the first three elements of μ^* to 0, the first three

elements of δ_1 to 1, and forcing the top-left corner of Φ^* to have the form

$$\Phi^*_{\text{TL}} = \begin{bmatrix} \phi^*_1 & \phi^*_2 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & \phi^*_3 \end{bmatrix} \tag{12}$$

which allows us to estimate the eigenvalues of this corner of Φ^* without imposing that two of them are real.

We model prices of bonds with special spreads by incorporating two observable factors into the model, one for 10-year on-the-run bonds and another for all other ("off-the-run") bonds. Each factor is the average of the square root of the special spread for bonds of that type. In addition, for each individual CUSIP we add a bond-specific, sixth factor with zero prices of risk that incorporates idiosyncratic variation in that bond's special spread relative to the average. This means that the matrix Γ is pinned down as

$$\Gamma = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \end{bmatrix} \tag{13}$$

for off-the-run bonds and

$$\Gamma = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \tag{14}$$

for on-the-run bonds. Equations (13) and (14) imply that the special spreads of off- and on-the-run bonds are given by

$$\begin{aligned} y_t^i &= \left(X_t^{(4)} + x_t^i\right)^2 && \text{(off-the-run)} \\ y_t^i &= \left(X_t^{(5)} + x_t^i\right)^2 && \text{(on-the-run)} \end{aligned} \tag{15}$$

where $X_t^{(j)}$ is the jth element of X_t. Because $X_t^{(4)}$ and $X_t^{(5)}$ are averages of the square root of the observed special spreads y_t^i, equation (15) implies that the bond-specific repo factors x_t^i can be identified as residuals.[3]

Because the x_t^i factors are idiosyncratic, we assume they carry no risk prices, so that their dynamics under the risk-neutral and physical measures are the same. We assume these idiosyncratic factors evolve according to

$$x_{t+1}^i = \rho x_t^i + \sigma_x \varepsilon_{t+1}^i \tag{16}$$

where i indexes individual CUSIPs and ε_{t+1}^i is a standard normal random variable that is iid over time and independent of the aggregate VAR shocks ε_{t+1}. We assume x_t^i is unconditionally mean-zero in order to allow the average repo spreads to be governed by the aggregate repo factors.

Off-the-run bonds have price loadings given by equation (7) with Γ defined in equation (13). On-the-run bonds have their price loadings (denoted with stars) given by equation (7) with Γ defined in equation (14), with one modification: instead of the initial condition $C_0^* = B_0^* = A_0^* = 0$, they have the initial condition $C_m^* = C_m$, $B_m^* = B_m$, and $A_m^* = A_m$, where m is the maturity at which the bond goes off-the-run. In other words, instead of maturity and paying \$1 per unit of face-value, the initial condition for the on-the-run loadings is that at a known maturity m they become an off-the-run bond and inherit the off-the-run loadings. The Treasury issues new bonds of each maturity at regular intervals, so that when a particular on-the-run bond will go off-the-run is known with certainty; therefore, so is its maturity at that date, which is constant for the life of the bond.

[Figure 3 about here.]

[3]Because we are allowing the 10-year on-the-run special-spread to have its "own" factor, x_t^i for this bond will be identically zero.

Figure 3 illustrates how the model prices on- and off-the-run bonds, i.e. bonds that are and are not exposed to the special spread factor $X_t^{(5)}$. The three solid lines are the loadings on the first three latent factors of the model, expressed in yields (i.e. $-B_n'/n$).[4] The dotted lines plot the estimated yield loadings for 10-year on-the-run bonds. For times to maturity of 9-3/4 years and less, the two sets of loadings are identical, and hence solid and dotted lines are indistinguishable. But because on-the-run bonds follow equation (7) with Γ from equation (14) for $n \in [9\ 3/4, 10]$, while off-the-run bonds use Γ from equation (13), for this range of maturities the loadings differ. Figure 3 shows that on-the-run bonds are more exposed to the slope factor, and less exposed to the curvature factor, for the three months that they are on-the-run, than would be a similar security that did not have special spreads driven by the on-the-run special-spread factor. This differential exposure to the level, slope, and curvature factors is in addition to the direct exposure of the bond to the time-varying special dividend inherent in y_t^i.

Because the first three factors are latent, any non-zero values in the top-right corner of Φ^* (the dependence of future latent factors on the repo factors, under the risk-neutral measure) can be rotated away by re-defining the latent factors appropriately. However, the same cannot be said for the lower-left values of Φ^*, thanks to the structure of equations (13) and (14).

Given these assumptions, the total set of risk-neutral parameters to be estimated is given by

parameter	size	number of free parameters
δ_0	1×1	1
$\delta_1 = [1,1,1,0,0,0]'$	6×1	0
$\mu^* = [0,0,0,\mu_4^*,\mu_5^*,0]'$	6×1	2
$\Phi^* = \begin{bmatrix} \Phi_{\text{TL}}^* & \vec{0}_{3\times 2} & \vec{0}_{1\times 3} \\ \Phi_{\text{BL}}^* & \Phi_{\text{BR}}^* & \vec{0}_{1\times 3} \\ \vec{0}_{1\times 3} & \vec{0}_{1\times 2} & \rho \end{bmatrix}$	6×6	$3 + 2\times 3 + 2 \times 2 + 1 = 14$

for a total of 17 free risk-neutral parameters in the full model with 2 repo factors. Of course

[4] Because the identification assumption in equation (12) does not lead to a nice picture, for Figure 3 the risk-neutral parameters have been rotated so that these loadings have the usual level-slope-curvature shapes. In addition, the factors have been scaled to be mean-zero, so that these loadings represent the loadings at the average X_t, i.e. ignoring the effects of the C_n loadings.

simpler models, for example a model without special spreads at all, have fewer parameters, as described in the next section.

Given these risk-neutral parameters, identification proceeds as follows. We recover the fourth and fifth factors $X_t^{(4)}$ and $X_t^{(5)}$ as the average of the square roots of the off- and on-the-run special spreads, respectively. Given these factors, we estimate x_t^i for each CUSIP i using equation (15). We then estimate ρ and σ_x from equation (16) via OLS, pooling across CUSIPs. For each cross-section t, we then estimate the first 3 (latent) factors in X_t to minimize the sum of squared residuals in equation (11).

To estimate risk-neutral parameters we proceed in steps. First we choose the parameters δ_0 and Φ^*_{TL} to minimize the sum of squared pricing residuals on off-the-run bonds. This gives us factors X_t at each date; holding those factors fixed, we search over the remaining parameters in Φ^*_{BL} and Φ^*_{BR} to minimize the sum of squared pricing residuals on on-the-run bonds. We iterate this process to convergence, ensuring that Σ is consistent with its estimate from equation (1) using ordinary least squares on the filtered X_t.[5]

4 Results

In this section we estimate models of increasing complexity to illustrate how incorporating time-varying prices of risk and special spreads on both on- and off-the-run bonds are both necessary to fit the cross-section of bond prices. We estimate 4 models. First, we estimate a standard 3-factor model ignoring special spreads completely; this model has only 4 risk-neutral parameters (δ_0 and the 3 parameters in Φ^*_{TL}). Next, we use these parameters and the estimated dynamics of the 10-year on-the-run special spread to price the special spreads on the 10-year on-the-run bond in a risk-neutral fashion. To do so, we set μ^* and Φ^* such that the fourth element of λ is zero, and the fourth row and column of Λ are all zeroes. λ and Λ are defined in equation (8); we obtain μ and Φ (and Σ) via OLS on the estimated factors X_t.

[Table 2 about here.]

[5] Strictly speaking, because the X_t are estimates, and not observed, we should estimate equation (1) taking the measurement error into account, as done for example by Pancost (2017). However, as shown by Pancost (2017), in practice there are so many bonds in each cross-section that the X_t are measured with sufficient precision that this step has no appreciable effect on the results.

Third, we incorporate time-varying risk premia into the 10-year on-the-run special spread by estimating four parameters in Φ^*_{BL} (3) and Φ^*_{BR} (1) when we add a fourth factor to the model equal to the square root of the 10-year on-the-run special spread. Panel A of Table 2 reports these parameters. Finally, we add a fifth factor equal to the average square root of the special spreads on all other bonds and estimate all parameters of Φ^*_{BL} and Φ^*_{BR}; we report these parameters in Panel B of Table 2.

Before describing results from estimating time-varying risk premia on special spreads, we briefly analyze some features of the 3-factor model ignoring special spreads completely. Figure 4 plots the implied fitting error in level prices at each date across all the off-the-run bonds. The fit is generally good, although it varies over time: in particular the model fits relatively less well at the beginning of the sample, during the height of the financial crisis after the failure of Lehman Brothers. Pancost (2017) examines the pricing residuals of a similar model in detail during this period and in late 2008; he finds that these errors are related to large price differences between bonds older and younger than fifteen years that cannot be explained by their different coupon levels and maturities.

[Figure 4 about here.]

Next, we explore in a reduced-form fashion how much the price fit can be improved by incorporating special spreads. In Table 3 we regress price residuals from the first estimated model (no special spreads at all) on special spreads and past pricing residuals:

$$\eta_{i,t} = \alpha_i + \beta_1 y^i_t + \beta_2 \eta_{i,t-1} + \xi_{i,t},$$

where the $\eta_{i,t}$ are estimated from equation (11) and the 3-factor model. Duffie (1996) shows in a static setting that a security on special should have a higher price than an equivalent security that is not on special, and the price difference should be increasing in specialness. All four columns of Table 3 confirm that the model without special spreads underprices bonds on special, and more so the higher the special spread. This remains the case even when including the lagged price residual, which is a strong predictor of future price residuals.

[Table 3 about here.]

Figure 5 plots the price residuals on the 10-year on-the-run bond for our three estimates incorporating special spreads. The black line is the price residual assuming that the risky special spreads are priced as risk-neutral dividends. It is very similar to the price residual plotted in the top panel of Figure 1. If the profits from short-selling the on-the-run bond and going long a bond with similar cash-flows, for example the first off-the-run bond, were about zero on average, then this line would also hover near zero: the profits from selling the more-expensive bond would be roughly offset by the cost of borrowing that bond (earning a negative interest rate) in the repo market. This is what Krishnamurthy (2002) finds for the 30-year on-the-run bond over his 1995–1999 sample period: the special spread and cash price are roughly consistent with risk-neutral repo specials. In our sample period, the 10-year on-the-run price premium is a good deal higher, and special spreads are not high enough to wipe out the trading profits he considers.

The red and blue lines in Figure 5 plot price residuals for models that allow for the special dividend to command a time-varying risk premium. The red line plots the price residual after allowing for time-varying risk premia on the special spread dividend alone, setting special spreads on all off-the-run bonds to zero. This model assumes that on the 10-year auction day when the current on-the-run bond goes off-the-run, its special spread immediately drops to zero and stays there for the remaining life of the bond. Thus the holders of the 10-year bond are only compensated for repo risk (and earn the special spread dividend) for the first three months of the bond's life. Time-varying risk premia on these three months of risk does a bit to explain the price premium in the early part of the sample, but at the end of the sample the fit is substantially worse. The sawtooth pattern in the red line in Figure 5 corresponds to auction dates; the model without special spreads on off-the-run bonds predicts sharp jumps in the bond prices that don't seem to be in the data.

[Figure 5 about here.]

Finally, the blue line in Figure 5 plots the price residuals after incorporating special spreads, including time-varying risk premia, on off-the-run as well as on-the-run bonds. Although these spreads are usually small (see Table 1), they are important for pricing on-the-run special-spread risk because the 10-year on-the-run bond becomes an off-the-run bond after 3 months. Thus in the 5-factor model the on-the-run bond is exposed to special spread risk for an additional 9 3/4 years,

which as can be seen from the figure is very important in matching the on-the-run price, though there remains variation in the price residual that our model does not capture. In particular, the blue line displays a similar sawtooth pattern to the red line, probably due to the fact that factor loadings are "kinked" on auction dates (see Figure 3). Allowing for time-varying risk premia on the special spreads of off-the-run bonds reduces the kinks, as can be seen in Figure 5, but does not eliminate them. Figure 3 of D'Amico, Fan and Kitsul (2017) shows the dynamics of the special spread for the 10-year note over the auction cycle; we conjecture that incorporating such rich auction-cycle dynamics into the DTSM would eliminate the kinks in Figure 5.

To quantify the amount of variation explained by our estimated models, denote the three time-series vector of residuals plotted in Figure 5 as η_0, η_1, and η_2. We define R^2 of the estimates in the usual way as

$$R_i^2 \equiv 1 - \frac{\eta_i' \eta_i}{\eta_0' \eta_0},$$

for $i \in \{1, 2\}$. With this definition we have $R_1^2 = 0.12$, and $R_2^2 = 0.78$. This implies that the time-varying risk-premia on the special spread risk factors, on both on- and off-the-run bonds, account for about 78% percent of the variation in the on-the-run premium.

[Figure 6 about here.]

We estimate the risk-neutral parameters pertaining to off-the-run bonds, i.e. the fourth rows of μ^* and Φ^* in the five-factor model, only to fit the prices of off-the-run bonds. Given those parameters and factors X_t, we then estimate the fifth row of μ^* and Φ^* to fit the 10-year on-the-run prices. Figure 6 plots the *change* in the standard deviation of off-the-run price residuals η in equation (11) as we move from moving from the 3-factor model with no special spreads, to the 5-factor model. In fact the fit is worse on some days, though overall the fit is marginally better, and usually on the order of around 5 bps of par value. Although economically-speaking this is a small change, coming from a standard deviation on the order of about 50 bps (see Figure 4), these off-the-run special spreads are crucial in fitting the on-the-run bond price. Morevoer they are statistically significant at conventional levels.

5 Conclusion

We estimate a dynamic no-arbitrage term structure model directly on individual Treasury securities, explicitly including securities that feature a large and time-varying on-the-run premium. We link this on-the-run premium directly to time-varying special repo rates on these securities, and show that the two prices can be made largely consistent with each other only after incorporating time-varying risk premia on the special spread risk factors and including the special spreads of off-the-run bonds.

This paper has only begun to scratch the surface of what is possible with the special repo rate data. Price residuals in both Figures 4 and 5 display predictable components, in particular jumps on auctions dates. Figures 2 and 3 of D'Amico, Fan and Kitsul (2017) show that special-spread dynamics are tightly linked to the auction cycle, so it is entirely possible that a model incorporating a richer, auction-centered dynamic process for special spreads could match these cash-price residuals. Moreover, data on the bid/cover ratio is publicly available from the Treasury's website; we speculate that CUSIPs from auctions with higher bid/cover ratios are "hotter" in the specials market, and that this may be an exogenous driver of both their special spread and cash price. We hope in the future to specify a tractable model incorporating such data.

A Proofs

A.1 Proof of Proposition 1

The result follows by induction. First note that a zero-coupon bond pays \$1 at maturity, so that $A_0 = 0$, $B_0' = \vec{0}$, and $C_0 = \vec{0}$ as in equation (7) prices bonds at maturity. Next, fix n and assume that at any time t, the price of an $n-1$ period bond satisfies

$$\log P_t^{(n-1)} = A_{n-1} + B_{n-1}' X_t + X_t' C_{n-1} X_t. \tag{17}$$

It then suffices to show that equation (17) implies equation (7) for bonds with maturity n.

The log price of an n-period zero-coupon bond at time t with special spread y_t is given by

$$\begin{aligned}
\log P_t^{(n)} &= y_t + \log E_t \left\{ \frac{M_{t+1}}{M_t} P_{t+1}^{(n-1)} \right\} \\
&= X_t' \Gamma X_t + \log E_t \left\{ \frac{M_{t+1}}{M_t} \exp\left\{ A_{n-1} + B_{n-1}' Z_{t+1} + Z_{t+1}' C_{n-1} Z_{t+1} \right\} \right\} \\
&= X_t' \Gamma X_t + \log E_t \exp \Bigg\{ -\delta_0 - \delta_1' X_t - \frac{1}{2} \lambda_t' \lambda_t - \lambda_t' \varepsilon_{t+1} \\
&\quad + A_{n-1} + B_{n-1}' \left(\mu + \Phi Z_t + \Sigma \varepsilon_{t+1} \right) \\
&\quad + \left(\mu + \Phi Z_t + \Sigma \varepsilon_{t+1} \right)' C_{n-1} \left(\mu + \Phi Z_t + \Sigma \varepsilon_{t+1} \right) \Bigg\} \\
&= X_t' \Gamma X_t + -\delta_0 - \delta_1' X_t - \frac{1}{2} \lambda_t' \lambda_t + A_{n-1} + B_{n-1}' \left(\mu + \Phi X_t \right) + \left(\mu + \Phi X_t \right)' C_{n-1} \left(\mu + \Phi X_t \right) \\
&\quad + \log E_t \exp \left\{ m' \varepsilon_{t+1} + \varepsilon_{t+1}' \Sigma' C_{n-1} \Sigma \varepsilon_{t+1} \right\}
\end{aligned} \tag{18}$$

where the second line uses equations (5) and (17), the next line plugs in equations (1) and (3), and m in the last line is given by

$$\begin{aligned}
m &\equiv -\lambda_t + \Sigma' B_{n-1} + 2\Sigma' C_{n-1} \left(\mu + \Phi X_t \right) \\
&= -\lambda + \Sigma' B_{n-1} + 2\Sigma' C_{n-1} \mu + \left(2\Sigma' C_{n-1} \Phi - \Lambda \right) X_t \\
&= d_{n-1} + \tilde{D}_{n-1} X_t
\end{aligned}$$

where the last line defines notation. Because ε_{t+1} is a standard multivariate normal random vari-

able, we have that

$$\begin{aligned}\log E_t \exp\left\{m'\varepsilon_{t+1} + \varepsilon'_{t+1}\Sigma' C_{n-1}\Sigma\varepsilon_{t+1}\right\} &= \frac{1}{2}m'G_{n-1}m + \frac{1}{2}\log|G_{n-1}|, \\ &= \frac{1}{2}d'_{n-1}G_{n-1}d_{n-1} + d'_{n-1}G_{n-1}\tilde{D}_{n-1}X_t \\ &\quad + \frac{1}{2}X'_t\tilde{D}'_{n-1}G_{n-1}\tilde{D}_{n-1}X_t + \frac{1}{2}\log|G_{n-1}|\end{aligned} \tag{19}$$

where $G_{n-1} = \left[I - 2\Sigma' C_{n-1}\Sigma\right]^{-1}$ and $|G_{n-1}|$ denotes the determinant of G_{n-1}. Equation (19) holds provided G_{n-1} is positive semi-definite, and can be derived by completing the square.

Plugging equation (19) into equation (18) and combining quadratic, linear, and scalar terms yields the following loadings:

$$\begin{aligned}C_n &= \Gamma - \frac{1}{2}\Lambda'\Lambda + \Phi' C_{n-1}\Phi + \frac{1}{2}\tilde{D}'_{n-1}G_{n-1}\tilde{D}_{n-1} \\ B'_n &= -\delta'_1 - \lambda'\Lambda + B'_{n-1}\Phi + 2\mu' C_{n-1}\Phi + d'_{n-1}G_{n-1}\tilde{D}_{n-1} \\ A_n &= -\delta_0 - \frac{1}{2}\lambda'\lambda + A_{n-1} + B'_{n-1}\mu + \mu' C_{n-1}\mu + \frac{1}{2}\log|G_{n-1}| + \frac{1}{2}d'_{n-1}G_{n-1}d_{n-1}\end{aligned} \tag{20}$$

where

$$\begin{aligned}\tilde{D}_{n-1} &= -\Lambda + 2\Sigma' C_{n-1}\Phi \\ G_{n-1} &= \left[I - 2\Sigma' C_{n-1}\Sigma\right]^{-1} \\ d_{n-1} &= -\lambda + \Sigma' B_{n-1} + 2\Sigma' C_{n-1}\mu\end{aligned}$$

and $C_0 = \vec{0}$, $B_0 = \vec{0}$, and $A_0 = 0$. The remainder of the proof consists of showing that equation (20) is equivalent to equation (7). To do so, I use the fact (proven below in Lemma 1) that the matrix $C_{n-1}D_{n-1} = C_{n-1}\Sigma G_{n-1}\Sigma^{-1}$ is symmetric.

For notational simplicity I drop all the $n-1$ subscripts.

First, write the C_n loadings in equation (20) as

$$\begin{aligned}
C_n &= \Gamma - \frac{1}{2}\Lambda'\Lambda + \Phi' C \Phi + \frac{1}{2}\Lambda' G \Lambda \\
&\quad + 2\Phi' C \Sigma G \Sigma' C \Phi - \Lambda' G \Sigma' C \Phi - \Phi' C \Sigma G \Lambda \\
&= \Gamma - \frac{1}{2}\Lambda G G^{-1} \Lambda + \frac{1}{2}\Lambda' G \Lambda - \Lambda' G \Sigma' C \Phi \\
&\quad + \Phi' C \Phi + 2\Phi' C \Sigma G \Sigma' C \Phi - \Phi' C \Sigma G \Lambda \\
&= \Gamma + \frac{1}{2}\Lambda' G \left(-G^{-1} + I\right) \Lambda - \Lambda' G \Sigma' C \Phi \\
&\quad + \Phi' C \Sigma G \Sigma' \left(J^{-1} + 2C\right) \Phi - \Phi' C \Sigma G \Lambda,
\end{aligned}$$

where $J \equiv \Sigma G \Sigma'$ so that

$$\begin{aligned}
J^{-1} &= \Sigma'^{-1} G^{-1} \Sigma^{-1} \\
&= \Sigma\prime - 1 \left(I - 2\Sigma' C \Sigma\right) \Sigma^{-1} \\
&= \Sigma'^{-1} \Sigma^{-1} - 2C.
\end{aligned} \tag{21}$$

Plugging in equation (21) and the fact that $G^{-1} = I - 2\Sigma' C \Sigma$ and rearranging yields

$$\begin{aligned}
C_n &= \Gamma + \Lambda' G \Sigma' C \Sigma \Lambda - \Lambda' G \Sigma' C \Phi \\
&\quad + \Phi' C \Sigma G \Sigma''^{-1} \Sigma^{-1} \Phi - \Phi' C \Sigma G \Lambda \\
&= \Gamma + \Lambda' G \Sigma' C \left(\Sigma \Lambda - \Phi\right) \\
&\quad + \Phi' C \Sigma G \left(\Sigma^{-1} \Phi - \Lambda\right) \\
&= \Gamma - \Lambda' G \Sigma'^{*} + \Phi'^{-1} \Phi^{*} \\
&= \Gamma + \left(\Phi'^{-1} - \Lambda' \Sigma''^{-1} G \Sigma' C\right) \Phi^{*} \\
&= \Gamma + \left(\Phi' C D - \Lambda' \Sigma' \left(C D\right)'\right) \Phi^{*},
\end{aligned}$$

where $H \equiv C \Sigma G \Sigma^{-1}$. Then applying Lemma 1 gives the C_n loadings in equation (7).

The B_n loadings in equation (20) are given by

$$\begin{aligned}
B_n' &= -\delta_1' - \lambda'\Lambda + B'\Phi + 2\mu' C\Phi \\
&\quad + \lambda' G\Lambda - 2\lambda' G\Sigma' C\Phi - B'\Sigma G\Lambda \\
&\quad + 2B'\Sigma G\Sigma' C\Phi - 2\mu' C\Sigma G\Lambda + 4\mu' C\Sigma G\Sigma' C\Phi \\
&= -\delta_1' - \lambda'\Lambda + \lambda' G\Lambda - 2\lambda' G\Sigma' C\Phi \\
&\quad + B'\Phi - B'\Sigma G\Lambda + 2B'\Sigma G\Sigma' C\Phi \\
&\quad + 2\mu' C\Phi - 2\mu' C\Sigma G\Lambda + 4\mu' C\Sigma G\Sigma' C\Phi \\
&= -\delta_1' + \lambda' G\left(-G^{-1} + I\right)\Lambda - 2\lambda' G\Sigma' C\Phi \\
&\quad + B'\Sigma G\left(G^{-1}\Sigma^{-1}\Phi - \Lambda\right) + 2B'\Sigma G\Sigma' C\Phi \\
&\quad + 2\mu' C\Sigma G\left(G^{-1}\Sigma^{-1}\Phi - \Lambda + 2\Sigma' C\Phi\right) \\
&= -\delta_1' + 2\lambda' G\Sigma' C\left(\Sigma\Lambda - \Phi\right) \\
&\quad + B'\Sigma G\left(\left(\Sigma^{-1} - 2\Sigma' C\right)\Phi - \Sigma^{-1}\Sigma\Lambda\right) + 2B'\Sigma G\Sigma' C\Phi \\
&\quad + 2\mu' C\Sigma G\left(\left(\Sigma^{-1} - 2\Sigma' C\right)\Phi - \Lambda + 2\Sigma' C\Phi\right) \\
&= -\delta_1' - 2\lambda' G\Sigma'^{*} \\
&\quad + B'^{-1}\Phi^{*} \\
&\quad + 2\mu'^{-1}\Phi^{*},
\end{aligned}$$

using $G^{-1} = I - 2\Sigma' C\Sigma$. Further combining terms and applying Lemma 1 yields

$$\begin{aligned}
B_n' &= -\delta_1' + \left(-2\lambda' G\Sigma' C + 2\mu'^{-1} + B'^{-1}\right)\Phi^{*} \\
&= -\delta_1' + \left(2\left(\mu' + \lambda'\Sigma'\right) C\Sigma G\Sigma^{-1} + B'^{-1}\right)\Phi^{*} \\
&= -\delta_1' + \left(2\mu'^{*} CD + B' D\right)\Phi^{*}
\end{aligned}$$

where the second line uses the implication from Lemma 1 that $G\Sigma' C = \Sigma'^{-1}$.

The A_n loadings in equation (20) are given by

$$\begin{aligned}
A_n &= -\delta_0 + A - \frac{1}{2}\lambda'\lambda + B'\mu + \mu' C\mu + \frac{1}{2}\log|G| \\
&\quad + \frac{1}{2}\lambda' G\lambda + \frac{1}{2}B'\Sigma G\Sigma' B + 2\mu' C\Sigma G\Sigma' C\mu \\
&\quad - \lambda' G\Sigma' B - 2\lambda' G\Sigma' C\mu + 2\mu' C\Sigma G\Sigma' B \\
&= -\delta_0 + A + \frac{1}{2}\log|G| + \frac{1}{2}B'\Sigma G\Sigma' B \\
&\quad - \frac{1}{2}\lambda'^{-1}\lambda + \frac{1}{2}\lambda' G\lambda - 2\lambda' G\Sigma' C\mu \\
&\quad + B'\mu - \lambda' G\Sigma' B + 2\mu' C\Sigma G\Sigma' B \\
&\quad + \mu' C\mu + 2\mu' C\Sigma G\Sigma' C\mu \\
&= -\delta_0 + A + \frac{1}{2}\log|G| + \frac{1}{2}B'\Sigma G\Sigma' B \\
&\quad + \frac{1}{2}\lambda' G\left(-G^{-1} + I\right)\lambda - 2\lambda' G\Sigma' C\mu \\
&\quad + B'\mu - B'\Sigma G\lambda + 2B'\Sigma G\Sigma' C\mu \\
&\quad + \mu' CJ\left(J^{-1} + 2C\right)\mu \\
&= -\delta_0 + A + \frac{1}{2}\log|G| + \frac{1}{2}B'\Sigma G\Sigma' B \\
&\quad + \lambda' G\Sigma' C\left(\Sigma\lambda - 2\mu\right) \\
&\quad + B'\Sigma G\left(G^{-1}\Sigma^{-1}\mu - \lambda\right) + 2B'\Sigma G\Sigma' C\mu \\
&\quad + \mu'^{-1}\mu
\end{aligned}$$

where again $J \equiv \Sigma G\Sigma'$, so that the last line applies equation (21). Using the fact that $G^{-1}\Sigma^{-1} =$

$\Sigma^{-1} - 2\Sigma' C$, we have that

$$\begin{aligned}
A_n &= -\delta_0 + A + \frac{1}{2}\log|G| + \frac{1}{2}B'\Sigma G\Sigma' B \\
&\quad + \lambda' G\Sigma' C\left(\Sigma\lambda - 2\mu\right) + \mu'^{-1}\mu \\
&\quad + B'\Sigma G\left(\left(\Sigma^{-1} - 2\Sigma' C\right)\mu - 2\lambda\right) + 2B'\Sigma G\Sigma' C\mu \\
&= -\delta_0 + A + \frac{1}{2}\log|G| + \frac{1}{2}B'\Sigma G\Sigma' B \\
&\quad + \lambda' G\Sigma' C\left(\Sigma\lambda - 2\mu\right) + \mu'^{-1}\mu \\
&\quad + B'^{-1}\left(\mu - \Sigma\lambda\right) - 2B'\Sigma G\Sigma' C\mu + 2B'\Sigma G\Sigma' C\mu \\
&= -\delta_0 + A + \frac{1}{2}\log|G| + \frac{1}{2}B'\Sigma G\Sigma' B + B'^{-1}\mu^* \\
&\quad + \lambda'\Sigma'^{-1}\left(\Sigma\lambda - 2\mu\right) + \mu'^{-1}\mu
\end{aligned}$$

where the last line applies Lemma 1. Rearranging terms and again applying Lemma 1 yields

$$\begin{aligned}
A_n &= -\delta_0 + A + \frac{1}{2}\log|G| + \frac{1}{2}B'\Sigma G\Sigma' B + B'^{-1}\mu^* \\
&\quad + \lambda'\Sigma'^{-1}\left(\Sigma\lambda - \mu\right) - \lambda'\Sigma'^{-1}\mu + \mu'^{-1}\mu \\
&= -\delta_0 + A + \frac{1}{2}\log|G| + \frac{1}{2}B'\Sigma G\Sigma' B + B'^{-1}\mu^* \\
&\quad - \lambda'\Sigma'^{-1}\left(\mu - \Sigma\lambda\right) + \left(\mu' - \lambda'\Sigma'\right)C\Sigma G\Sigma^{-1}\mu \\
&= -\delta_0 + A + \frac{1}{2}\log|G| + \frac{1}{2}B'\Sigma G\Sigma' B + B'^{-1}\mu^* \\
&\quad + \left(\mu' - \lambda'\Sigma'\right)C\Sigma G\Sigma^{-1}\left(\mu - \Sigma\lambda\right)
\end{aligned}$$

which is equation (7).

Lemma 1. *The matrix $C_{n-1}D_{n-1} = C_{n-1}\Sigma G_{n-1}\Sigma^{-1}$ is symmetric for all n.*

Proof. Using equation (20), so long as Γ is symmetric, then C_{n-1} and G_{n-1} are both symmetric for all n. In what follows I drop the $n-1$ subscripts. Let $H \equiv C\Sigma G\Sigma^{-1}$, so that I need to show that $H = H' = \Sigma'^{-1}G\Sigma' C$.

By definition, $G^{-1} = 1 - 2\Sigma' C\Sigma$, so that

$$\begin{aligned}
\Sigma G^{-1}\Sigma^{-1} &= \Sigma\left(1 - 2\Sigma' C\Sigma\right)\Sigma^{-1} \\
&= I - 2\Sigma\Sigma' C \qquad (22) \\
\Sigma'^{-1} G^{-1}\Sigma' &= \Sigma'^{-1}\left(I - 2\Sigma' C\Sigma\right)\Sigma' \\
&= I - 2C\Sigma\Sigma'. \qquad (23)
\end{aligned}$$

Then we have that

$$\begin{aligned}
H &= C\Sigma G\Sigma^{-1} \\
&= \underbrace{\left(\Sigma'^{-1} G\Sigma'\right)\left(\Sigma'^{-1} G^{-1}\Sigma'\right)}_{=I} C\Sigma G\Sigma^{-1} \\
&= \left(\Sigma'^{-1} G\Sigma'\right)\underbrace{\left(I - 2C\Sigma\Sigma'\right)}_{\text{by equation (23)}} C\Sigma G\Sigma^{-1} \\
&= \left(\Sigma'^{-1} G\Sigma'\right)\left(C - 2C\Sigma\Sigma' C\right)\Sigma G\Sigma^{-1} \\
&= \left(\Sigma'^{-1} G\Sigma' C\right)\left(I - 2\Sigma\Sigma' C\right)\Sigma G\Sigma^{-1} \\
&= H'\underbrace{\left(\Sigma G^{-1}\Sigma^{-1}\right)}_{\text{by equation (22)}}\Sigma G\Sigma^{-1} \\
&= H'.
\end{aligned}$$

∎

References

Adrian, Tobias, Michael Fleming, and Erik Vogt. 2017. "An Index of Treasury Market Liquidity: 1991–2017." *Federal Reserve Bank of New York Staff Report 827*. 11

Ahn, Dong-Hyun, Robert F. Dittmar, and A. Ronald Gallant. 2002. "Quadratic Term Structure Models: Theory And Evidence." *The Review of Financial Studies*, 15(1): 243–288. 6

Andreasen, Martin M., and Bent Jesper Christensen. 2015. "The SR Approach: A New Estimation Procedure For Non-Linear And Non-Gaussian Dynamic Term Structure Models." *Journal of Econometrics*, 184(2): 420–451. 9

Ang, Andrew, and Monika Piazzesi. 2003. "A No-Arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables." *Journal of Monetary Economics*, 50(4): 745–787. 3, 5, 7

Bauer, Michael D., and Glenn D. Rudebusch. 2014. "The Signaling Channel For Federal Reserve Bond Purchases." *International Journal of Central Banking*, 10(3): 233–289. 3

Breach, Tomas, Stefania D'Amico, and Athanasios Orphanides. 2016. "The Term Structure And Inflation Uncertainty." Federal Reserve Bank of Chicago Working Paper Series WP-2016-22. 7

Buraschi, Andrea, and Davide Menini. 2002. "Liquidity Risk And Specialness." *Journal of Financial Economics*, 64(2): 243–284. 2

Christensen, Jens H.E., Francis X. Diebold, and Glenn D. Rudebusch. 2011. "The Affine Arbitrage-Free Class Of Nelson?OOsiegel Term Structure Models." *Journal of Econometrics*, 164(1): 4–20. 3

Creal, Drew D, and Jing Cynthia Wu. 2016. "Bond Risk Premia in Consumption-based Models." Working paper. 3

D'Amico, Stefania, Don H. Kim, and Min Wei. 2010. "Tips from TIPS: the Informational Content of Treasury Inflation-Protected Security Prices." *Federal Reserve Board Finance and Economics Discussion Series*, 2010–19. 3

D'Amico, Stefania, Roger Fan, and Yuriy Kitsul. 2017. "The Scarcity Value Of Treasury Collateral: Repo Market Effects Of Security-Specific Supply And Demand Factors." Forthcoming, *Journal of Financial and Quantitative Analysis*. 2, 9, 19, 20

Diebold, Francis X., Glenn D. Rudebusch, and S. Borağan Aruoba. 2006. "The Macroeconomy and the Yield Curve: a Dynamic Latent Factor Approach." *Journal of Econometrics*, 131(1-2): 309–338. 3

Duffie, Darrell. 1996. "Special Repo Rates." *The Journal of Finance*, 51(2): 493–526. 2, 3, 4, 5, 17

Fama, Eugene F., and Robert R. Bliss. 1987. "The Information in Long-Maturity Forward Rates." *The American Economic Review*, 77(4): 680–692. 3

Gürkaynak, Refet S., Brian Sack, and Jonathan H. Wright. 2007. "The U.S. Treasury Yield Curve: 1961 to the Present." *Journal of Monetary Economics*, 54(8): 2291–2304. 3, 10, 11, 12, 31

Hamilton, James D., and Jing Cynthia Wu. 2012. "The Effectiveness Of Alternative Monetary Policy Tools In A Zero Lower Bound Environment." *Journal of Money, Credit and Banking*, 44: 3–46. 3

Jordan, Bradford D., and Susan D. Jordan. 1997. "Special Repo Rates: An Empirical Analysis." *The Journal of Finance*, 52(5): 2051–2072. 2

Joslin, Scott, Kenneth J. Singleton, and Haoxiang Zhu. 2011. "A New Perspective On Gaussian Dynamic Term Structure Models." *Review of Financial Studies*, 24(3): 926–970. 3

Kim, Don H. 2004. "Time-varying risk and return in the quadratic-gaussian model of the term-structure." Ph.D. Dissertation, Stanford University. 7

Kim, Don H., and Athanasios Orphanides. 2012. "Term Structure Estimation With Survey Data On Interest Rate Forecasts." *Journal of Financial and Quantitative Analysis*, 47(1): 241–272. 3

Kim, Don H., and Kenneth J. Singleton. 2012. "Term Structure Models And The Zero Bound: An Empirical Investigation Of Japanese Yields." *Journal of Econometrics*, 170(1): 32–49. 6

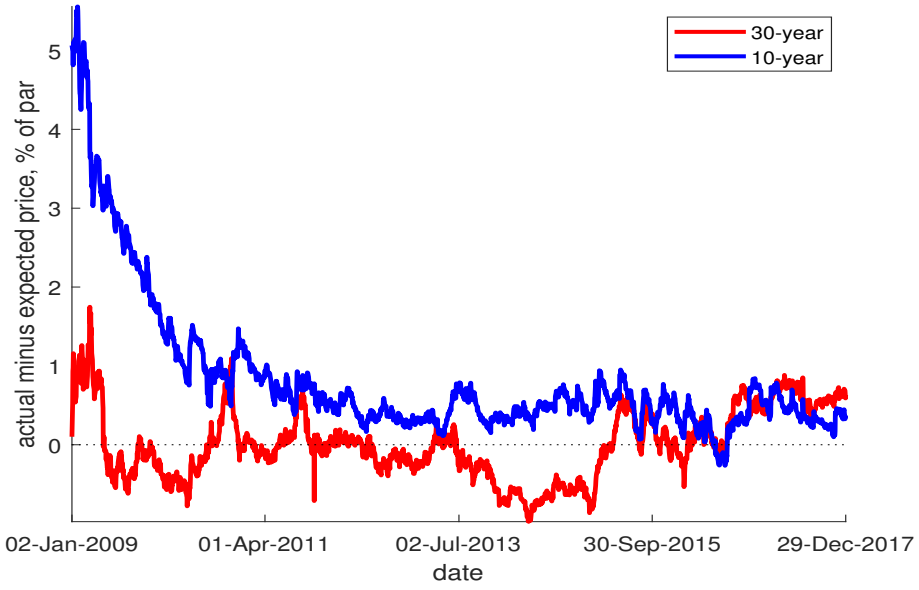
Krishnamurthy, Arvind. 2002. "The Bond/Old-Bond Spread." *Journal of Financial Economics*, 66(2–3): 463–506. 2, 3, 11, 18

Musto, David, Greg Nini, and Krista Schwarz. 2017. "Notes on Bonds: Illiquidity Feedback During the Financial Crisis." Working Paper. 2

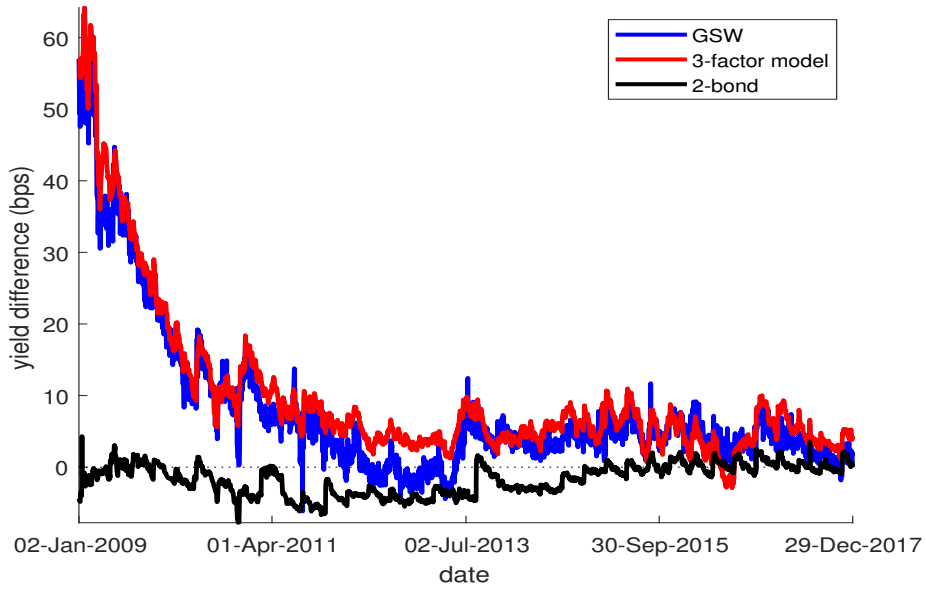
Pancost, N. Aaron. 2017. "Zero-Coupon Yields and the Cross-section of Bond Prices." Working Paper, University of Chicago. 2, 3, 12, 16, 17

Rudebusch, Glenn D., and Tao Wu. 2008. "A Macro-Finance Model of the Term Structure, Monetary Policy and the Economy." *The Economic Journal*, 118(530): 906–926. 3

Vayanos, Dimitri, and Pierre-Olivier Weill. 2008. "A Search-Based Theory Of The On-The-Run Phenomenon." *The Journal of Finance*, 63(3): 1361–1398. 4



(a) On-the-Run Price Premia, 10- and 30-year



(b) On-the-Run Yield Premia, 10-year

Figure 1. On-the-Run Premia and Special Rates
The top panel plots price residuals for the 30- and 10-year on-the-run bonds over time, where the residual is defined in equation (11) as the actual price minus the model-implied price in % of par value, and the parameters are given in Table 2, Panel A. The factors for these price residuals are estimated only on off-the-run bonds. The bottom panel plots the difference between the yields to maturity of the 10-year on-the-run bond and the 10-year first off-the-run bond (black line), a synthetic bond with the same cash-flows as the 10-year on-the-run bond, according to the Gürkaynak, Sack and Wright (2007) estimated yield curve (blue line), and according to our esitmated 3-factor model (red line). Each line subtracts the on-the-run yield from the alternative yield.

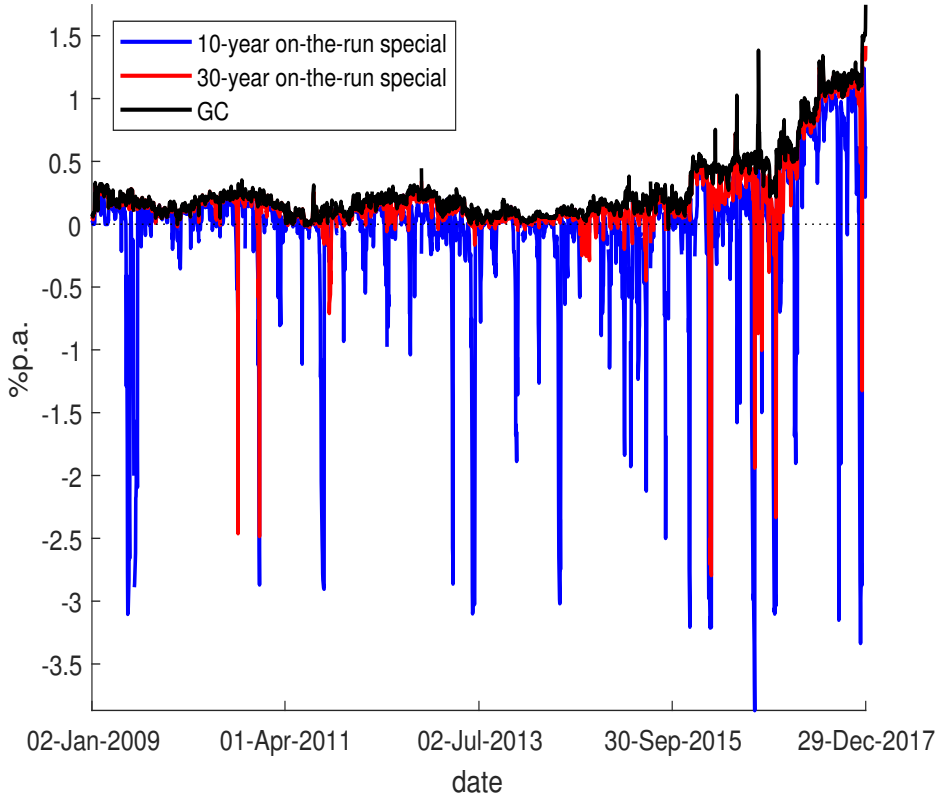


Figure 2. On-the-Run Special Rates
The figure plots the GC repo (black line) rate along with the special rates for the 30- and 10-year on-the-run bonds (blue and red lines, respectively) over time, in annualized percent.

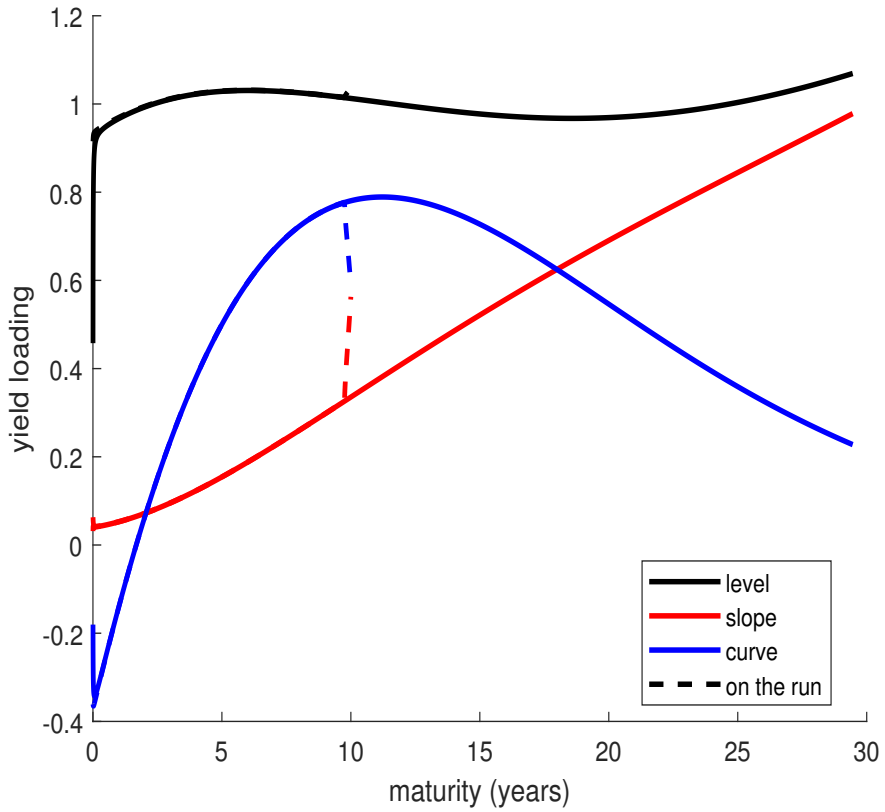


Figure 3. The figure plots the yield loadings $-\frac{1}{n}B_n'$ as a function of time to maturity n in years for the 5-factor model whose parameters are reported in Panel B of Table 2. Only the latent-factor loadings are plotted (the first 3). The loadings are translated so that the factors are mean-zero and rotated so that these loadings have the traditional level-slope-curvature shape. The solid lines are loadings for off-the-run bonds; the dotted lines plot the loadings for 10-year on-the-run bonds.

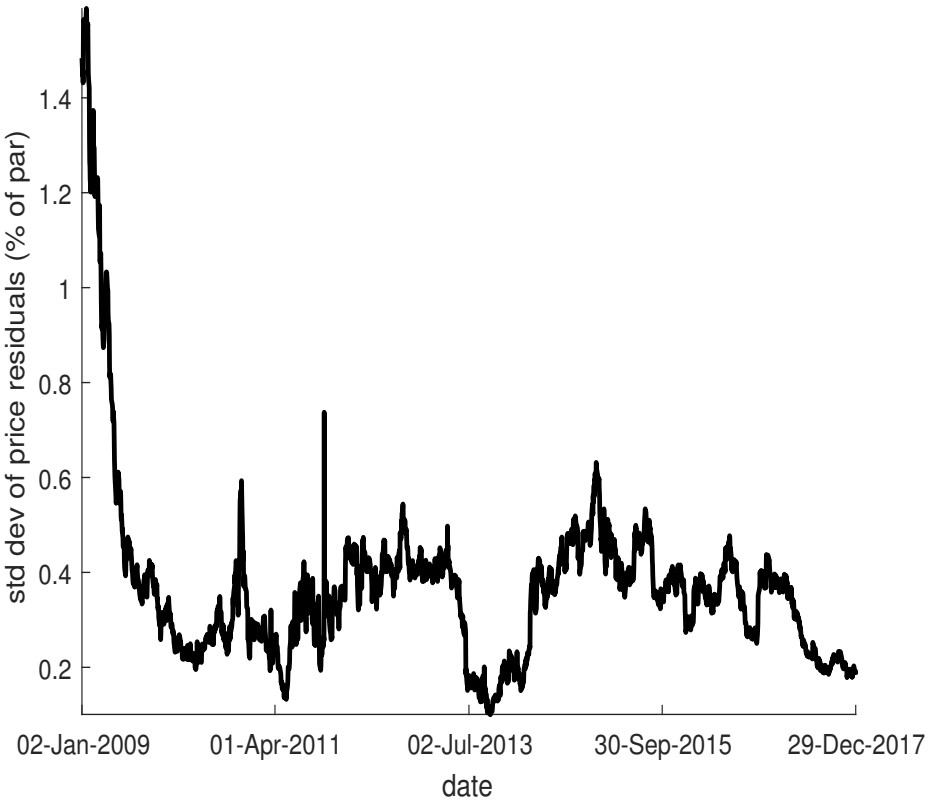


Figure 4. The figure plots the the standard deviation of the residuals from equation (11) for each date in the sample, estimated only on off-the-run bonds and setting all special spreads to zero. The parameters of this 3-factor model are reported in Panel A of Table 2. The values are reported in percent of par.

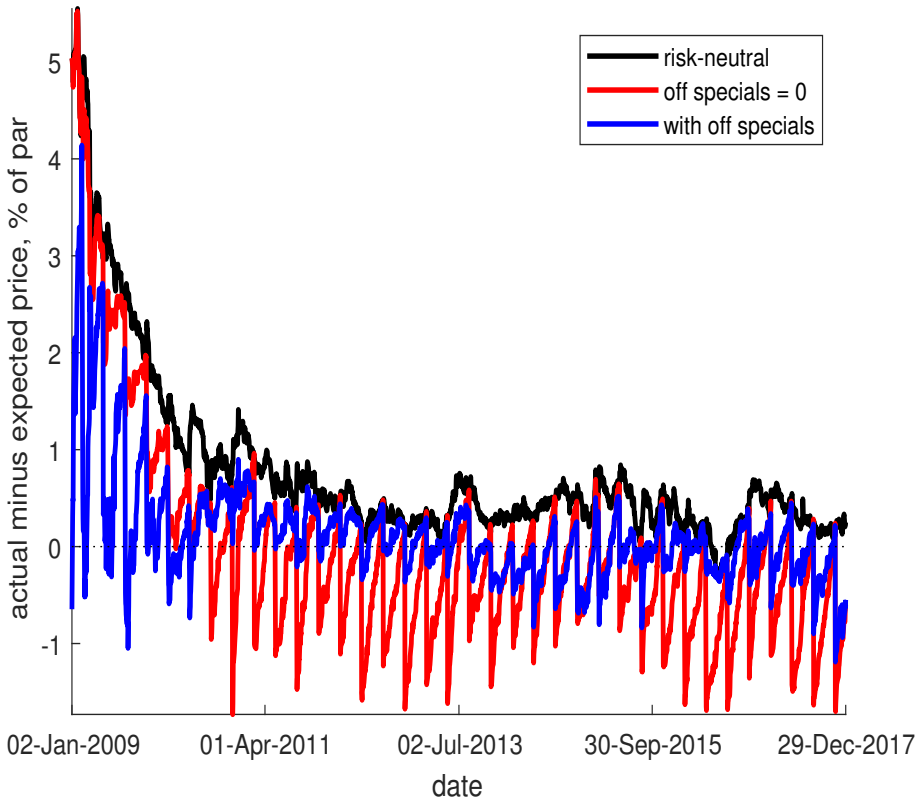


Figure 5. The figure plots the price residuals from equation (11) on the 10-year on-the-run bond for three different estimates of the model: risk-neutral special spreads (black), i.e. a 4-factor model the fourth row and column of Λ and the fourth element of λ equal to zero, and μ^* and Φ^* defined in equation (8); a 4-factor model with time-varying risk-premia (red), i.e. μ^* and Φ^* freely estimated but special spreads on off-the-run bonds set to zero; and a 5-factor model with time-varying risk premia on both on- and off-the-run bonds (blue).

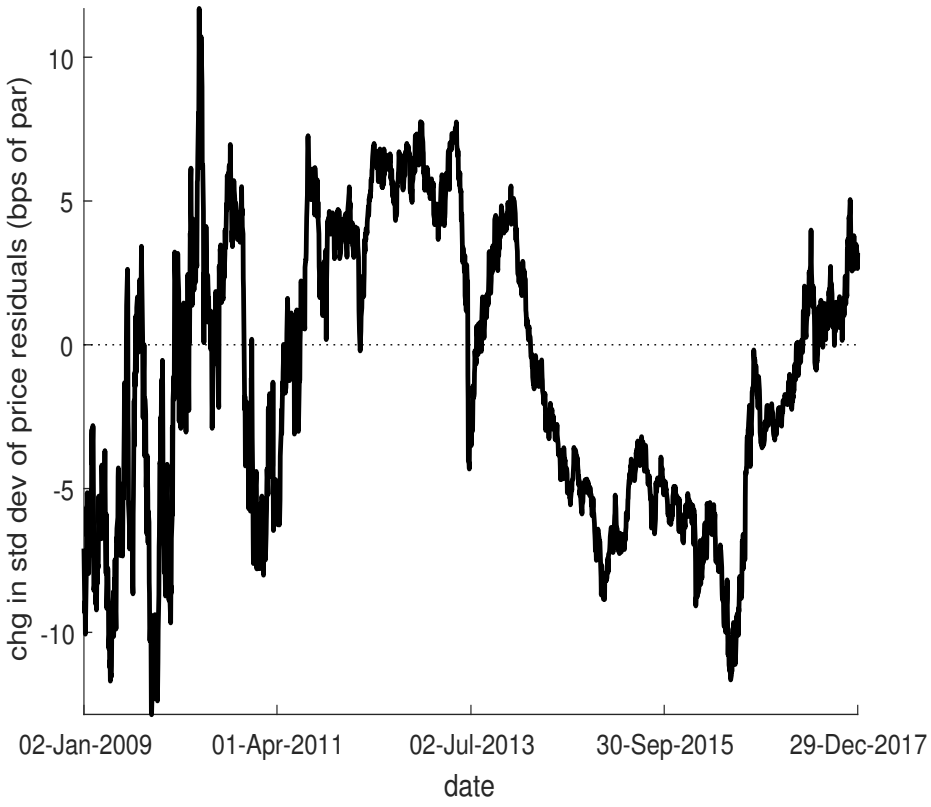


Figure 6. The figure plots the difference in the standard deviation of the residuals from equation (11) between the model with only 3 factors, and the model with 5 factors incorporating time-varying risk-premia on off-the-run special spreads, for each date in the sample, estimated only on off-the-run bonds. The values are reported in basis points of par.

Table 1. Sample Summary Statistics
The table reports summary statistics for our sample from January 2, 2009 to December 29, 2017, by maturity at issuance in years (first column). The second column reports the average number of bonds in each cross-section with the indicated maturity at issuance. The third column reports the number of off-the-run bonds that are on special, i.e. have a positive y according to equation (4). The fourth and fifth columns report the average specialness spread, in bps, for off- and on-the-run bonds, respectively. The last column reports the average price residual (actual minus expected), in percent of par, of on-the-run bonds from the 3-factor model without repo specials. Numbers in parenthese are standard deviations.

Maturity (years)	Avg # Bonds	% On Special (off-the-run)	Avg Spread (off-the-run)	Avg Spread (on-the-run)	Avg Price Residual (on-the-run)
All	220	84.7	4.79	19.7	0.127
			(6.15)	(41.6)	(0.605)
2	11.2	87.3	5.37	20.5	0.0259
			(8.8)	(38.7)	(0.371)
3	21.7	88.9	5.72	21.1	0.0658
			(7.6)	(36)	(0.248)
5	46.9	84.5	3.83	24.7	0.00523
			(6.04)	(41.7)	(0.244)
7	47.8	86.4	4.85	6.18	-0.046
			(5.47)	(11)	(0.41)
10	35.2	83.2	3.44	35.4	0.833
			(3.69)	(66.5)	(0.931)
30	58.6	82.4	5.86	8.97	-0.125
			(6.43)	(23.2)	(0.526)

Table 2. The table reports the estimated parameters of the models discussed in the text. Panel A reports the parameters of the 3- and 4-factor models, in which the first 3 factors are latent and the fourth factor applies to special spreads of the 10-year on-the-run bonds. Panel B reports the parameters for the 5-factor model in which the first 3 factors are latent, the 4th factor applies to special spreads of off-the-run bonds, and the 5th factor applies to special spreads of on-the-run bonds. Within each panel, the top row reports the risk-neutral parameters for the 3 latent factors of the model, with the top-left submatrix of Φ^* given by equation (12). The second row of each panel reports the risk-neutral parameters that pertain to the repo side of the model. The third row of each panel reports the Cholesky decomposition of the variance covariance matrix, Σ for each each model after estimating a VAR on the factors.

Panel A: 3- and 4-Factor Models

δ_0	ϕ_1^*	ϕ_2^*	ϕ_3^*	ρ	σ_x
-6.63e-06	1	-3.08e-07	1	0.909	0.0013
μ^*		Φ^*_{BL}		Φ^*_{BR}	
-0.0189	-1.56e+05	64.1	661	-0.947	
		Σ			
3.58e-09					
-2.07e-06	1.55e-06				
-1.28e-07	-1.31e-08	8.67e-07			
2.66e-6	-1.14e-06	6.69e-07	0.00112		

Panel B: 5-Factor Model

δ_0	ϕ_1^*	ϕ_2^*	ϕ_3^*	ρ	σ_x
-8.95e-06	1	-2.82e-07	1	0.795	0.00038
μ^*		Φ^*_{BL}		Φ^*_{BR}	
-0.000693	-6.71e+03	-6.66	0.175	0.757	0
0.137	6.32e+05	653	-1.09e+03	2.13	-0.976
		Σ			
4.08e-09					
-2.42e-06	1.6e-06				
-2.98e-07	2.12e-07	8.48e-07			
-1.93e-05	9.67e-06	-1.15e-05	0.000296		
-0.000147	3.69e-05	2.37e-05	0.000268	0.00103	

Table 3. The table reports regression results from estimating $\eta_{i,t} = \alpha_i + \beta_1 y_t^i + \beta_2 \eta_{i,t-1} + \xi_{i,t}$, where $\eta_{i,t}$ are estimated from equation (11) with the 3-factor model ignoring all special spreads, and the special spread y_t^i is defined in equation (4). Special spreads in the regression are annualized, and the price residuals are in percent of par value. The table reports t-statistics, clustered at the CUSIP level, in parentheses. The first two columns do not include CUSIP-level fixed effects, while the third and fourth columns do.

y_t^i	60.30***	0.514***	43.43***	0.483***
	(8.240)	(4.664)	(6.668)	(4.067)
$\eta_{i,t-1}$		0.989***		0.986***
		(851.9)		(616.6)
Observations	496,420	495,792	496,420	495,792
R-squared	0.019	0.984	0.012	0.977
CUSIP FE	NO	NO	YES	YES
# of CUSIP			628	628

t-statistics in parentheses, clustered at CUSIP level
*** p<0.01, ** p<0.05, * p<0.1