Bond Risk Premiums at the Zero Lower Bound

Martin M. Andreasen, Kasper Jørgensen, and Andrew Meldrum*

August 22, 2019

Abstract

This paper documents a significantly stronger relationship between the slope of the yield curve and future excess bond returns on Treasuries from 2008-2015 than before 2008. This new predictability result is not matched by the standard shadow rate model with Gaussian factor dynamics, but extending the model with regime-switching in the (physical) dynamics of the factors at the lower bound resolves this shortcoming. This extended model implies that, as the short rate rises away from the lower bound, bond risk premiums behave as they did before 2008.

Keywords: Dynamic term structure model, bond return predictability, shadow rate model, structural break, regime-switching.

JEL: E43, E44, G12.

*Andreasen, mandreasen@econ.au.dk, Aarhus University, CREATES, and Danish Finance Institute; Jørgensen, kasper.joergensen@frb.gov, Board of Governors of the Federal Reserve System; Meldrum, andrew.c.meldrum@frb.gov, Board of Governors of the Federal Reserve System. We thank Eric Engstrom, Joachim Grammig, Thomas B. King, and Cynthia Wu for useful comments and discussions, as well as participants at the 6th University of York Asset Pricing Workshop. The analysis and conclusions are those of the authors and do not indicate concurrence by the Board of Governors of the Federal Reserve System or other members of the research staff of the Board.

1 Introduction

One of the most widely used dynamic term structure models (DTSMs) in recent years has been the Gaussian shadow rate model (SRM) proposed by Black (1995) (see Kim and Singleton (2012), Christensen and Rudebusch (2015), Bauer and Rudebusch (2016), Wu and Xia (2016), and Andreasen and Meldrum (2018), among others). The popularity of the SRM stems from its ability to enforce the zero lower bound (ZLB) on nominal bond yields while preserving a (near-)linear relationship between bond yields and pricing factors away from the ZLB. However, in this paper we show that this standard version of the SRM has an important shortcoming in its ability to match key facts about the predictability of bond returns and hence the dynamics of bond risk premiums. We address this shortcoming by proposing a new SRM with regime-switching factor dynamics at the ZLB.

Beginning with the seminal studies of Fama and Bliss (1987) and Campbell and Shiller (1991), the finance literature has devoted substantial attention to regressions of excess bond returns on the slope of the yield curve, that is, the "yield spread" between long- and short-maturity bonds. These regressions are commonly used as diagnostic tests of whether a DTSM can explain deviations from the expectations hypothesis and generate predictability in excess bond returns. Specifically, Dai and Singleton (2002) require that a well-specified model should imply population regression coefficients that match those estimated from the data, and they show that a 3-factor Gaussian affine term structure model (ATSM) estimated on a data set when U.S. bond yields were away from the ZLB can satisfy this test.[1] However, Bauer and Rudebusch (2016) show that ATSMs imply material conditional probabilities of negative nominal interest rates during the recent ZLB episode, which shows the importance of also examining the performance of DTSMs that enforce the ZLB.

While the SRM provides the most popular way of enforcing the ZLB, its ability to match important facts about return predictability has received relatively little attention, perhaps

[1] Other prominent studies that adopt this diagnostic test include Bansal and Zhou (2002), Wachter (2006), Rudebusch and Wu (2007), Bansal and Shaliastovich (2013), and Kung (2015).

because the task of assessing ZLB-consistent models is complicated by the nonlinearities in the data close to the ZLB. A key assumption underpinning the diagnostic test of Dai and Singleton (2002) is that there is a stable linear relationship between excess bond returns and the yield spread in the data. Indeed, many studies of DTSMs use data sets that start during the 1980s or early 1990s because of a possible structural break in this linear relationship during the 1980s documented by Rudebusch and Wu (2007). In this paper, we document a further shift in the linear relationship between excess returns and the yield spread that appears to coincide with short-term interest rates reaching the ZLB in the United States in 2008. In this case, however, it would be premature to simply exclude data from before 2008 when estimating a SRM, for three reasons. First, it would leave us with a very short sample of data to estimate the model, exacerbating existing small-sample problems.[2] Second, it is not yet clear whether the break in 2008 is permanent or whether it is a temporary phenomenon associated with proximity to the ZLB. Third, it is worth exploring whether nonlinear term structure models such as the SRM can generate the observed state-dependency in the relationship between excess returns and the yield spread. To our knowledge, the only previous study considering the ability of the SRM to satisfy the Dai and Singleton (2002) diagnostic test is Andreasen and Meldrum (2018), who focus only on the narrower question of whether the standard SRM satisfies the test when interest rates are away from the ZLB.

To assess whether a term structure model is capable of explaining the shift in the regression coefficients at the ZLB, we propose a modification of the Dai and Singleton (2002) test. Specifically, we argue that a well-specified DTSM should match the slope coefficients both when the short rate is close to the ZLB and when it is away from the lower bound. We show that the standard SRM matches the slope coefficients away from the lower bound but that the model is unable to match the shift in these slope coefficients at the ZLB. This result suggests that simply enforcing the ZLB by truncating the short rate at zero in a Gaussian SRM is not sufficient to capture the change in bond yield dynamics that occurred at the

[2] For discussions of small-sample problems in DTSMs, see, for example, Bauer, Rudebusch and Wu (2012), Kim and Orphanides (2012), and Wright (2014).

ZLB.

Given this finding, a natural question is whether we can extend the standard SRM to match the observed shift in the relationship between excess returns and the yield spread. We consider two extensions that both allow for a structural change in the parameters determining the prices of risk and hence the time-series dynamics of the factors, as in Bansal and Zhou (2002) and Giacoletti, Laursen and Singleton (2018). The first extension we consider introduces regime-switching in the SRM by allowing the time-series dynamics of the pricing factors to change at the lower bound. We motivate this extension by a desire to accommodate the effects of various unconventional monetary policy measures that should have a largely temporary effect on bond yields during the ZLB episode. The second extension we consider allows for a permanent break in the time-series dynamics of the pricing factors in 2008 to accommodate long-lasting effects of recent developments, such as a possible "secular stagnation," as proposed by Summers (2015). Our results show that the SRM with regime-switching is able to match the shift in bond return predictability at the ZLB, meaning that this shift can be attributed to a temporary re-pricing of risk among bond investors at the lower bound that will revert once the short rate rises sufficiently far from the ZLB. We also find that the SRM with a permanent break does not improve on the performance of the standard SRM, suggesting that bond market investors do not perceive the shift in the relationship between excess returns and the yield spread to be permanent.

The remainder of this paper proceeds as follows. Section 2 documents the shift in bond return predictability during the recent ZLB period. Section 3 shows that the standard SRM cannot match this new empirical finding. Section 4 presents the two modifications of the standard SRM. Section 5 shows that our key conclusions are robust to including data on long-horizon interest rate surveys in the data set used to estimate the models, as proposed by Kim and Orphanides (2012). Section 6 shows that our conclusions about the SRM with a permanent break in the factor dynamics are robust to an alternative way of determining the timing of the break. Section 7 provides some concluding comments.

2 Shift in Bond Return Predictability at the ZLB

This section documents a significant shift in the ability of the yield spread to predict future excess bond returns during the recent ZLB period. We proceed by showing instability in the classic bond return predictability regressions in Section 2.1, while Section 2.2 presents our new empirical finding about the nature of this instability. Section 2.3 provides various robustness checks.

2.1 Instability in the Classic Return Predictability Regression

The risk premium implied by a long-maturity bond is often measured by its expected return in excess of the return on a short-maturity bond. These expected returns are not directly observed and therefore typically estimated by regressing realized excess bond returns $rx_{t,t+h}^{(n)}$ on a set of predictors $\mathbf{z}_t$. That is, by running the regression

$$rx_{t,t+h}^{(n)} = \beta_0^{(n)} + \boldsymbol{\beta}^{(n)\prime}\mathbf{z}_t + \varepsilon_{t,t+h}^{(n)}, \tag{1}$$

where $\varepsilon_{t,t+h}^{(n)}$ is a residual. Here, $rx_{t,t+h}^{(n)} \equiv p_{t+h}^{(n-h)} - p_t^{(n)} + p_t^{(h)}$ denotes the excess return on an n-month bond relative to an h-period bond between time t and $t+h$, and $p_t^{(n)}$ denotes the log price at time t of a zero-coupon bond with n periods to maturity. A natural benchmark is the expectations hypothesis, which implies $\boldsymbol{\beta}^{(n)} = 0$ and hence no bond return predictability.

One of the most prominent and robust predictors of excess bond returns is the slope of the yield curve, as used in the seminal work of Campbell and Shiller (1991).[3] Letting $y_t^{(n)}$ denote the yield at time t on an n period zero-coupon bond, the specification in Campbell

[3] A number of other yield curve and macroeconomic variables have recently been shown to predict excess returns (see, for example, Cochrane and Piazzesi (2005), Ludvigson and Ng (2009), Joslin, Priebsch and Singleton (2014), Cieslak and Povala (2015), and Bauer and Rudebusch (2019)).

and Shiller (1991) is equivalent to

$$rx_{t,t+h}^{(n)} = \beta_0^{(n)} + \beta_1^{(n)} \left(y_t^{(n)} - y_t^{(h)} \right) + \varepsilon_{t,t+h}^{(n)}, \tag{2}$$

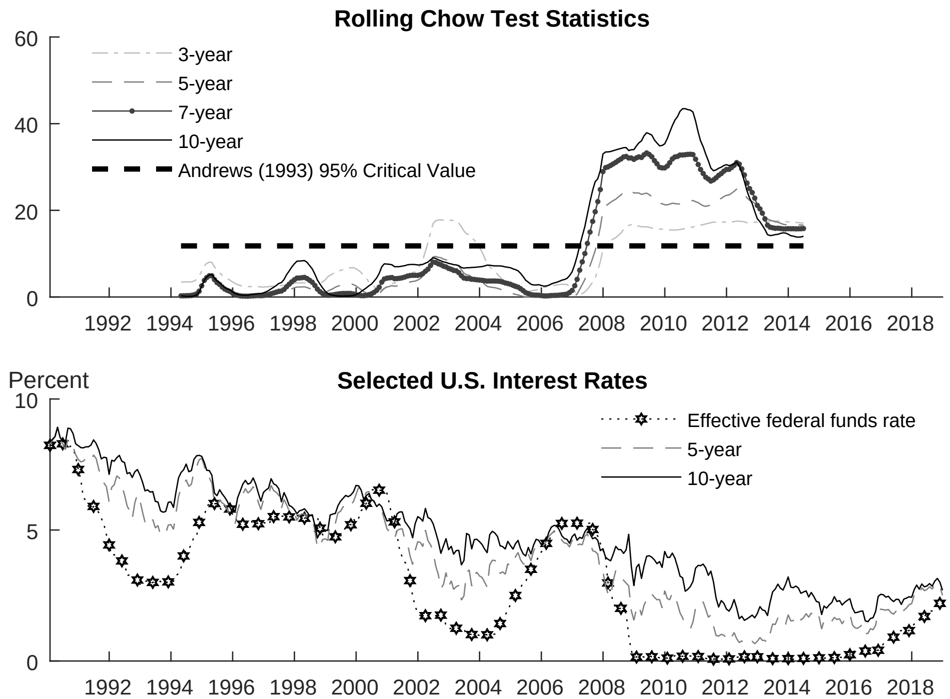
where the yield spread $y_t^{(n)} - y_t^{(h)}$ measures the slope of the yield curve.

A key assumption underpinning equation (2) is that there is a stable linear relationship between excess bond returns and the yield spread. However, the recent financial crisis has generated several unusual developments in the bond market that may have affected the relationship. The most obvious is probably that U.S. short-term interest rates were constrained by the ZLB from late 2008 to late 2015. This introduced an obvious nonlinearity in the yield curve that most likely caused the yield spread $y_t^{(n)} - y_t^{(h)}$ to be smaller than it would otherwise have been, because the short-term yield $y_t^{(h)}$ was constrained from below. Such a slope compression is likely to have increased $\beta_1^{(n)}$ because a given yield spread during the ZLB period carried a stronger signal about future bond returns than before the financial crisis.

A second important development was the Large-Scale Asset Purchase (LSAP) programs that the Federal Reserve used during the ZLB period to stimulate the U.S. economy. These programs were mainly designed to temporarily lower long-term yields by reducing bond risk premiums, as discussed in D'Amico, English, Lopez-Salido and Nelson (2012), Li and Wei (2013), and Bonis, Ihrig and Wei (2017). Thus, the LSAP programs are likely to have affected both the level of excess bond returns and the yield spread, implying that the coefficients in equation (2) may be affected.

Finally, there may have been a permanent reduction in expectations about the long-term level of the short rate, perhaps due to concerns about a secular stagnation in the U.S. (see, for instance, Summers (2015), Laubach and Williams (2016), Del Negro, Giannone, Giannoni and Tambalotti (2017), Christensen and Rudebusch (2019)). Such a shift in expectations also seems capable of affecting the coefficients in equation (2).

Figure 1: Chow Tests and Historical Bond Yields
The top chart reports the Chow test statistics for breaks in equation (2) using at least 15 percent of the observations for the pre- and post-break sample. The reported 95 percent critical value for the maximum Chow test statistic is taken from Andrews (1993). The bottom chart shows selected historical bond yields. The yields are end-month from Gürkaynak, Sack and Wright (2007) starting in January 1990 and ending in December 2018.



Given these considerations, we formally test whether the coefficients in equation (2) are constant across time using a monthly sample of U.S. bond yields from Gürkaynak et al. (2007) between January 1990 and December 2018, adopting a 1-year holding period ($h = 12$). We apply separate tests to regressions in equation (2) for various maturities. Given a sample observed at times $t = 1, 2, ..., T$, the null hypothesis is that there is no structural break between periods $\tau_1 > \lceil \kappa T \rceil$ and $\tau_2 < \lfloor (1 - \kappa) T \rfloor$ and the alternative hypothesis is that there is a single break at an unknown time τ_B where $\tau_1 \leq \tau_B \leq \tau_2$. Setting $\kappa > 0$ means that we ensure that the pre- and post-break periods each contain at least a minimum proportion of the total sample. For all candidate break points $\tau = \tau_1, \tau_1 + 1, ..., \tau_2$, we compute the Chow test statistic

$$F(\tau) = \frac{SSR_{1,T} - (SSR_{1,\tau-1} + SSR_{\tau,T})}{(SSR_{1,\tau-1} + SSR_{\tau,T})/(T-2k)}, \tag{3}$$

where $SSR_{i,j}$ denotes the sum of squared residuals for equation (2) estimated on a sample from period i to period j and the number of parameters in the restricted model is $k = 2$. The Quandt (1960) likelihood ratio statistic for a break at an unknown time is given by

$$QLR = \max_{\tau=\tau_1,\ldots,\tau_2} F(\tau). \tag{4}$$

Andrews (1993) tabulates the relevant critical values and we refer the reader to Stock (1994) for further details.

The top chart in Figure 1 shows the Chow test statistics $F(\tau)$ for returns on 3-, 5-, 7-, and 10-year bonds for all of the considered break points, where we set $\kappa = 0.15$ to ensure that the pre- and post-break samples contain at least 15 percent of the observations. The dashed line shows the Andrews (1993) 95 percent critical value for the QLR test statistic. The bottom chart in Figure 1 shows time series of various U.S. interest rates. We find that the Chow test statistics for all bond yields increase sharply around the time the effective federal funds rate reached the ZLB in 2008, and reach their maximum values, well above the Andrews (1993) 95 percent critical value, fairly soon thereafter.

In summary, the classic predictability regressions in equation (2) display evidence of a break around 2008 when the U.S. short rate approached the ZLB.

2.2 Formalizing the Shift in Bond Return Predictability

The next step is to examine the nature of this break in the classic return predictability regression. The first two of the unusual bond market developments mentioned above (that is, the ZLB episode and the LSAP programs) were specific to the period between 2008 and late 2015. These developments seem therefore likely to only have had an effect on bond yields that is largely contained to that period. Hence, a natural extension of equation (2) is to allow for a threshold effect depending on the level of the short rate. That is, to consider a regression of the form

$$rx_{t,t+h}^{(n)} = \beta_{0,1}^{(n)}\mathcal{I}_{\{r_t \geq c\}} + \beta_{0,2}^{(n)}\mathcal{I}_{\{r_t < c\}} + \left(\beta_{1,1}^{(n)}\mathcal{I}_{\{r_t \geq c\}} + \beta_{1,2}^{(n)}\mathcal{I}_{\{r_t < c\}}\right)\left(y_t^{(n)} - y_t^{(h)}\right) + \varepsilon_{t,t+h}^{(n)}, \quad (5)$$

where $\mathcal{I}_{\{r_t \geq c\}}$ is an indicator function that takes a value of 1 at time t if the short rate $r_t \geq c$ and 0 otherwise. As a result, equation (5) allows the regression coefficients in equation (2) to shift when r_t becomes sufficiently low.[4]

On the other hand, concerns about secular stagnation and a lower long-term level for the short rate are not specific to the ZLB period and may therefore affect bond yields for several years after the ZLB period. Hence, an alternative extension of equation (2) is to allow for a permanent break at time τ. That is, to consider a regression of the form

$$rx_{t,t+h}^{(n)} = \beta_{0,1}^{(n)}\mathcal{I}_{\{t < \tau\}} + \beta_{0,2}^{(n)}\mathcal{I}_{\{t \geq \tau\}} + \left(\beta_{1,1}^{(n)}\mathcal{I}_{\{t < \tau\}} + \beta_{1,2}^{(n)}\mathcal{I}_{\{t \geq \tau\}}\right)\left(y_t^{(n)} - y_t^{(h)}\right) + \varepsilon_{t,t+h}^{(n)}, \quad (6)$$

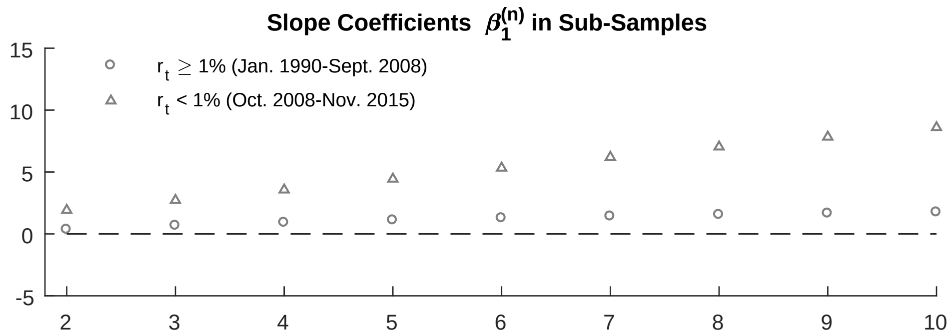
where $\mathcal{I}_{\{t < \tau\}}$ is an indicator function that takes a value of 1 at time t if $t < \tau$ and 0 otherwise.

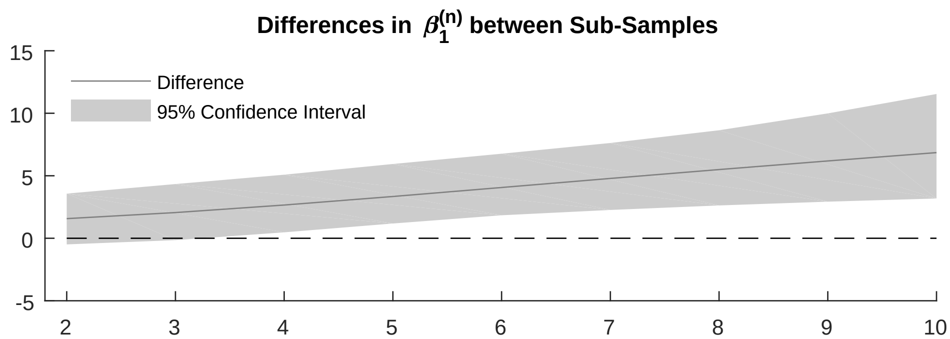
Given that there is only one ZLB period for the postwar U.S. economy in our sample, it is hard to tell which of the nonlinear specifications in equations (5) and (6) we should prefer. To avoid taking a view on the best specification at this stage, we simply estimate the standard predictability regression in equation (2) using two separate samples. The first sample ("Regime 1") runs from January 1990 to September 2008, excluding December 2003 when the short rate was slightly below 1 percent. The second sample ("Regime 2") runs from October 2008 to November 2015, which is the period where the short rate was close to the ZLB. Because the first sample includes only observations before October 2008 with a short rate above 1 percent, and the second sample includes only observations after October

[4] An alternative specification would be to allow the coefficients of a return regression to transition smoothly as the short rate approaches the ZLB. However, in most of our data set the short rate is either well away from the ZLB or very close to it, so there is likely to be insufficient information to estimate the parameters of such a smooth transition reasonably precisely.

Figure 2: Bond Return Predictability

The top chart reports the slope coefficient in $rx_{t,t+h}^{(n)} = \beta_0^{(n)} + \beta_1^{(n)} \left(y_t^{(n)} - y_t^{(h)} \right) + \varepsilon_{t,t+h}^{(n)}$ with $h = 12$ when estimated from January 1990 to September 2008 (excluding December 2003) where $r_t \geq 0.01$, and when estimated from October 2008 to November 2015 where $r_t < 0.01$. The bottom chart reports the difference in the slope estimates at a given maturity. The 95 percent confidence interval for these differences are computed using a block bootstrap with 5,000 repetitions and a block window of 24 months. In each bootstrap sample it is required that there are at least 50 observations where the short rate is above and below 1 percent, respectively. The x-axes report maturity in years, and all yields are end-month from Gürkaynak et al. (2007) with r_t measured by the effective federal funds rate.





2008 with a short rate below 1 percent, the estimates are consistent with either equation (5) with $c = 0.01$ or equation (6) with $\tau = 225$ (October 2008 is the 225th month of our sample).

We estimate the regressions for both regimes using monthly observations of yields and excess returns on bonds with maturities from 2 to 10 years using a 1-year holding period as above. The gray circles in the top chart of Figure 2 for the pre-ZLB period reveal the usual empirical pattern that the slope coefficients in equation (2) are positive and increase with maturity. Our new empirical finding is that these slope coefficients are larger and increase faster with maturity during the recent ZLB period (the gray triangles) when compared to the pre-ZLB period. The bottom chart in Figure 2 shows that these differences in the slope coefficients are significant at the 5 percent level for maturities beyond 3 years.[5] Thus, recent developments in the bond market imply a shift in bond return predictability with a much stronger relationship between the yield spread and excess bond returns than observed in the pre-ZLB period.

2.3 Robustness

We examine the robustness of this shift in bond return predictability in Table 1, where $\Delta\beta_1^{(n)} \equiv \beta_{1,2}^{(n)} - \beta_{1,1}^{(n)}$ refers to the change in the slope coefficient in equation (2). The first column examines the importance of the holding period by considering biannual bond returns ($h = 6$) instead of the annual horizon we use elsewhere. The shift in the slope coefficients is again positive, and it is significant beyond the 4-year maturity.

Our new finding is also robust to measuring the slope of the yield curve by the second principal component (PCA2) of bond yields, as shown by the second column in Table 1.

The next column shows that the shift in bond return predictability is also robust to replacing the yield spread by the forward spread $f_t^{(n,h)} - y_t^{(h)}$, as in Fama and Bliss (1987), where the forward rate is given by $f_t^{(n,h)} = \log\left(P_t^{(n+h)}/P_t^{(n)}\right)$.

The fourth robustness check reported in Table 1 adds the Chicago Fed National Activity Index (CFNAI) as in Joslin et al. (2014) and the inflation trend factor of Cieslak and Povala (2015) as additional regressors to equation (2) to control for the effect of macro variables. Table 1 shows that we again find a positive and significant shift in the ability of the yield spread to predict future bond returns during the ZLB period.

The fifth column in Table 1 controls for the decline in the natural (or equilibrium) nominal rate i_t^*, which Bauer and Rudebusch (2019) show transmits to long-term bond yields. We use the approach in Christensen and Rudebusch (2019) among others and measure the natural

[5] Unreported results show that allowing for a break in this way roughly doubles the R^2 statistics across the considered maturities.

Table 1: Bond Return Predictability: Robustness

Column (1) estimates equation (2) with $h = 6$. Column (2) replaces $y_t^{(n)} - y_t^{(h)}$ in equation (2) by the second principal component of bond yields. Column (3) replaces $y_t^{(n)} - y_t^{(h)}$ in equation (2) by the forward spread $f_t^{(n,h)} - y_t^{(h)}$. Column (4) adds CFNAI and the inflation trend factor of Cieslak and Povala (2015) as two additional macro control variables to equation (2). Column (5) adds $\mathcal{I}_{\{r_t<c\}} i_t^*$ to equation (2) to control for the effect of the natural nominal short rate i_t^*, when measured by the expected 3-month Treasury bill yield from 5 to 10 years ahead in the Blue Chip Economic Indicators survey. These alternative specifications are estimated from January 1990 to September 2008 (excluding December 2003) where $r_t \geq 0.01$ and from October 2008 to November 2015 where $r_t < 0.01$. Standard errors are provided in parenthesis using a block bootstrap with 5,000 repetitions and a block window of 24 months. In each bootstrap sample it is required that there are at least 50 observations where the short rate is above and below 1 percent, respectively. Significance at the 10, 5, and 1 percent level is denoted by $\star$, $\star\star$, and $\star\star\star$. All yields are end-month from Gürkaynak et al. (2007) with r_t measured by the effective federal funds rate.

	(1)		(2)		(3)		(4)		(5)	
	$h = 6$		PCA2		$f_t^{(n,h)} - y_t^{(h)}$		Macro		Trend	
n	$\beta_1^{(n)}$	$\Delta\beta_1^{(n)}$	$\beta_1^{(n)}$	$\Delta\beta_1^{(n)}$	$\beta_1^{(n)}$	$\Delta\beta_1^{(n)}$	$\beta_1^{(n)}$	$\Delta\beta_1^{(n)}$	$\beta_1^{(n)}$	$\Delta\beta_1^{(n)}$
2	0.65	0.83	0.04	0.45	0.17	0.79	0.12	2.09*	0.34	1.19
	(0.57)	(0.67)	(0.25)	(0.28)	(0.50)	(0.51)	(0.99)	(1.13)	(0.99)	(1.00)
3	0.84	1.15	0.19	1.11**	0.34	1.05*	0.49	2.39*	0.66	1.65
	(0.63)	(0.80)	(0.46)	(0.53)	(0.57)	(0.61)	(1.07)	(1.30)	(1.07)	(1.13)
4	0.96	1.53	0.39	1.95***	0.47	1.49**	0.75	2.84**	0.90	2.25*
	(0.64)	(0.96)	(0.62)	(0.70)	(0.59)	(0.68)	(1.07)	(1.36)	(1.08)	(1.23)
5	1.05*	1.92*	0.64	2.88***	0.57	2.04***	0.97	3.41**	1.10	3.03**
	(0.63)	(1.10)	(0.74)	(0.84)	(0.60)	(0.70)	(1.05)	(1.36)	(1.05)	(1.34)
6	1.10*	2.31*	0.91	3.82***	0.66	2.67***	1.15	4.06***	1.26	4.02***
	(0.61)	(1.28)	(0.83)	(1.00)	(0.61)	(0.71)	(1.02)	(1.36)	(1.02)	(1.46)
7	1.15*	2.67**	1.19	4.70***	0.72	3.34***	1.32	4.76***	1.41	5.23***
	(0.60)	(1.24)	(0.90)	(1.21)	(0.62)	(0.85)	(1.01)	(1.43)	(0.99)	(1.58)
8	1.19**	3.01**	1.46	5.50***	0.78	4.06***	1.46	5.48***	1.54	6.63***
	(0.59)	(1.31)	(0.95)	(1.47)	(0.63)	(1.13)	(0.99)	(1.61)	(0.97)	(1.68)
9	1.22**	3.32**	1.73*	6.19***	0.83	4.87***	1.58	6.20***	1.65*	8.19***
	(0.59)	(1.41)	(1.00)	(1.78)	(0.65)	(1.47)	(0.99)	(1.88)	(0.96)	(1.74)
10	1.25**	3.62**	1.99*	6.78***	0.88	5.78***	1.69*	6.91***	1.74*	9.87***
	(0.59)	(1.54)	(1.03)	(2.11)	(0.67)	(1.85)	(0.99)	(2.21)	(0.95)	(1.78)

rate by its long-horizon expectations from 5 to 10 years into the future. Specially, we measure i_t^* by the average expected 3-month Treasury bill yield from 5 to 10 years ahead, as reported biannually in the Blue Chip Economic Indicators survey.[6] Given that i_t^* operates as a level factor affecting all yields equally, it should cancel out in the yield spread away from the ZLB but not close to the bound, where short-term interest rates cannot respond one-to-one with i_t^*. This suggests that i_t^* should only affect excess bond returns close to the ZLB, and

[6] The forecast horizons are not the same in each edition of the survey, so we linearly interpolate the mean expectations across respondents to get short rate expectations at the desired horizon. This biannual series is then extended to a monthly frequency by linear interpolation. The data source is Wolters Kluwer Legal and Regulatory Solutions U.S., Blue Chip Economic Indicators.

we therefore add the regressor $\mathcal{I}_{\{r_t<c\}}i_t^*$ to equation (2). The last column in Table 1 shows that controlling for i_t^* cannot explain our new empirical result, as we also in this case find a positive and significant shift in the ability of the yield spread to predict bond returns.[7]

Thus, the shift in bond return predictability documented in Section 2.2 is robust to several modifications and extensions of the classic predictability regression in (2).

3 A Shortcoming of the Standard Shadow Rate Model

This section shows that the standard SRM with Gaussian factor dynamics cannot explain the shift in bond return predictability documented in Section 2. We proceed by describing the model in Section 3.1, our estimation method in Section 3.2, and the inability of the standard SRM to generate stronger return predictability from the yield spread at the ZLB in Section 3.3.

3.1 A Shadow Rate Model

Suppose that the U.S. economy can be described by a set of factors collected in $\mathbf{x}_t$ of dimension $n_x \times 1$ that evolves as

$$\mathbf{x}_{t+1} = \mathbf{h}\left(\mathbf{x}_t\right) + \mathbf{\Sigma}\boldsymbol{\varepsilon}_{t+1}^{\mathbb{P}}. \tag{7}$$

The function $\mathbf{h}\left(\mathbf{x}_t\right)$ is potentially nonlinear and $\boldsymbol{\varepsilon}_{t+1}^{\mathbb{P}}$ is an $n_x \times 1$ vector of independent standard Gaussian innovations under the physical probability measure $\mathbb{P}$, denoted $\boldsymbol{\varepsilon}_{t+1}^{\mathbb{P}} \sim \mathcal{NID}\left(\mathbf{0}, \mathbf{I}\right)$. Throughout, we consider models with the standard 3 pricing factors (that is, $n_x = 3$). The stochastic discount factor M_{t+1} is assumed to have the flexible form

$$M_{t+1} = \exp\left\{-r_t - \boldsymbol{\lambda}\left(\mathbf{x}_t\right)' \boldsymbol{\lambda}\left(\mathbf{x}_t\right) - \boldsymbol{\lambda}\left(\mathbf{x}_t\right)\boldsymbol{\varepsilon}_{t+1}^{\mathbb{P}}\right\},$$

[7] Unreported results show that we also find a positive and significant shift in $\beta_1^{(n)}$ if we simply add our measure of i_t^* as an extra regressor in equation (2) to account for the effect of i_t^* away from the ZLB and close to the lower bound.

where r_t is the one-period nominal short rate and $\boldsymbol{\lambda}(\mathbf{x}_t)$ is a potentially nonlinear function for the market prices of risk. Many macroeconomic models relate M_{t+1} to consumption and inflation. The precise structural underpinnings of M_{t+1} are, however, not provided in reduced-form DTSMs to reduce the risk of model misspecification, although the factors in $\mathbf{x}_t$ are linked to the U.S. economy at an overall level. The short rate is given by

$$r_t = \max\{0, s_t\}, \tag{8}$$

where the use of the maximum function following Black (1995) constrains r_t from below at zero and $s_t \equiv \alpha + \boldsymbol{\beta}'\mathbf{x}_t$ is an unconstrained "shadow rate." To make the bond pricing tractable, it is assumed that the market prices of risk are given by

$$\boldsymbol{\lambda}(\mathbf{x}_t) = \boldsymbol{\Sigma}^{-1}(\mathbf{h}(\mathbf{x}_t) - \boldsymbol{\Phi}\boldsymbol{\mu} - (\mathbf{I} - \boldsymbol{\Phi})\mathbf{x}_t), \tag{9}$$

because it implies a first-order vector auto-regression (VAR) under the risk-neutral probability measure $\mathbb{Q}$. That is,

$$\mathbf{x}_{t+1} = \boldsymbol{\Phi}\boldsymbol{\mu} + (\mathbf{I} - \boldsymbol{\Phi})\mathbf{x}_t + \boldsymbol{\Sigma}\boldsymbol{\varepsilon}_{t+1}^{\mathbb{Q}}, \tag{10}$$

where $\boldsymbol{\varepsilon}_{t+1}^{\mathbb{Q}} \sim \mathcal{NID}(\mathbf{0}, \mathbf{I})$.

The standard SRM with Gaussian factor dynamics is obtained by letting $\mathbf{h}(\mathbf{x}_t) = \mathbf{h}_0 + \mathbf{H}_x\mathbf{x}_t$, implying that equation (7) reduces to

$$\mathbf{x}_{t+1} = \mathbf{h}_0 + \mathbf{H}_x\mathbf{x}_t + \boldsymbol{\Sigma}\boldsymbol{\varepsilon}_{t+1}^{\mathbb{P}}. \tag{11}$$

Yields do not have closed-form expressions in this version of the SRM, and we therefore use the second-order approximation of Priebsch (2013). For identification, we impose the standard restrictions that $\boldsymbol{\beta}' = \mathbf{1}$, $\boldsymbol{\mu} = \mathbf{0}$, $\boldsymbol{\Sigma}$ is lower triangular, and $\boldsymbol{\Phi}$ is in Jordan form with increasing diagonal elements (see Joslin, Singleton and Zhu (2011)). A preliminary analysis reveals that the first eigenvalue of $\mathbf{\Phi}$ is often indistinguishable from zero, meaning that the pricing factors have a near unit root under $\mathbb{Q}$. This implies that the unconditional mean of the shadow rate under $\mathbb{Q}$ is badly identified, and we therefore impose $\mathbf{\Phi}(1,1) = \alpha = 0$ to ensure identification (see Hamilton and Wu (2012)).[8] Finally, the $\mathbb{P}$ transition parameters $\mathbf{h}_0$ and $\mathbf{H}_x$ are unrestricted.

3.2 Estimation Method and Data

Most previous studies that estimate ZLB-consistent DTSMs use quasi-maximum likelihood (QML) methods. However, the joint optimization of all parameters required by standard QML is computationally challenging and the asymptotic properties of these QML estimators are unknown in nonlinear models such as the SRM. We overcome these limitations by using the sequential regression (SR) approach of Andreasen and Christensen (2015), which is computationally simpler and provides asymptotically Gaussian parameter estimates.

To apply the SR approach to the standard SRM, it is convenient to define two vectors with partly overlapping sub-sets of parameters. First, $\boldsymbol{\theta}_1$ collects the "$\mathbb{Q}$ parameters" in equation (10) that determine the cross-sectional relationship between the pricing factors and bond yields. Second, $\boldsymbol{\theta}_2$ collects the "$\mathbb{P}$ parameters" in equation (11) that determine the time-series dynamics of the pricing factors. Because $\mathbf{\Sigma}$ appears in both $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$, it is convenient to partition these vectors further as $\boldsymbol{\theta}_1 = \begin{bmatrix} \boldsymbol{\theta}_{11}' & vech(\mathbf{\Sigma})' \end{bmatrix}'$ and $\boldsymbol{\theta}_2 = \begin{bmatrix} \boldsymbol{\theta}_{22}' & vech(\mathbf{\Sigma})' \end{bmatrix}'$, where $\boldsymbol{\theta}_{11} = \begin{bmatrix} \mathbf{\Phi}(2,2) & \mathbf{\Phi}(3,3) \end{bmatrix}'$ and $\boldsymbol{\theta}_{22} = \begin{bmatrix} \mathbf{h}_0' & vec(\mathbf{H}_x)' \end{bmatrix}'$.

The SR approach proceeds in three steps. At step 1, $\boldsymbol{\theta}_1$ and the pricing factors $\mathbf{x}_t$ are jointly estimated using nonlinear regressions from $n_{y,t}$ yields with maturities $m_1, m_2, \ldots, m_{n_{y,t}}$ in period t. The observed yield with maturity m_j at time t is denoted $y_t^{(m_j)} = g_{m_j}(\mathbf{x}_t; \boldsymbol{\theta}_1) + v_{m_j,t}$, where $g_{m_j}(\mathbf{x}_t; \boldsymbol{\theta}_1)$ is the model-specific function that relates the pricing factors to the

[8] Previous studies that have imposed these restrictions on $\mathbf{\Phi}(1,1)$ and α include Christensen and Rudebusch (2015).

cross section of yields and $v_{m_j,t}$ is a measurement error. We assume that these measurement errors have zero means and finite, positive-definite covariance matrices. For a given value of $\boldsymbol{\theta}_1$, we obtain the pricing factors in each period by

$$\widehat{\mathbf{x}}_t(\boldsymbol{\theta}_1) = \underset{\mathbf{x}_t \in \mathbb{R}^{n_x}}{\arg\min} \; \frac{1}{2n_{y,t}} \sum_{j=1}^{n_{y,t}} \left(y_t^{(m_j)} - g_{m_j}(\mathbf{x}_t; \boldsymbol{\theta}_1) \right)^2 . \tag{12}$$

The estimate of $\boldsymbol{\theta}_1$ is then given by minimizing the sum of squared residuals from equation (12). That is,

$$\widehat{\boldsymbol{\theta}}_1^{step1} = \underset{\boldsymbol{\theta}_1 \in \boldsymbol{\Theta}_1}{\arg\min} \frac{1}{2N} \sum_{t=1}^{T} \sum_{j=1}^{n_{y,t}} \left(y_t^{(m_j)} - g_{m_j}(\widehat{\mathbf{x}}_t(\boldsymbol{\theta}_1); \boldsymbol{\theta}_1) \right)^2 , \tag{13}$$

where $\widehat{\boldsymbol{\theta}}_1^{step1}$ denotes the step 1 estimate of $\boldsymbol{\theta}_1$, $N = \sum_{t=1}^{T} n_{y,t}$, and $\boldsymbol{\Theta}_1$ is the feasible domain of $\boldsymbol{\theta}_1$. Andreasen and Christensen (2015) show that $\widehat{\boldsymbol{\theta}}_1^{step1}$ is consistent and asymptotically Gaussian for $n_{y,t} \to \infty$ for all t given standard regularity conditions.

At step 2 of the SR approach, the time-series parameters $\boldsymbol{\theta}_2$ are estimated using the estimated factors $\widehat{\mathbf{x}}_t\left(\widehat{\boldsymbol{\theta}}_1^{step1}\right)$ from step 1 with a correction for estimation uncertainty in these factors. As shown in Andreasen and Christensen (2015), this procedure corresponds to running a modified VAR, where all second moments are corrected for estimation uncertainty in $\widehat{\mathbf{x}}_t\left(\widehat{\boldsymbol{\theta}}_1^{step1}\right)$. The details are provided in Appendix A.

At step 3 of the SR approach, the estimates of $\boldsymbol{\Sigma}$ from step 1 and 2 can be combined optimally and, conditional on the optimal estimate of $\boldsymbol{\Sigma}$, the remaining $\mathbb{Q}$ parameters $\boldsymbol{\theta}_{11}$ and the $\mathbb{P}$ parameters $\boldsymbol{\theta}_2$ are re-estimated. Preliminary results for our applications reveal that $\boldsymbol{\Sigma}$ is estimated very imprecisely at step 1 compared with step 2.[9] So, for simplicity, we

[9] This finding is similar to the result of Joslin et al. (2011) for affine DTSMs, because their estimates of $\boldsymbol{\Sigma}$ from the time-series dynamics of their (observed) factors hardly change when taking account of the cross section of bond yields.

settle by conditioning on the estimate $\widehat{\Sigma}^{step2}$ at step 2 and re-estimate $\boldsymbol{\theta}_{11}$ as

$$\widehat{\boldsymbol{\theta}}_{11}^{step3} = \underset{\boldsymbol{\theta}_{11}\in\boldsymbol{\Theta}_{11}}{\arg\min} \frac{1}{2N} \sum_{t=1}^{T} \sum_{j=1}^{n_{y,t}} \left(y_t^{(m_j)} - g_{m_j} \left(\widehat{\mathbf{x}}_t \left(\boldsymbol{\theta}_{11}, \widehat{\Sigma}^{step2} \right) ; \boldsymbol{\theta}_{11}, \widehat{\Sigma}^{step2} \right) \right)^2, \tag{14}$$

where $\widehat{\boldsymbol{\theta}}_{11}^{step3}$ denotes the step 3 estimate of $\boldsymbol{\theta}_{11}$ and $\boldsymbol{\Theta}_{11}$ is the feasible domain of $\boldsymbol{\theta}_{11}$. Finally, we update the estimate of $\boldsymbol{\theta}_2$ by re-running step 2 using the estimated factors from step 3, that is $\widehat{\mathbf{x}}_t \left(\widehat{\boldsymbol{\theta}}_{11}^{step3}, \widehat{\Sigma}^{step2} \right)$.

The SR approach is designed for a setting with a large cross section, and we therefore include more yields than typically used when estimating DTSMs. That is, we represent the yield curve by 25 points, using the 3-month yield, the 6-month yield, yields in the 1-year to 3-year range at 3-month intervals, and yields in the 3- to 10-year range at 6-month intervals. For the 3- and 6-month yields, we use Treasury bill yields, while yields with longer maturities are obtained from the Gürkaynak et al. (2007) data set. Our monthly sample covers the period from January 1990 through December 2018, where the starting point is chosen to reduce the possibility of a structural break associated with the shift in U.S. monetary policy during the 1980s (see Rudebusch and Wu (2007)).

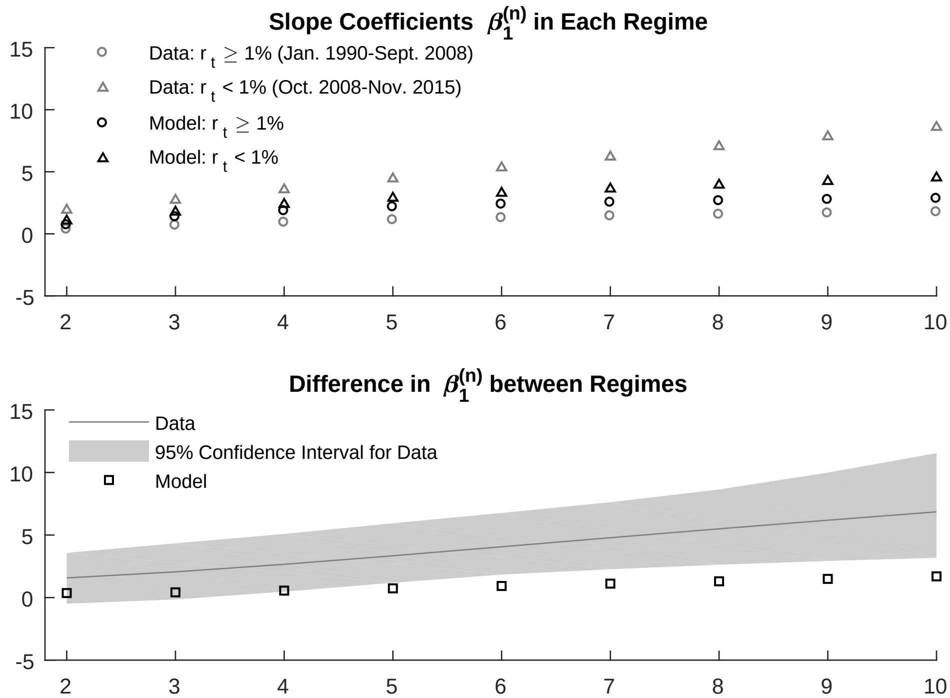
3.3 Matching the Shift in Bond Return Predictability

We next examine whether the standard SRM can explain the shift in bond return predictability when the short rate approaches the ZLB. Given that the standard SRM has constant parameters, it is obvious that it cannot match this shift by generating a permanent break in the regression coefficients, as implied by equation (6). However, this model may be able to match the shift in bond return predictability through a threshold effect due to its nonlinear mapping between the pricing factors and bond yields. We therefore simulate a sample of 1 million observations from the standard SRM and estimate equation (5) with a threshold of 1 percent.

When the short rate is above 1 percent, the top chart in Figure 3 shows that the standard

Figure 3: Bond Return Predictability: Standard SRM

The top chart shows the ability of the standard SRM to match the shift in the slope coefficients in equation (5) with $h = 12$ and $c = 0.01$ using a simulated sample of 1 million observations. The corresponding data moments are from equation (2) with $h = 12$ when estimated from January 1990 to September 2008 (excluding December 2003) where $r_t \geq 0.01$, and when estimated from October 2008 to November 2015 where $r_t < 0.01$. The bottom chart shows the difference in the slope coefficients in equation (5) at a given maturity as implied by the standard SRM and the data. All x-axes show the maturity in years.



SRM (the black circles) matches closely the corresponding data moment (the gray circles). This finding is perhaps not too surprising given that the standard SRM away from the ZLB reduces to the Gaussian affine model, which is able to match these moments (see, for instance, Dai and Singleton (2002)).[10] Turning to the more interesting case when the short rate is below 1 percent, we find that the standard SRM (the black triangles) is indeed able

[10] Our results for the pre-ZLB period contrast somewhat with the findings of Andreasen and Meldrum (2018). They show that once the ZLB period is included in the estimation sample, the standard SRM does not match the desired slope coefficients conditional on yields being away from the ZLB. Unreported results indicate that the difference in the results is because bond yields are away from the ZLB for a greater proportion of our sample than in the sample considered in Andreasen and Meldrum (2018), implying that the presence of the ZLB period has a relatively small effect on the estimated time-series dynamics in our case. The unreported results also show that our main results that the slope coefficients implied by the SRM do not shift at the ZLB is robust to considering a shorter sample.

to generate larger slope coefficients when compared to the loadings away from the ZLB. But this increase in the slope coefficients is insufficient to match the shift in the data. This is shown by the bottom chart in Figure 3, as the differences between the model-implied slope coefficients lie outside the 95 percent confidence interval for the corresponding data moment for maturities exceeding 3 years. Finally, in the Online Appendix we show that these results are robust to reducing the value of the slope coefficients in equation (5) for both the model and the data to 0.25 percent.

Thus, it appears that simply enforcing the ZLB by the standard SRM is insufficient to capture the shift in bond return predictability that occurred during the recent ZLB period.

4 Two Extended Shadow Rate Models

This section proposes two extensions of the standard SRM by introducing nonlinearities in the $\mathbb{P}$ dynamics of the pricing factors. Section 4.1 describes the two extensions, which allow for i) regime-switching in the pricing factors or ii) a permanent break in the factor dynamics. Section 4.2 shows that it is impossible to distinguish between these two models (or the standard SRM) from the in-sample fit to the cross-section of yields, but that the models have substantially different implications for the time-series dynamics of bond yields. Section 4.3 shows that the SRM with regime-switching goes a long way in explaining the shift in bond return predictability documented in Section 2, whereas the extension with a permanent break performs even worse than the standard SRM.

4.1 Nonlinear Factor Dynamics in the SRM

As discussed in Section 2, there are at least two interpretations of the recent shift in bond return predictability. If we believe that the shift is temporary, as implied by the threshold regression in equation (5), it seems natural to accommodate regime-switching in the $\mathbb{P}$ dynamics of the pricing factors. That is, to replace equation (11) with

$$\mathbf{x}_{t+1} = \mathbf{h}_0^{(1)} \mathcal{I}_{\{r_t \geq c\}} + \mathbf{h}_0^{(2)} \mathcal{I}_{\{r_t < c\}} + \left(\mathbf{H}_x^{(1)} \mathcal{I}_{\{r_t \geq c\}} + \mathbf{H}_x^{(2)} \mathcal{I}_{\{r_t < c\}} \right) \mathbf{x}_t + \boldsymbol{\Sigma} \boldsymbol{\varepsilon}_{t+1}^{\mathbb{P}}, \tag{15}$$

where both the level and the persistence of the factors may change when r_t falls below the threshold c. In our application, we consider a threshold of 1 percent, which means that the two regimes implied by the model are consistent with the two regimes for the predictability regression in equation (2). The short rate and the $\mathbb{Q}$ dynamics remain given by equations (8) and (10), respectively, meaning that bond prices also for this extended model can be computed using the second-order approximation of Priebsch (2013). We refer to this modified shadow rate model with regime-switching dynamics as the R-SRM.

Another possibility is to interpret the shift in bond return predictability as being permanent, as implied by equation (6). In this case, it seems more natural to allow for a permanent break in the $\mathbb{P}$ dynamics of the pricing factors. That is, to replace equation (11) with

$$\mathbf{x}_{t+1} = \mathbf{h}_0^{(1)} \mathcal{I}_{\{t < \tau\}} + \mathbf{h}_0^{(2)} \mathcal{I}_{\{t \geq \tau\}} + \left(\mathbf{H}_x^{(1)} \mathcal{I}_{\{t < \tau\}} + \mathbf{H}_x^{(2)} \mathcal{I}_{\{t \geq \tau\}} \right) \mathbf{x}_t + \boldsymbol{\Sigma} \boldsymbol{\varepsilon}_{t+1}^{\mathbb{P}}, \tag{16}$$

where both the level and the persistence of the factors are allowed to change after the τth month in the sample. In our application, we set $\tau = 225$ to obtain a break in October 2008, which implies that the two regimes in the model are consistent with the two regimes for the predictability regression in equation (2). We again compute bond yields using the second-order approximation of Priebsch (2013), as the short rate and the $\mathbb{Q}$ dynamics remain given by equations (8) and (10), respectively. We refer to this modified shadow rate model with a permanent break as the B-SRM.

These two modifications of the standard SRM do not change the cross-sectional relationship between the pricing factors and yields, but correspond to introducing nonlinearities in the market prices of risk given by equation (9). To guide intuition for why such an extension of the standard SRM has the potential to explain the shift in equation (2), consider the

ATSM where $r_t = s_t$ with closed-form expressions for bond yields and returns. The slope coefficient in equation (2) with $h = 1$ is then

$$\beta_1^{(n)} = \frac{(n-1)\,\mathbf{b}_{n-1}\boldsymbol{\lambda}_x\mathbb{V}\left[\mathbf{x}_t\right]\left(\frac{1}{n}\mathbf{b}_n+\boldsymbol{\beta}'\right)'}{\left(\frac{1}{n}\mathbf{b}_n+\boldsymbol{\beta}'\right)\mathbb{V}\left[\mathbf{x}_t\right]\left(\frac{1}{n}\mathbf{b}_n+\boldsymbol{\beta}'\right)'}, \tag{17}$$

where $\boldsymbol{\lambda}_x = \mathbf{H}_x - (\mathbf{I}-\boldsymbol{\Phi})$ are the loadings of the market prices of risk on $\mathbf{x}_t$ and yields are given by $y_t^{(n)} = -\frac{1}{n}\left(a_n + \mathbf{b}_n'\mathbf{x}_t\right)$ (see, for example, Duffee (2002) and Joslin et al. (2011)). Thus, allowing for a shift in $\mathbf{H}_x$ without changing $\boldsymbol{\Phi}$ induces a shift in $\boldsymbol{\lambda}_x$ that may explain the observed shift in the predictability regression in equation (2). Our approach is therefore related to Bansal and Zhou (2002), who introduce regime-switching dynamics into the prices of risk in an ATSM, and to Giacoletti et al. (2018), who allow the parameters of the prices of risk in an ATSM to vary period-by-period.

An even more flexible extension of the standard SRM than implied by the R-SRM and the B-SRM would also allow for regime-switching in i) the short rate in equation (8), ii) the $\mathbb{Q}$ dynamics in equation (10), and iii) the conditional covariance matrix $\boldsymbol{\Sigma}$, as considered in the regime-switching model of Dai, Singleton and Yang (2007). Such a model would display greater flexibility in fitting the cross section of bond yields than implied by the standard SRM and the two extensions we propose. However, previous studies have shown that the standard SRM is able to fit the cross section of bond yields closely, both when yields are away from and close to the ZLB (see, for example, Christensen and Rudebusch (2015) and Andreasen and Meldrum (2018)). Hence, any improvements to the cross-sectional fit in a more flexible SRM seems likely to be economically marginal and would increase the risk of over-fitting the data. We therefore prefer to consider more parsimonious models that only allow for structural changes in the $\mathbb{P}$ dynamics.

Estimation of the R-SRM and B-SRM proceeds as described in Section 3.2, except that step 2 of the SR approach is extended by running either a threshold vector autoregression (for the R-SRM) or an autoregression with a break (for the B-SRM). This implies that the additional parameters introduced in the R-SRM and the B-SRM in comparison to the standard SRM are estimated in closed form and without any additional computational cost. The details are described in Appendix B.

4.2 In-Sample Results

We start our comparison of the standard SRM, the R-SRM, and the B-SRM by examining their in-sample fit to bonds yields.[11] The results reported in Table 2 show that the root mean squared errors (RMSEs) between observed and model-implied yields are almost identical for the three models and below 10 basis points at all maturities. This result is fairly unsurprising because all the models have the same short rate specification and $\mathbb{Q}$ dynamics, meaning that the cross-sectional mapping between the pricing factors and bond yields only differs slightly due to different estimates of $\boldsymbol{\Sigma}$ and $\boldsymbol{\Phi}$.

Table 2: In-Sample Fit to Bond Yields

This table reports the root mean squared errors (RMSEs) in annualized basis points between actual and model-implied yields at selected maturities at step 3 of the SR approach.

	Maturity (in months)				
	6	12	24	60	120
SRM	8.28	9.93	2.60	5.37	9.81
R-SRM	8.28	9.93	2.59	5.37	9.78
B-SRM	8.28	9.93	2.59	5.37	9.75

However, the models have different implications for the time-series dynamics of bond yields due to the nonlinearities introduced in the $\mathbb{P}$ dynamics. For the R-SRM, we find a significant shift in the persistence of the pricing factors, as we reject the null hypothesis of $\mathbf{H}_x^{(1)} = \mathbf{H}_x^{(2)}$ in a Wald test (p-value = 0.000). On the other hand, the intercepts in (15) do not appear to shift when we approach the ZLB (p-value = 0.34). For the B-SRM, we get qualitatively the same results, as we find a significant shift in the persistence of the pricing factors but not in their intercepts after the break in October 2008.

[11] We provide parameter estimates for the three models in the Online Appendix.

One way to illustrate these differences in $\mathbb{P}$ dynamics is to compare the model-implied term premiums $TP_t^{(n)}$, which are defined as

$$TP_t^{(n)} = y_t^{(n)} - \frac{1}{n}\sum_{i=0}^{n-1} \mathbb{E}_t\left[r_{t+i}\right]. \tag{18}$$

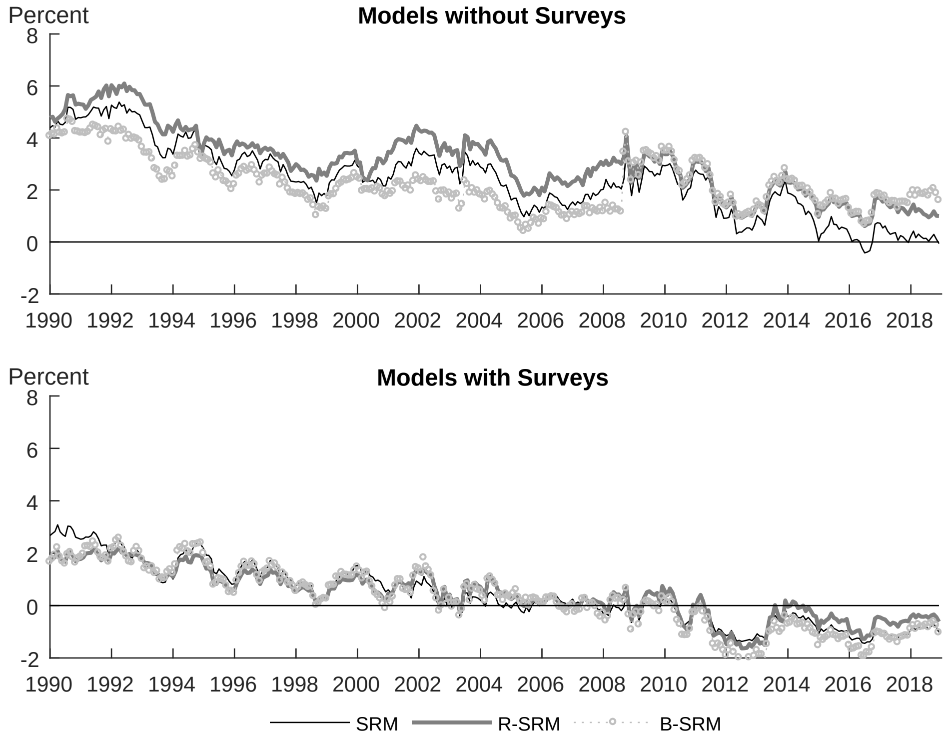
Given that the three models provide the same close fit to bond yields, any differences in $TP_t^{(n)}$ must primarily reflect different model-implied projections of the short rate as determined by the $\mathbb{P}$ dynamics of the pricing factors. The top chart in Figure 4 shows the 10-year term premiums implied by the three models. Over much of the sample, the three term premium estimates are highly correlated. However, there are also important differences. For instance, the term premium implied by the R-SRM before 2008 is somewhat higher when compared to the term premium implied by the B-SRM. This is because the R-SRM assigns the last three years of the sample (that is, December 2015 to December 2018) with low interest rates to the regime away from the ZLB, whereas these observations in the B-SRM are in the post-break regime. This implies that the R-SRM for the pre-ZLB regime has slightly lower expected short rate paths than the B-SRM, which then generates higher term premiums in the R-SRM than in the B-SRM. Another interesting difference appears toward the end of the sample, where the B-SRM implies higher term premiums than either the standard SRM or the R-SRM. This result arises because the model with the permanent break implies a permanent downward shift in the long-run mean of the short rate, which generates a lower expected short rate path in the B-SRM and hence higher term premiums than in the two other models.

4.3 Matching the Shift in Bond Return Predictability

We next consider whether the two extensions of the standard SRM can match the shift in bond return predictability documented in Section 2. For the R-SRM, it is natural to view this shift as being generated by the proximity of the short rate to the ZLB. Thus, we adopt

Figure 4: Model-Implied 10-Year Term Premiums

This figure reports model-implied 10-year term premiums. For the R-SRM, we compute the expected average short rate over the next 10 years in a given period using Monte Carlo integration with 1 million draws.

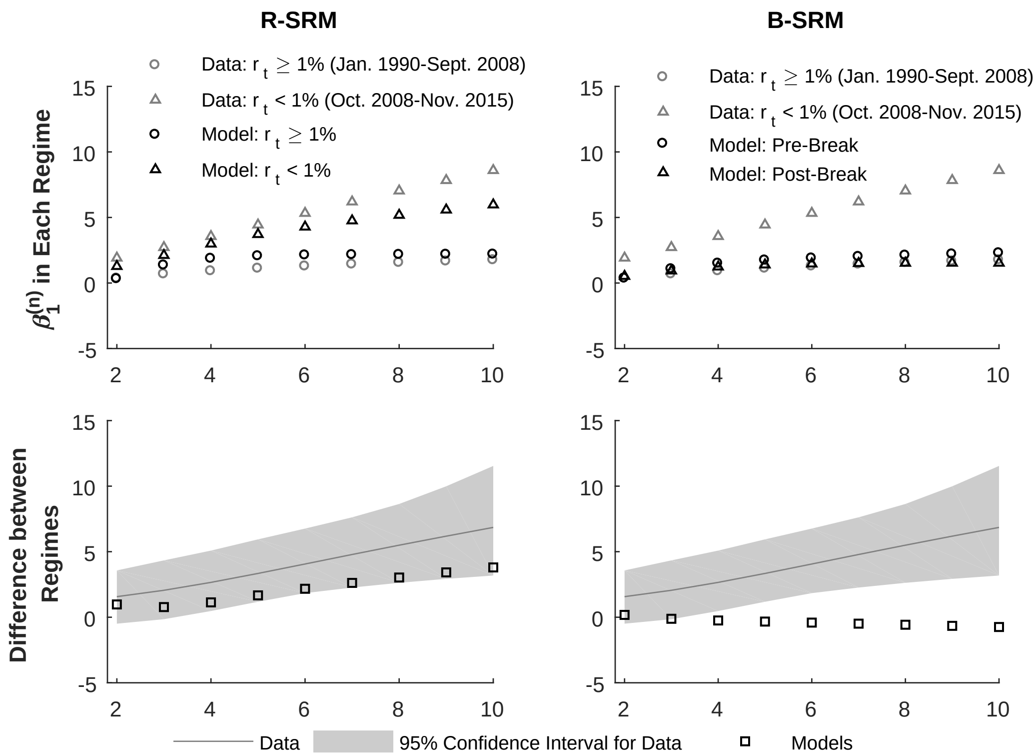


the same procedure as for the standard SRM in Section 3.3 and estimate equation (5) with a threshold of 1 percent using a simulated sample of 1 million observations from the R-SRM. The upper left chart in Figure 5 shows that the R-SRM preserves the satisfying ability of the standard SRM to generate the desired degree of return predictability in the pre-ZLB period when $r_t \geq 0.01$. Even more encouraging is the ability of the R-SRM to generate a much stronger relationship between the yield spread and excess bond returns close to the ZLB with $r_t < 0.01$ than in the pre-ZLB period. This shift in return predictability is not as strong as seen in the data, but the lower left chart in Figure 5 shows that the increase in the slope coefficients lie within the 95 percent confidence interval for the sample moment at all maturities.[12]

[12] In the Online Appendix we show that repeating the simulation with a lower threshold of 0.25 percent does not materially change these results.

Figure 5: Bond Return Predictability: Extended SRMs

The left charts show the ability of the R-SRM to match the shift in the slope coefficients in equation (5) with $h = 12$ and $c = 0.01$ using a simulated sample of 1 million observations. The right charts show the ability of the B-SRM to match the shift in the slope coefficients in equation (6) with $h = 12$ using a simulated sample of 1 million observations with $\tau = 500,001$. The corresponding data moments are from equation (2) with $h = 12$ when estimated from January 1990 to September 2008 (excluding December 2003) where $r_t \geq 0.01$, and when estimated from October 2008 to November 2015 where $r_t < 0.01$. All x-axes show the maturity in years.



For the B-SRM, it seems natural to view the shift in bond return predictability as the result of a permanent break. Thus, we obtain the model-implied predictability coefficients for the B-SRM by estimating equation (6) on a simulated sample of 1 million observations with a break at $\tau = 500,001$ to generate two regimes of $500,000$ observations. The top right chart in Figure 5 shows that the B-SRM matches the degree of bond predictability in the pre-ZLB sample, but the model is simply unable to explain the shift in bond predictability after the break in 2008.

In summary, the R-SRM goes a long way in explaining the stronger link between the yield spread and excess bond returns during the recent ZLB period, meaning that this shift in return predictability can be attributed to a temporary re-pricing of risk. The results are less encouraging for the B-SRM with a permanent break, as this model is unable to generate stronger bond predictability after the break in 2008.

5 Matching Survey Expectations

This section shows that the conclusions of the previous section are robust to including long-horizon surveys of short-term interest rate expectations in the data set used to estimate the models. In Section 5.1 we show that none of the three models in Sections 3 and 4 can match these surveys when the models are estimated using only data on bond yields as done above, which suggests that the models may not provide plausible implications for bond risk premiums. As a robustness check, we therefore consider whether our key conclusions are affected by including surveys in the data set used to estimate the models. Section 5.2 explains how survey expectations can be included when estimating the three SRMs using the SR approach. Section 5.3 shows that this modification allows the three models to match survey expectations reasonably well and that our key conclusion from Section 4 is unaffected, that is, only the R-SRM is able to match the documented shift in bond return predictability.

5.1 Out-of-Sample Fit to Surveys

We explore the ability of the three SRMs to match the expected short rate in 5 years and the average expected short rate from 5 to 10 years ahead. That is, we focus on long-horizon survey expectations and disregard more short-term expectations. This is because it is unclear whether surveys represent the mean or the mode of respondents' probability distributions and the model-implied mean and mode are likely to be materially different near the ZLB.[13] The distinction between mean and mode should matter less at long horizons, where the ZLB is likely to have a much smaller effect on the conditional model-implied short rate distributions. Another benefit of including long-term short rate expectations is that they serve as observable proxies for i_t^*, and hence allow us to explore whether the models match the evolution in the natural nominal short rate.

Our empirical source for these long-horizon expectations is the Blue Chip Economic Indicators survey, where repondents every 6 months report a point expectation for the 3-month Treasury bill yield at horizons up to 11 years ahead. The forecast horizons are not the same in each edition of the survey, so we linearly interpolate the mean expectations across respondents at different horizons to get a measure of the average expected 3-month Treasury bill yield at a constant horizon of 5 years and from 5 to 10 years ahead - that is, $\mathbb{E}_t\left[y_{t+59}^{(3)}\right]$ and $\mathbb{E}_t\left[\frac{1}{60}\sum_{i=60}^{119} y_{t+i}^{(3)}\right]$, respectively. These expectations are evaluated exactly in the SRM and B-SRM and by Monte Carlo integration in the R-SRMs. We exploit the close correlation between $\mathbb{E}_t\left[y_{t+i}^{(3)}\right]$ and $\mathbb{E}_t\left[r_{t+i}\right]$ at long forecast horizons to approximate $\mathbb{E}_t\left[y_{t+59}^{(3)}\right]$ by $\mathbb{E}_t\left[r_{t+59}\right]$ and $\mathbb{E}_t\left[\frac{1}{60}\sum_{i=60}^{119} y_{t+i}^{(3)}\right]$ by $\mathbb{E}_t\left[\frac{1}{60}\sum_{i=60}^{119} r_{t+i}\right]$. This substantially reduces the computational burden, because we avoid to repeatedly solve for the three-month yield in the SRMs.

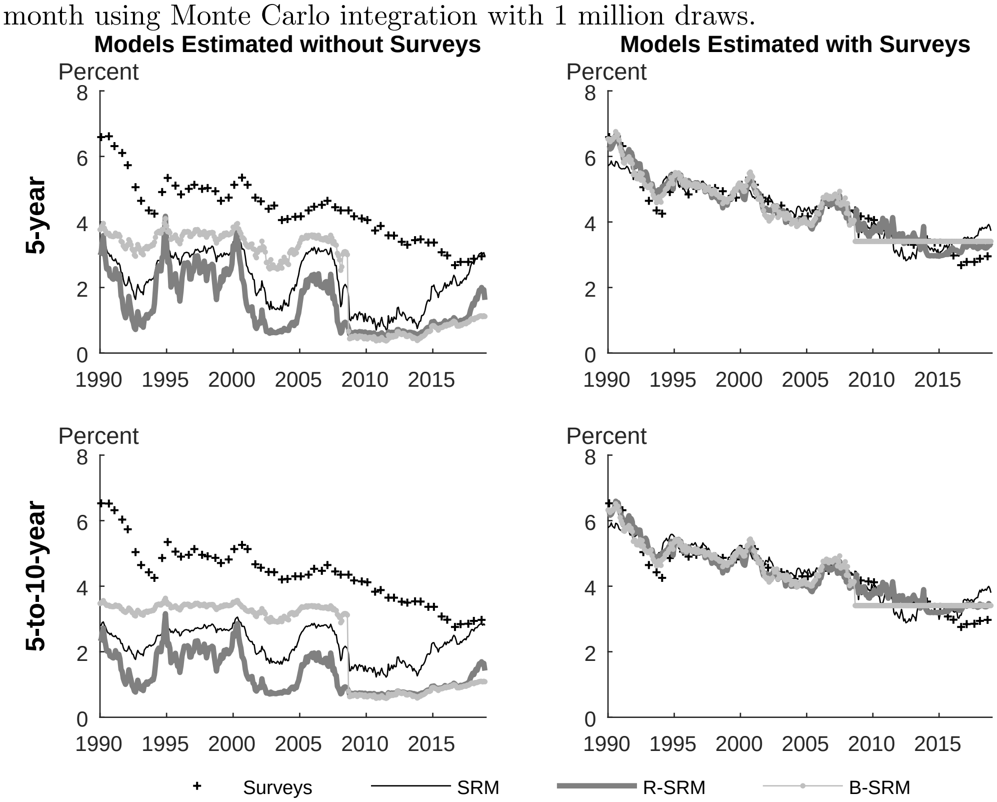
The left charts in Figure 6 show the survey-based measure at the 5 year horizon (top left) and at the 5-to-10 year ahead (bottom left) along with the corresponding expectations in the three SRMs. The survey-based measures (the black plus signs) generally decline at both

[13] The SRM and extensions of it are only able to address differences between the mean and mode of the conditional short rate distribution to a limited extent because the model implies that the mean must always lie above the mode.

horizons throughout the sample, from about 6.5 percent to about 3 percent. In contrast, the short rate expectations implied by the standard SRM (the black line) are typically lower than 3 percent, have no clear downward trend, and display only a weak correlation with the two surveys. The short rate expectations from the R-SRM (the dark gray line) are generally even further below the survey-based measure but are otherwise very similar to those from the standard SRM. Finally, the B-SRM (the light gray line) implies long-horizon expectations that are relatively stable (with the obvious exception of the drop at the time of the break in October 2008), and these expectations differ therefore also substantially from the surveys.

Figure 6: Matching Long-Horizon Short Rate Surveys

This figure shows the ability of the SRMs to match the expected short rate in 5 years and the average expected short rate from 5 to 10 years ahead as implied by responses to Blue Chip Economic Indicators surveys. The left charts show the results for models estimated using only a panel of bond yields, while the right charts show the results for models estimated using a data set that includes a panel of bond yields and surveys. For the R-SRM, we compute the expected average short rate at each month using Monte Carlo integration with 1 million draws.



Thus, none of the considered models are able to produce plausible short rate expectations at long horizons, and hence match the downward trend in the survey-based measure of the natural nominal short rate, when these models are estimated solely based on a panel of bond yields.

5.2 Incorporating Surveys into the SR Approach

The reason that the SRMs do not produce short rate expectations at long horizons that match those from surveys is most likely explained by the fairly short span of our sample (from 1990 to 2018). Given the high persistence in bond yields, it is well-known that short samples imply estimates of the $\mathbb{P}$ dynamics that may be biased and subject to substantial estimation uncertainty (see Bauer et al. (2012), Kim and Orphanides (2012), and Wright (2014)). The fact that the models do not match surveys well is potentially concerning because if the models do not produce plausible short rate expectations, we should presumably also be skeptical about their implications for the term premiums reported in Figure 4.

One well-known way to mitigate these problems is to incorporate survey data when estimating DTSMs, as proposed by Kim and Orphanides (2012). We incorporate surveys at step 2 of the SR approach when estimating the $\mathbb{P}$ parameters in $\boldsymbol{\theta}_2$, but not at step 1 and 3 when determining the $\mathbb{Q}$ parameters and the pricing factors, because they are already accurately identified from the considered panel of bond yields. Although this assumption implies a degree of simplification, it means that we can continue to separate the $\mathbb{Q}$ and $\mathbb{P}$ parameters, which is the key computational advantage of the SR approach.

More formally, at step 2 of the SR approach, we allow surveys to be measured with errors, which we denote by $\eta_t^{(5y)}$ and $\eta_t^{(5-10y)}$ for the short rate expectations in 5 years and from 5 to 10 years ahead, respectively. Each of these errors are assumed to be independent and identically distributed with zero mean and standard deviation σ_η. For the standard SRM and the R-SRM, we augment the moment conditions for the $\mathbb{P}$ parameters by the first and second moments of $\eta_t^{(5y)}$ and $\eta_t^{(5-10y)}$ when $r_t \geq 0.01$ and when $r_t < 0.01$. That is, a total of 8 additional moment conditions. For the B-SRM, we include the first and second moments of $\eta_t^{(5y)}$ and $\eta_t^{(5-10y)}$ both before and after the break point, which also imply 8 additional moment conditions. Further details are provided in Appendix C.

For our application, we assume that the measurement errors in surveys have a standard deviation of 10 basis points, i.e. $\sigma_\eta = 0.1$, which seems reasonable given the fairly stable evolution in the considered surveys.[14]

5.3 Estimation Results with Surveys

The in-sample fit to bond yields is basically unaffected by the inclusion of surveys, and the three SRMs therefore display the same satisfying fit as reported in Table 2.[15] For the R-SRM we once again find a significant shift in the persistence of the pricing factors but not in the intercepts. The shift in the persistence of the pricing factors is also significant in the B-SRM (p-value = 0.000), but now we also find a significant shift in the intercepts (p-value = 0.047) whereas it was insignificant without surveys.

The right charts in Figure 6 show that the models track the overall evolution in surveys reasonably well once they are included in the estimation. That is, the standard SRM now matches the downward trend in long-term short rate expectations, and hence this empirical proxy for the equilibrium short rate. But it is also evident from Figure 6 that this model generates too high short rate expectations at the end of our sample, up to 1 percentage point higher than in the survey data. Figure 6 further shows that the R-SRM and the B-SRM are also able to match the downward trend in surveys, and that these models produce less elevated short rate expectations at the end of our sample when compared to the standard SRM.

The bottom panel of Figure 4 reveals that the inclusion of surveys in the estimation reduces the level of the 10-year term premium substantially in all three SRMs. The main

[14] A preliminary analysis also considered $\sigma_\eta = 0.75$ as in Kim and Orphanides (2012), but this implied a very low weight to surveys in our estimation, and hence fairly small effects of including surveys at the second step of the SR approach. The intention of this paper is not to suggest that $\sigma_\eta = 0.1$ is the optimum calibration but to show that our key conclusions about return predictability are not changed once the models fit the broad movements in the surveys.

[15] We provide parameter estimates for the three models including surveys in the estimation in the Online Appendix.

differences between the three models now only appear at the end of our sample, where the R-SRM implies a higher term premium than seen in the B-SRM.

Figure 7 finally explores the ability of the SRMs to match the shift in bond return predictability once surveys are included in the estimation. The encouraging finding is that the R-SRM remains able to match the shift, whereas both the standard SRM and the B-SRM fail to generate a sufficiently large shift in bond return predictability. In summary, we conclude that the models are capable of matching long-term short rate expectations reasonably closely without changing our fundamental conclusion that only the R-SRM is able to match the documented shift in bond return predictability at the ZLB.

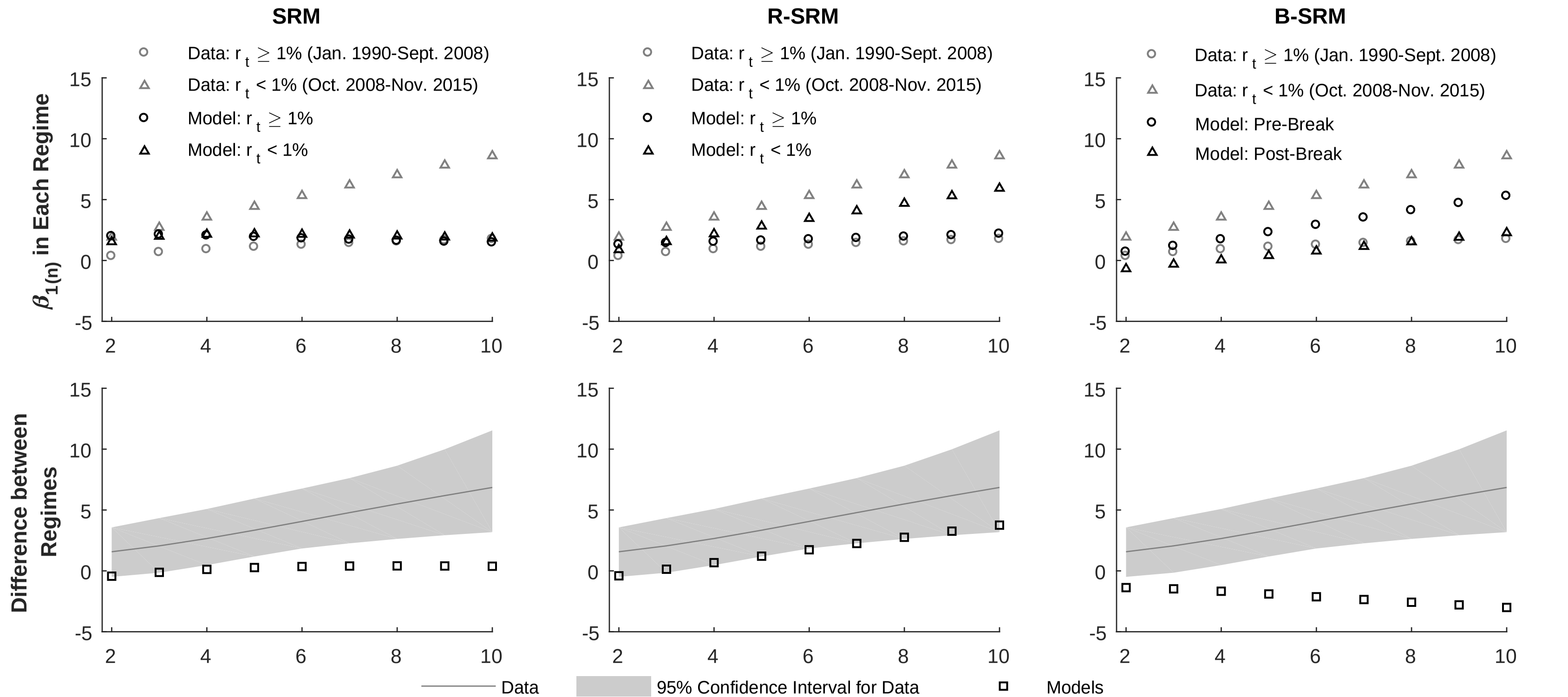
One caveat to this results is that all of the models, including the R-SRM, struggle to explain the low short rate expectations since 2015. That result suggests that we should not necessarily conclude that the *levels* of bond yields will revert back to where they were pre-crisis, once the short rate has moved away from the ZLB. In unreported results, we show that allowing for a permanent shift in the long-run mean of the factors in the R-SRM in 2015 allows the model to better match the long-horizon surveys. This permanent shift would be consistent with recent studies (such as Summers (2015), Laubach and Williams (2016), Del Negro et al. (2017), and Christensen and Rudebusch (2019)) arguing that the long-run natural rate of interest has undergone a persistent decline in recent years.

6 Alternative Timing of a Permanent Break

We have so far assumed that the break in the B-SRM coincides with the start of the ZLB period based on the instability of the predictability regression in equation (2). But a different timing of the break point may be more suitable for the $\mathbb{P}$ dynamics in the B-SRM and this may improve its performance. In this section, we therefore demonstrate that the performance of the B-SRM does not improve by changing the timing of the break point according to an alternative criterion.

Figure 7: Bond Return Predictability: SRMs with Surveys

The left and middle charts show the ability of the standard SRM and R-SRM, respectively, to match the shift in the slope coefficients in equation (5) with $h = 12$ and $c = 0.01$ using a simulated sample of 1 million observations. The right charts show the ability of the B-SRM to match the shift in the slope coefficients in equation (6) with $h = 12$ using a simulated sample of 1 million observations with $\tau = 500,001$. The corresponding data moments are from equation (2) with $h = 12$ when estimated from January 1990 to September 2008 (excluding December 2003) where $r_t \geq 0.01$, and when estimated from October 2008 to November 2015 where $r_t < 0.01$. All x-axes show the maturity in years.



The alternative approach for determining the timing of the break point is to examine the historical evolution of the estimated pricing factors directly. Here, we exploit a convenient property of the SR approach that the pricing factors are estimated non-parametrically without any assumption about the $\mathbb{P}$ dynamics, unlike typical QML estimators based on Kalman filtering.[16] This property of the SR approach implies that we can directly examine the $\mathbb{P}$ dynamics for a structural break using the pricing factors estimated at step 1 of the SR approach. Given that the timing of the break is unknown, we again compute Chow test statistics using a range of candidate breakpoints as proposed by Quandt (1960). For a candidate break point τ the Chow test statistic for the VAR model in equation (11) is given by

$$C(\tau) = \frac{\left(|\mathbf{\Sigma}_{1,T}| - |\mathbf{\Sigma}_{1,\tau-1} + \mathbf{\Sigma}_{\tau,T}|\right)/k}{|\mathbf{\Sigma}_{1,\tau-1} + \mathbf{\Sigma}_{\tau,T}|\,/\,(T-2k)}, \tag{19}$$

where $\mathbf{\Sigma}_{i,j}$ is the estimated covariance matrix of the innovations from equation (11) estimated on a sample from period i to period j and $k = n_x(n_x+1)$ is the number of parameters in the model without a break. We again assume that both the pre- and post-break samples contain at least 15 percent of the total number of observations. The solid line in Figure 8 shows the resulting Chow test statistics. These test statistics are generally higher during the ZLB period than before but the maximum test statistic does not come until November 2013. This result suggests that a break in the $\mathbb{P}$ dynamics during the fall of 2013 may be more appropriate than during the fall of 2008. One possible explanation for this result could be that long-horizon expectations of future interest rates did not decline until relatively recently, as discussed in Section 5. Another is that several bond yields did not appear to be constrained by the ZLB until after 2010, as shown in Swanson and Williams (2014).

Given this finding, we consider whether a version of the B-SRM with a break in November 2013 ($\tau = 287$) is able to match the documented shift in bond return predictability. To give the model the best possible chance of matching this shift, we also re-estimate the regression coefficients in the data using the same sample split. Hence, the gray markers on the top

[16] Here, we refer to the step 1 estimates of the pricing factors in the SR approach.

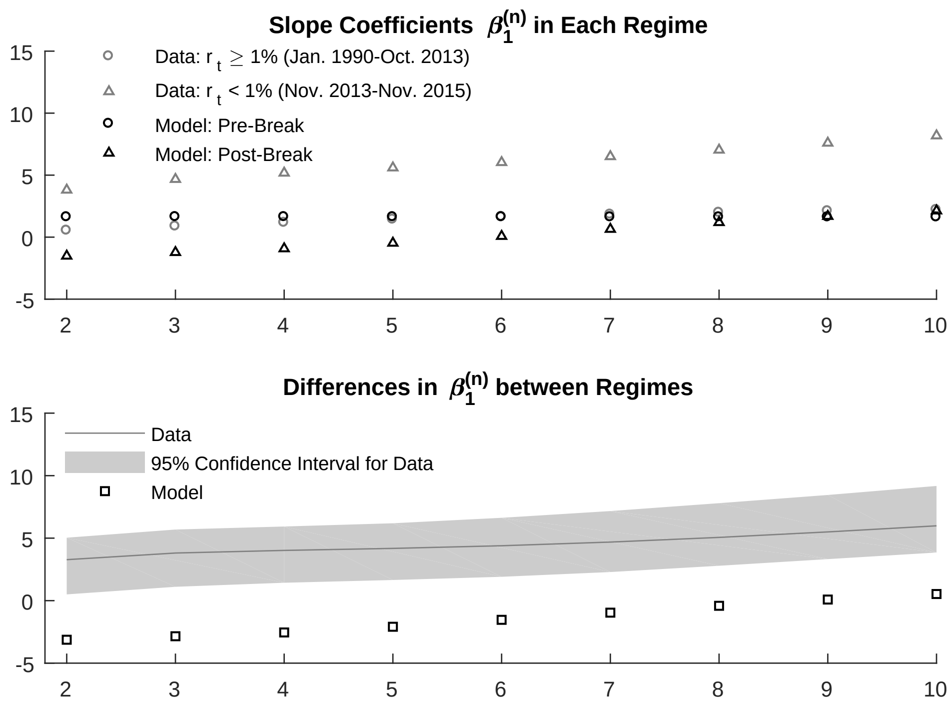
Figure 8: Chow Tests for a Break in the Pricing Factors
This chart reports the Chow test statistic for breaks in the estimated pricing factors in the standard SRM (first step estimates) using at least 15 percent of the observations for the pre- and post-break sample.

chart in Figure 9 show the slope estimates in equation (2) for the sub-sample from January 1990 through October 2013 ("Regime 1") and the sub-sample from November 2013 through December 2018 ("Regime 2"). The black markers are for the B-SRM estimated using surveys on short rate expectations as described in Section 5. The clear message from Figure 9 is that this alternative timing for the break point does not change our conclusion from above, as the B-SRM also in this case is unable to match the recent shift in bond return predictability.

7 Conclusion

This paper documents a shift in the predictability of excess bond returns during the recent ZLB period that a standard SRM cannot replicate. This shows that simply enforcing the ZLB by truncating the short rate at zero in a Gaussian model is insufficient to properly

Figure 9: Bond Return Predictability: Different Break Point in the B-SRM
The top chart shows the slope coefficient in equation (6) with $h = 12$ using a simulated sample of 1 million observations with $\tau = 500,001$ using an estimated version of the B-SRM with a break in November 2013. The corresponding data moments are from equation (2) with $h = 12$ when estimated from January 1990 to October 2013, and when estimated from November 2013 to December 2018. The bottom chart shows the difference in the slope estimates at a given maturity. The 95 percent confidence interval for these differences are computed using a block bootstrap with 5,000 repetitions and a block window of 24 months. In each bootstrap sample it is required that there are at least 50 observations in each regime. The x-axes reports maturity in years, and all yields are end-month from Gürkaynak et al. (2007)



capture the change in bond yield dynamics that occurred at the ZLB. We find that this new predictability result is consistent with a SRM that allows the time-series dynamics of the pricing factors to change at the lower bound. In contrast, a SRM that introduces a permanent break in the pricing factors in 2008 is not able to explain this shift in bond return predictability at the ZLB. These conclusions are robust to including surveys of future short-term interest rate expectations in the data set used to estimate the models and to determining the timing of a potential permanent break directly from the estimated factors.

Our results suggest that once the short rate has risen sufficiently far from the ZLB, bond risk premiums will behave as they did before the financial crisis. That may be consistent with factors that were largely contained to the ZLB episode, such as the Federal Reserve's unconventional monetary policies, having a temporary effect on the predictability of excess bond returns. However, we also note that none of the models we consider are able to match the low level of long-horizon surveys since 2015, which is consistent with various recent studies that point to a more persistent recent decline in the long-horizon natural interest rate.

References

Andreasen, M. M. and Christensen, B. J. (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.

Andreasen, M. M., Engsted, T., Möller, S. V. and Sander, M. (2016), 'Bond Market Asymmetries across Recessions and Expansions: New Evidence on Risk Premia', *Working paper* .

Andreasen, M. M. and Meldrum, A. C. (2018), 'A Shadow Rate or a Quadratic Policy Rule? the Best Way to Enforce the Zero Lower Bound in the United States', *Journal of Financial and Quantitative Analysis* **Forthcoming**.

Andrews, D. W. K. (1993), 'Tests for Parameter Instability and Structural Change With Unkown Change Point', *Econometrica* **61**(4), 821–856.

Bansal, R. and Shaliastovich, I. (2013), 'A long-run risk explanation of predictability puzzles in bond and currency markets', *Review of Financial Studies* **Vol. 26**, 1–33.

Bansal, R. and Zhou, H. (2002), 'Term Structure of Interest Rates with Regime Shifts', *Journal of Finance* **57**(5).

Bauer, M. D. and Rudebusch, G. D. (2016), 'Monetary Policy Expectations at the Zero Lower Bound', *Journal of Money, Credit, and Banking* **48**(7), 1439–1465.

Bauer, M. D. and Rudebusch, G. D. (2019), 'Interest Rates Under Falling Stars', *Federal Reserve Bank of San Francisco Working Paper 2017-16* .

Bauer, M. D., Rudebusch, G. D. and Wu, J. C. (2012), 'Correcting Estimation Bias in Dynamic Term Structure Models', *Journal of Business and Economic Statistics* **30**(3), 454–467.

Black, F. (1995), 'Interest Rates as Options', *Journal of Finance* **50**(5), 1371–1376.

Bonis, B., Ihrig, J. and Wei, M. (2017), 'Projected Evolution of the SOMA Portfolio and the 10-year Treasury Term Premium Effect', *FEDS Notes, Washington: Board of Governors of the Federal Reserve System,* .

Campbell, J. Y. and Shiller, R. J. (1991), 'Yield Spreads and Interest Rate Movements: A Bird's Eye View', *Review of Economic Studies* **58**(3), 495–514.

Christensen, J. H. E. and Rudebusch, G. D. (2015), 'Estimating Shadow-Rate Term Structure Models with Near-Zero Yields', *Journal of Financial Econometrics* **13**(2), 226–259.

Christensen, J. H. E. and Rudebusch, G. D. (2019), 'A New Normal for Interest Rates? Evidence from Inflation-Indexed Debt', *Federal Reserve Bank of San Francisco Working Paper Series* .

Cieslak, A. and Povala, P. (2015), 'Expected Returns in Treasury Bonds', *Review of Financial Studies* **28**(10), 2859–2901.

Cochrane, J. H. and Piazzesi, M. (2005), 'Bond Risk Premia', *American Economic Review* **95**(1), 138–160.

Dai, Q. and Singleton, K. J. (2002), 'Expectation Puzzles, Time-Varying Risk Premia, and Affine Models of the Term Structure', *Journal of Financial Economics* **63**(3), 415–441.

Dai, Q., Singleton, K. J. and Yang, W. (2007), 'Regime Shifts in a Dynamic Term Structure Model of U.S. Treasury Bond Yields', *Review of Financial Studies* **20**(5), 1669–1706.

D'Amico, S., English, W., Lopez-Salido, D. and Nelson, E. (2012), 'The Federal Reserve's Large-Scale Asset Purchase Programmes: Rationale and Effects', *Economic Journal* **122**, F415–F446.

Del Negro, M., Giannone, D., Giannoni, M. P. and Tambalotti, A. (2017), 'Safety, Liquidity, and the National Rate of Interest', *Federal Reserve Bank of New York, Staff Reports No. /812* pp. 1–52.

Duffee, G. (2002), 'Term Premia and Interest Rate Forecast in Affine Models', *Journal of Finance* **57**(1), 405–443.

Fama, E. F. and Bliss, R. R. (1987), 'The Information in Long-Maturity Forward Rates', *American Economic Review* **77**(4), 680–692.

Giacoletti, M., Laursen, K. T. and Singleton, K. J. (2018), 'Learning and Risk Premiums in an Arbitrage-free Term Structure Model', *Stanford Graduate School of Businesss Working Paper* **3670**.

Gürkaynak, R. S., Sack, B. and Wright, J. H. (2007), 'The U.S. Treasury Yield Curve: 1961 to the present', *Journal of Monetary Economics* **54**(8), 2291–2304.

Hamilton, J. D. and Wu, J. C. (2012), 'Identification and Estimation of Gaussian Affine Term Structure Models', *Journal of Econometrics* **168**(2), 315–331.

Joslin, S., Priebsch, M. and Singleton, K. J. (2014), 'Risk Premiums in Dynamic Term Structure Models with Unspanned Macro Risk', *Journal of Finance* **69**(3), 453–468.

Joslin, S., Singleton, K. J. and Zhu, H. (2011), 'A New Perspective on Gaussian Dynamic Term Structure Models', *Review of Financial Studies* **24**(3), 926–970.

Kim, D. H. and Orphanides, A. (2012), 'Term Structure Estimation with Survey Data on Interest Rate Forecasts', *Journal of Financial and Quantitative Analysis* **47**(1), 241–272.

Kim, D. H. and Singleton, K. J. (2012), 'Term Structure Models and the Zero Bound: An Empirical Investigation of Japanese Yields', *Journal of Econometrics* **170**(1), 32–49.

Kung, H. (2015), 'Macroeconomic Linkages between Monetary Policy and the Term Structure of Interest Rates', *Journal of Financial Economics* **115**(1), 42–57.

Laubach, T. and Williams, J. C. (2016), 'Measuring the National Rate of Interest Redux', *Business Economics* **51**(2), 57–67.

Li, C. and Wei, M. (2013), 'Term Structure Modeling with Supply Factors', *International Journal of Central Banking* **9**(1), 3–39.

Ludvigson, S. C. and Ng, S. (2009), 'Macro Factors in Bond Risk Premia', *Review of Financial Studies* **22**(12), 5027–5067.

Priebsch, M. A. (2013), 'Computing Arbitrage-Free Yields in Multi-Factor Gaussian Shadow-Rate Term Structure Models', *Board of Governors of the Federal Reserve System (U.S.) Finance and Economics Discussion Series 2013-63* .

Quandt, R. E. (1960), 'Tests of the Hypothesis That a Linear Regression System Obeys Two Separate Regimes', *Journal of the American Statistical Association* **55**(290), 324–330.

Rudebusch, G. D. and Wu, T. (2007), 'Accounting for a Shift in Term Structure Behavior with No-Arbitrage and Macro-Finance Models', *Journal of Money, Credit, and Banking* **39**(2-3), 395–422.

Stock, J. (1994), Unit Roots, Structural Breaks, and Trends, *in* R. Engle and D. McFadden, eds, 'Handbook of Econometrics', Amsterdam: Elsevier, pp. 2740–2843.

Summers, L. H. (2015), 'Demand Side Secular Stagnation', *American Economic Review: Papers and Proceedings* **105**(5), 60–65.

Swanson, E. T. and Williams, J. C. (2014), 'Measuring the Effect of the Zero Lower Bound on Medium- and Longer-Term Interest Rates', *American Economic Review* **104**(10), 3154–3185.

Wachter, J. A. (2006), 'A consumption-based model of the term structure of interest rates', *Journal of Financial Economics* **Vol. 79**, 365–399.

Wright, J. H. (2014), 'Term Premia and Iinflation Uncertainty: Empirical Evidence from an International Panel Dataset: Reply', *American Economic Review* **104**(1), 338–341.

Wu, J. C. and Xia, F. D. (2016), 'Measuring the Macroeconomic Impact of Monetary Policy at the Zero Lower Bound', *Journal of Money, Credit, and Banking* **48**(2-3), 253–291.

Appendix A: Step 2 in the SR Approach

To estimate $\boldsymbol{\theta}_2$ in Step 2 of the SR approach we follow Andreasen and Christensen (2015) and use the moment conditions

$$\begin{bmatrix} \mathbb{E}_t\left[\widehat{\boldsymbol{\varepsilon}}^{\mathbb{P}}_{t+1}\right] \\ vec\left(\mathbb{E}_t\left[\widehat{\boldsymbol{\varepsilon}}^{\mathbb{P}}_{t+1}\widehat{\mathbf{x}}_t\left(\boldsymbol{\theta}_1\right)'\right]\right) \\ vech\left(\mathbb{V}_t\left[\widehat{\boldsymbol{\varepsilon}}^{\mathbb{P}}_{t+1}\right]\right) \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ vec\left(\mathbb{C}_t\left[\mathbf{u}_{t+1},\mathbf{u}_t\right] - \mathbf{H}_x\mathbb{V}_t\left[\mathbf{u}_t\right]\right) \\ vech\begin{pmatrix} \mathbb{V}_t\left[\boldsymbol{\varepsilon}^{\mathbb{P}}_{t+1}\right] + \mathbb{V}_t\left[\mathbf{u}_t\right] + \mathbf{H}_x\mathbb{V}_t\left[\mathbf{u}_t\right]\mathbf{H}_x' \\ -\mathbb{C}_t\left[\mathbf{u}_{t+1},\mathbf{u}_t\right]\mathbf{H}_x' - \mathbf{H}_x\mathbb{C}_t\left[\mathbf{u}_t,\mathbf{u}_{t+1}\right] \end{pmatrix} \end{bmatrix}, \quad (20)$$

where $\widehat{\boldsymbol{\varepsilon}}^{\mathbb{P}}_t$ denotes the estimated values of $\boldsymbol{\varepsilon}^{\mathbb{P}}_t$ and $\mathbf{u}_t \equiv \widehat{\mathbf{x}}_t\left(\widehat{\boldsymbol{\theta}}_1^{step1}\right) - \mathbf{x}_t$. Consistent estimates of the second moments $\mathbb{V}_t\left[\mathbf{u}_{t+1}\right]$, $\mathbb{C}_t\left[\mathbf{u}_{t+1},\mathbf{u}_t\right]$, and $\mathbb{C}_t\left[\mathbf{u}_t,\mathbf{u}_{t+1}\right]$ follow from step 1 of the SR approach. Thus, $\boldsymbol{\theta}_2$ can be estimated using the generalized method of moments (GMM), which has a closed-form solution as shown in Andreasen and Christensen (2015)

Appendix B: Estimating the Extended SRMs

The parameters to be estimated at step 2 can again be written as $\boldsymbol{\theta}_2 = \begin{bmatrix} \boldsymbol{\theta}_{22}' & vech\left(\boldsymbol{\Sigma}\right)' \end{bmatrix}'$, but now we have

$$\boldsymbol{\theta}_{22} = \begin{bmatrix} \mathbf{h}_0^{(1)\prime} & vec\left(\mathbf{H}_x^{(1)}\right)' & \mathbf{h}_0^{(2)\prime} & vec\left(\mathbf{H}_x^{(2)}\right)' \end{bmatrix}'.$$

For the R-SRM, consider the revised set of moment conditions are

$$\begin{bmatrix} \mathbb{E}_t\left[\widehat{\boldsymbol{\varepsilon}}^{\mathbb{P}}_{t+1}\mathcal{I}_{\{r_t\geq c\}}\right] \\ \mathbb{E}_t\left[\widehat{\boldsymbol{\varepsilon}}^{\mathbb{P}}_{t+1}\mathcal{I}_{\{r_t< c\}}\right] \\ vec\left(\mathbb{E}_t\left[\widehat{\boldsymbol{\varepsilon}}^{\mathbb{P}}_{t+1}\widehat{\mathbf{x}}^{(1)}_t\left(\boldsymbol{\theta}_1\right)'\right]\right) \\ vec\left(\mathbb{E}_t\left[\widehat{\boldsymbol{\varepsilon}}^{\mathbb{P}}_{t+1}\widehat{\mathbf{x}}^{(2)}_t\left(\boldsymbol{\theta}_1\right)'\right]\right) \\ vech\left(\mathbb{V}_t\left[\widehat{\boldsymbol{\varepsilon}}^{\mathbb{P}}_{t+1}\right]\right) \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ vec\left(\mathbb{C}_t\left[\mathbf{u}^{(1)}_{t+1},\mathbf{u}^{(1)}_t\right] - \mathbf{H}^{(1)}_x\mathbb{V}_t\left[\mathbf{u}^{(1)}_t\right]\right) \\ vec\left(\mathbb{C}_t\left[\mathbf{u}^{(2)}_{t+1},\mathbf{u}^{(2)}_t\right] - \mathbf{H}^{(2)}_x\mathbb{V}_t\left[\mathbf{u}^{(2)}_t\right]\right) \\ vech\left(\mathbb{V}_t\left[\boldsymbol{\varepsilon}^{\mathbb{P}}_{t+1}\right] + \boldsymbol{\Omega}_{t+1}\right) \end{bmatrix}, \tag{21}$$

where $\widehat{\mathbf{x}}^{(1)}_t\left(\boldsymbol{\theta}_1\right) = \widehat{\mathbf{x}}_t\left(\widehat{\boldsymbol{\theta}}_1\right)\mathcal{I}_{\{r_t\geq c\}}$, $\widehat{\mathbf{x}}^{(2)}_t\left(\boldsymbol{\theta}_1\right) = \widehat{\mathbf{x}}_t\left(\widehat{\boldsymbol{\theta}}_1\right)\mathcal{I}_{\{r_t< c\}}$, $\mathbf{u}^{(1)}_t = \mathbf{u}_t\mathcal{I}_{\{r_t\geq c\}}$, $\mathbf{u}^{(2)}_t = \mathbf{u}_t\mathcal{I}_{\{r_t> c\}}$, and

$$\begin{aligned} \boldsymbol{\Omega}_{t+1} \;=\; & \mathbb{V}_t\left[\mathbf{u}_{t+1}\right] + \mathbf{H}^{(1)}_x\mathbb{V}_t\left[\mathbf{u}^{(1)}_t\right]\mathbf{H}^{(1)\prime}_x + \mathbf{H}^{(2)}_x\mathbb{V}_t\left[\mathbf{u}^{(2)}_t\right]\mathbf{H}^{(2)\prime}_x \\ & -\mathbb{C}_t\left[\mathbf{u}^{(1)}_{t+1},\mathbf{u}^{(1)}_t\right]\mathbf{H}^{(1)\prime}_x - \mathbf{H}^{(1)}_x\mathbb{C}_t\left[\mathbf{u}^{(1)}_t,\mathbf{u}^{(1)}_{t+1}\right] \\ & -\mathbb{C}_t\left[\mathbf{u}^{(2)}_{t+1},\mathbf{u}^{(2)}_t\right]\mathbf{H}^{(2)\prime}_x - \mathbf{H}^{(2)}_x\mathbb{C}_t\left[\mathbf{u}^{(2)}_t,\mathbf{u}^{(2)}_{t+1}\right]. \end{aligned}$$

Andreasen, Engsted, Möller and Sander (2016) provide the closed-form solution for $\boldsymbol{\theta}_2$ using these moment conditions. It is easy to verify that a similar result applies for the B-SRM, where regimes are defined based on the time point in the sample instead of the value of r_t.

Appendix C: The SR Approach with Surveys

The model parameters to be estimated at step 2 are the same as when surveys are not included. For the standard SRM and the R-SRM, we consider the following revised set of moment conditions

$$
\mathbf{R}_t = \begin{bmatrix}
\mathbb{E}_t\left[\widehat{\boldsymbol{\varepsilon}}^{\mathbb{P}}_{t+1}\mathcal{I}_{\{r_t\geq c\}}\right] \\
\mathbb{E}_t\left[\widehat{\boldsymbol{\varepsilon}}^{\mathbb{P}}_{t+1}\mathcal{I}_{\{r_t< c\}}\right] \\
vec\left(\mathbb{E}_t\left[\widehat{\boldsymbol{\varepsilon}}^{\mathbb{P}}_{t+1}\widehat{\mathbf{x}}^{(1)}_t\left(\boldsymbol{\theta}_1\right)'\right]\right) \\
vec\left(\mathbb{E}_t\left[\widehat{\boldsymbol{\varepsilon}}^{\mathbb{P}}_{t+1}\widehat{\mathbf{x}}^{(2)}_t\left(\boldsymbol{\theta}_1\right)'\right]\right) \\
vech\left(\mathbb{V}_t\left[\widehat{\boldsymbol{\varepsilon}}^{\mathbb{P}}_{t+1}\right]\right) \\
\mathbb{E}_t\left[\eta^{(5y)}_t\mathcal{I}_{\{r_t\geq c\}}\right] \\
\mathbb{E}_t\left[\eta^{(5y)}_t\mathcal{I}_{\{r_t< c\}}\right] \\
\mathbb{E}_t\left[\eta^{(5-10y)}_t\mathcal{I}_{\{r_t\geq c\}}\right] \\
\mathbb{E}_t\left[\eta^{(5-10y)}_t\mathcal{I}_{\{r_t< c\}}\right] \\
\mathbb{E}_t\left[\left(\eta^{(5y)}_t\right)^2\mathcal{I}_{\{r_t\geq c\}}\right] \\
\mathbb{E}_t\left[\left(\eta^{(5y)}_t\right)^2\mathcal{I}_{\{r_t< c\}}\right] \\
\mathbb{E}_t\left[\left(\eta^{(5-10y)}_t\right)^2\mathcal{I}_{\{r_t\geq c\}}\right] \\
\mathbb{E}_t\left[\left(\eta^{(5-10y)}_t\right)^2\mathcal{I}_{\{r_t< c\}}\right]
\end{bmatrix} - \begin{bmatrix}
\mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ vech\left(\mathbb{V}_t\left[\boldsymbol{\varepsilon}^{\mathbb{P}}_{t+1}\right]\right) \\ 0 \\ 0 \\ 0 \\ 0 \\ \sigma^2_\eta \\ \sigma^2_\eta \\ \sigma^2_\eta \\ \sigma^2_\eta
\end{bmatrix}. \tag{22}
$$

Here, we ignore the estimation uncertainty in the factors as accommodating this feature would make the estimation computationally infeasible.[17] With more moment conditions than parameters, we obtain the GMM estimates for $\boldsymbol{\theta}_2$ using numerical optimization of the squared residuals in equation (22), where we upscale the weight assigned to the survey-based moments by 6 to account for the fact that surveys are only observed once every 6 months. For the B-SRM, we use a similar estimation procedure with regimes defined based on the break point in the sample instead of the value of r_t. In the standard SRM and the B-SRM, all short rate expectations are computed in closed form. In the R-SRM, the short rate expectations are computed by Monte Carlo simulation.[18]

[17] Our experience with other SRMs suggests that the effects of allowing for estimation uncertainty in the factors are generally very small when representing the yield curve by 25 points each period.

[18] We find that a simulation using 500 simulated short rate paths generated using antithetic shocks is sufficiently accurate to allow estimation of the model.