Expectation and Duration at the Effective Lower Bound

Thomas B. King*
Federal Reserve Bank of Chicago

July 25, 2017

Abstract

I study equilibrium bond pricing with risk-averse arbitrageurs and an effective lower bound on nominal rates. The model exposes nonlinear interactions among short-rate expectations, bond supply, and term premia that are absent from affine models, and these features help it replicate the observed behavior of the yield curve near the ELB, including evidence about unconventional monetary policy. The impact of both short-rate expectations and bond supply are attenuated at the ELB. However, in simulations of the recent ELB episode in the U.S., the model implies that shocks to short-rate expectations influenced yields more than shocks to investors' duration-risk exposures.

*230 S. LaSalle Street, Chicago, Illinois, 60604. Phone: 312-322-5957. Email: thomas.king@chi.frb.org. For helpful comments and discussions on this paper and related work, I thank Stefania D'Amico, Robin Greenwood, Philippe Mueller, Andrea Vedolin, and seminar participants at the Federal Reserve Bank of Chicago, the 2015 Federal Reserve Day-Ahead Conference, and the 2015 Banque de France conference on Term Structure Modeling and the Zero Lower Bound. Zachry Wang provided excellent research assistance. The views expressed here do not reflect official positions of the Federal Reserve.

1 Introduction

Over the last decade, fixed-income markets have witnessed a combination of two extraordinary circumstances: massive changes in the quantity and structure of safe debt—including large-scale purchases of such debt by central banks—and the decline of short-term interest rates to their effective lower bound (ELB). This paper studies how these two phenonmena interact in an equilibrium model of arbitrage-free bond pricing.

To see empirically that the ELB may be important for the relationship between interest rates and debt supply, the top panel of Table 1 reports regressions of long-term yields on the weighted-average maturity of outstanding Treasury debt (WAM) and the one-year Treasury yield. The data are monthly, from 1971 through 2015. Using interactive dummies, I allow the coefficients on both variables to change after the ELB was reached in December 2008, but otherwise the regressions are identical to those of Greenwood and Vayanos (2014) (whose sample ended in 2007). Indeed, column 2 replicates the main result of that paper: long-term yields were significantly positively related to the duration risk held by investors in the pre-ELB period. A one-year increase in WAM pushed the 10-year yield up by 22 basis points in this sample, and, consistent with longer-term bonds having greater duration exposure, the coefficients were higher for longer maturities.

As column 3 shows, those coefficients all fall substantially and lose statistical significance during the ELB period. Despite the relatively few observations at the ELB, the t statistics reported in column 4 show that the declines in the coefficient values are significant.[1] The relationships between short- and long-term yields also change at the ELB, as shown in columns 5 through 7. Prior to 2008, the coefficients on the one-year yield were less than 1 and were monotonically decreasing in the maturity of the dependent variable. At the ELB, the coefficients rise above 2, with the 10- and 15-year yields now being more sensitive than the 5-year yield is. Again, t tests show that the differences across the two periods are statistically significant.[2] The remaining panels of the table show that the shifts in both sets of coefficients are robust to using the maturity-weighted debt-to-GDP ratio in place of WAM and the two-year yield in place of the one-year.

These results suggest important changes in the behavior of the yield curve and its

[1] The standard errors are calculated using the Newey and West (1987) procedure, with 36 lags, again following Greenwood and Vayanos (2014).

[2] Gilchrist, Lopez-Salido, and Zakrajsek (2015) document similar changes in the relationship between shorter- and longer-term yields at the ELB.

relationship to Treasury supply at the ELB. But what theoretical reasons do we have to expect such changes? I argue that at least three nonlinear mechanisms may be at work:

1. An increase in the quantity of longer-term bonds that investors hold raises the duration risk of their portfolios by an amount that depends directly on interest-rate volatility. Interest-rate volatility is lower when the short-rate distribution is truncated. Thus, if term premia are increasing in the amount of durtation risk held by investors, the effects of bond supply will be damped at the ELB.

2. At the ELB, near-term rate expectations are constrained and are unlikely to move much in response to shocks. Consequently, changes in expectations will have relatively larger effects on medium- and long-term yields.

3. Interest-rate volatility moves together with short-rate expectations at the ELB because an increase in the length of time that the ELB is expected to bind reduces near-term uncertainty about short rates.[3] If term premia depend on this uncertainty, changes in rate expectations will induce changes in term premia at the ELB. Again, these effects will be larger for longer maturities.

To formalize and quantify these possibilities, I incorporate the ELB into a structural model of bond pricing in the style of Vayanos and Vila (2009). In this type of model, the marginal investors are arbitrageurs with limited risk-bearing capacity. When they are given more long-term bonds to hold, the duration risk of their portfolios rises and they demand higher risk premia—a phenomenon sometimes known as the "duration channel" of bond supply. Many recent empirical studies on the effects of duration shocks have explicitly pointed to this framework for motivation and interpretation,[4] and other papers have extended and applied it in various ways.[5] Although those models have been useful for understanding the relationships between bond supply and bond yields, they have almost exclusively been developed under the assumption that the short-term interest rate follows a linear process. The three arguments just mentioned,

[3] Hattori, Schrimpf, and Sushko (2016) show that accomodative monetary-policy announcements during the ELB period caused declines in implied interest-rate volatility across the term structure.

[4] E.g., Gagnon, Raskin, Remache, and Sack (2011), Joyce, Lasaosa, Stevens, and Tong (2011), Swanson (2011), Krishnamurthy and Vissing-Jorgensen (2011), D'Amico and King (2013).

[5] E.g., Hamilton and Wu (2012), Greenwood and Vayanos (2014), King (2015), Altavilla, Carboni, and Motto (2015), Greenwood, Hanson, and Vayanos (2015), Haddad and Sraer (2015), Hayashi (2016), Malkhozov, Mueller, Vedolin, and Venter (2016).

together with the evidence in Table 1, suggest that the nonlinearity associated with the ELB may induce first-order changes in the yield curve's behavior. The situation in which the ELB binds is particularly important to consider because central banks have typically sought to exploit the duration channel through long-term bond purchases only after they have cut their traditional policy rates close to zero.

To model the ELB, I generalize the standard affine process for the short-term interest rate in Vayanos and Vila (2009) and its followers to a "shadow rate" process, following the use of that device in the recent empirical term-structure literature, such as Kim and Singleton (2012), Krippner (2012), and Wu and Xia (2016).[6] When the shadow rate is below the ELB, shocks to its value correspond to changes in investor beliefs about the length of time the ELB is expected to bind. Thus, they may capture both explicit forward guidance about the short-term interest rate and the "signaling channel" of asset purchases, through which expansions of the central bank's balance sheet might be viewed as a commitment to keep rates near zero for a longer time.[7]

When parameterized to match the unconditional moments of Treasury yields since 1971, the model delivers quantitatively accurate results in several respects. First, it replicates the basic features of the yield curve when the short rate is close to zero. Comparable models that ignore the ELB fail this test. Second, it reproduces the patterns reported in Table 1: the effects of shocks to Treasury supply become weaker at the ELB, and longer rates become "more sensitive" to shorter rates, with the latter coefficients switching from declining to increasing across maturities. Finally, when the model is subjected to shocks that approximate the Federal Reserve's unconventional monetary policy over the ELB period—bond-supply shocks that reduce the duration-weighted quantity of government debt by 18% and shadow-rate shocks that keep the short rate at the ELB for exactly seven years—it produces a cumulative yield-curve impact similar to what event studies suggest and a hump shape of the reaction in forward rates across maturities that matches the pattern observed in those studies. All of these empirical successes depend crucially on the three nonlinear mechanisms

[6] Bauer and Rudebusch (2014) argue that the shadow-rate specification does a good job of capturing yield-curve dynamics near the ELB, greatly outperforming traditional affine models. Notably, however, this literature has so far been dominated by atheroetical term structure models. This paper is among the first to incorporate a shadow-rate process into a structural model of the yield curve.

[7] Woodford (2012), Bauer and Rudebusch (2014), and Bhattarai, Eggertsson, and Gafarov (2015) argue for the importance of the signaling channel. As noted by Swanson (2017), because many announcements of asset purchases were accompanied by changes in the FOMC's communications about future short rates, it is impossible to distinguish empirically between the the effects of the signaling channel and those of forward guidance.

discussed above, which emerge endogenously in the model.

Having thus validated the model, I use it to address two quantitative questions. First, I ask through which channel unconventional monetary-policy shocks had their largest effects. The answer is that, given the magnitude of the observed shocks, changes in the expected path of the short rate were responsible for at least half of the cumulative downward shift in the yield curve during the ELB period. In addition, at longer maturities, about one-third of the total decline in yields is explained by the term-premium effects associated with reduced short-rate uncertainty at the ELB—a channel of unconventional policy that has been overlooked by previous literature. The bond-supply shocks account for less than 20% of the total decline in the ten-year yield, and at shorter maturities their contribution is even smaller. In other words, the model suggests that the duration channel of asset purchases was considerably less important than implicit or explicit forward guidance about the path of short-term rates.

Second, I ask the model whether the relative effectiveness of the two shocks changes in different environments. I find that bond-supply shocks are most powerful, relative to shadow-rate shocks, when the shadow rate is deeply negative and the amount of duration held by the market is high. In this situtation, the efficacy of both types of shocks is attenuated because of the damping effects associated with the ELB, but the attenuation is greater for the shadow-rate shocks. A negative shadow rate and a high quantity of market duration are precisely the conditions under which most Federal Reserve asset purchases were conducted. Thus, even though those purchases appear to have had only modest effects through the duration channel, their use could have been consistent with the Fed optimizing across its policy tools in the ELB environment.

This paper is related to several others in the recent literature. As noted above, a number of studies have used variants of the Vayanos-Vila (2009) framework to analyze the effects of fluctuations in bond supply in a linear environment. Hamilton and Wu (2012) briefly considered a version in which, once the short rate reached the ELB, investors believed that it would stay there with an exogenously given probability. However, because that probability was assumed to be constant, their model did not contain a mechanism for signaling or forward guidance. In addition, away from the ELB it priced bonds as if the ELB did not exist. Thus, their model lacked the key nonlinearities and interactions that drive most of my results.

Greenwood, Hanson, and Vayanos (2015) note the hump-shaped pattern in forward rates in response to unconventional-policy announcements and argue that expectations

of future changes in bond supply likely account for that pattern. While my model does not rule out their type of mechanism, it implies that the empirical hump in the forward curve can alternatively be explained by the non-monotonic effects of changes in short-rate expectations due to the ELB—a situation that cannot arise in their linear model. It also implies that the effects of bond-supply shocks are relatively modest at the ELB.

A final set of related papers are the empirical studies that have attempted to decompose the effects of unconventional policy into various channels. Krishnamurthy and Vissing-Jorgensen (2011, 2013) argue, based on event studies, that the evidence for the duration channel is weak, consistent with what my model implies. Swanson (2017) conducts event studies on unconventional policy to isolate a component reflecting short-rate expectations and a residual component that he essentially interprets as reflecting the duration channel. Although he concludes that the latter is important for long-term yields, his approach requires that factor loadings for interest rates in the ELB preiod were similar to those in the pre-ELB period. My model effectively allows for endogdous changes in loadings at the ELB and suggests that those changes could be quite substantial.[8]

2 Theoretical Framework

2.1 Investor behavior and equilibrium yields

I begin with the same portfolio-choice problem that forms the basis of the models in Vayanos and Vila (2009) and the several theoretical papers that have followed it. Investors have access to a continuum of zero-coupon bonds with maturities 0 to T. At each point in time t, they choose to hold a market-value quantity $x_t(\tau)$ of each maturity τ. Let $P_t^{(\tau)}$ represent the time-t price of a bond with remaining maturity τ. In addition, investors have access to a risk-free security that pays the instantaneous rate r_t. Investors' time-t wealth W_t is the sum of the market-value of the bond portfolio

[8] D'Amico, English, Lopez-Salido, and Nelson (2012) and Cahill, D'Amico, Li, and Sears (2013) present event-study evidence that asset purchases may also operate through a scarcity or "local supply" channel, whereby imperfect substitutability causes yields to fall by more for maturities where more purchases occurred. My model is silent about this type of phenomenon.

and the risk-free asset, and it thus evolves according to

$$dW_t = \int_0^T x_t(\tau) \frac{dP_t^{(\tau)}}{P_t^{(\tau)}} d\tau + r_t \left(W_t - \int_0^T x_t(\tau)\, d\tau \right) \tag{1}$$

Investors have mean-variance preferences, and thus, taking W_t as given, they choose quantities $x_t(\tau)$ to solve the problem

$$\max_{x_t(\tau)\forall\tau} \mathrm{E}_t[dW_t] - \frac{a}{2}\mathrm{var}_t[dW_t] \tag{2}$$

subject to (1), where a is absolute risk aversion and E_t and var_t represent expectation and variance conditional on the time-t state.

The first-order conditions for this problem can be written as

$$E_t\left[dp_t^{(\tau)}\right] = r_t + a\int_0^T x_t(\tau')\,\mathrm{cov}_t\left[dp_t^{(\tau)}, dp_t^{(\tau')}\right] d\tau' \tag{3}$$

for all τ, where $p_t^{(\tau)} = \log P_t^{(\tau)}$, cov_t denotes the covariance conditional on the time-t state, and τ' indexes the integration across maturities. Note that, under risk-neutrality ($a = 0$), all bonds have the same expected return, equal to the risk-free rate. Otherwise, the risk premium demanded for each bond is proportional to the covariance of that bond's price with the return on the whole portfolio of bonds.

The model is closed by assuming that the government exogenously supplies a time-varying quantity of bonds $s_t(\tau)$ at each maturity. A solution to the model is a set of state-contingent bond prices that clear the market. Specifically, market clearing requires

$$s_t(\tau) = x_t(\tau) \tag{4}$$

at each maturity τ and at each point in time t. Prices adjust to make (3) and (4) hold jointly in all states of the world. Solving the model is thus tantamount to solving for the conditional expectations and covariances in equation (3). Since the investors optimize without constraints on their portfolio weights, the equilibrium is arbitrage free.

The exogenous state variables in the model are r_t and $s_t(\tau)$. I assume that the short rate and the par value of debt outstanding ($s_t(\tau)/P_t^{(\tau)}$) are constant "within"

periods. That is, they jump discretely at regular intervals, normalized to unit length. This implies that $\frac{dP_h^{(\tau)}}{P_h^{(\tau)}} = p_{t+1}^{(\tau-1)} - p_t^{(\tau)}$ for all $h \in [t, t+1)$.[9] Since the yield on a τ-period bond has the usual relationship to its price, $y_t^{(\tau)} = -p_t^{(\tau)}/\tau$, it is then straightforward to show that yields are given by:

$$y_t^{(\tau)} = \frac{1}{\tau}\sum_{h=0}^{\tau} \mathrm{E}_t\left[r_{t+h}\right] + a\frac{1}{\tau}\mathrm{E}_t\left\{\sum_{h=0}^{\tau-1}\int_0^T \tau' s_{t+h}\left(\tau'\right)\mathrm{cov}_{t+h}\left[y_{t+h+1}^{(\tau-h-1)}, y_{t+h+1}^{(\tau')}\right]d\tau'\right\} \tag{5}$$

The first term on the left-hand side of (5) is the expectations component of yields; the second term is the term premium. The basic intuition for how asset purchases (or other fluctuations in bond supply) affect yields in this model is that they change the weights $s_t\left(\tau'\right)$ on the covariance terms in the term premium. A shock that shifts $s_t\left(\tau\right)$ toward lower-covariance assets—typically, those with shorter duration—will reduce yields through that term in period t. Note also that today's term premium depends not just on today's bond supply s_t, but also on the expected future values of supply s_{t+h}. Thus, if bond-supply shocks are persistent, they will also affect the expected value of the integral term in subsequent periods, leading to a further reduction in time-t yields.

Finally, it will also be instructive to examine forward rates. The one-period forward rate τ periods ahead is given by

$$f_t^{(1,\tau)} \equiv \tau y_t^{(\tau)} - (\tau - 1)y_t^{(\tau-1)} \tag{6}$$

2.2 The short rate

I assume that r_t follows the "shadow rate" process

$$r_t = \max\left[\widehat{r}_t, b\right] \tag{7}$$

where b is the lower bound on the short rate and

$$\widehat{r}_t = \mu_{\widehat{r}}(1 - \phi_{\widehat{r}}) + \phi_{\widehat{r}}\widehat{r}_{t-1} + e_t^{\widehat{r}} \qquad e_t^{\widehat{r}} \sim Niid\left(0, \sigma_{\widehat{r}}\right) \tag{8}$$

[9] The discretization makes little quantitative difference, and it becomes irrelevant as the length of the time interval goes to zero. It can be justified by the observation that monetary policy and debt issuance do not, in reality, adjust in infinitesimal increments in continuous time but rather move by sizeable amounts following periodic policy decisions.

for some parameters $\mu_{\widehat{r}}$, $\phi_{\widehat{r}}$, and $\sigma_{\widehat{r}}$. This is the discrete-time (single-factor) equivalent of the process used in the empirical shadow-rate literature mentioned in the introduction. As noted there, that literature generally shows that the shadow-rate specification performs well in describing the reduced-form dynamics of the yield curve at the ELB. Obviously, a special case that produces an affine specification for the short rate is $b = -\infty$. This will be a useful case for comparison, because it is the specification used in the previous theoretical literature on the duration channel.[10]

Since the shadow rate follows a Gaussian AR(1) process, the conditional distribution of future shadow rates h periods ahead is normal, with mean and variance given by the standard prediction equations

$$\mathrm{E}_t\left[\widehat{r}_{t+h}\right] = \mu_{\widehat{r}}(1 - \phi_{\widehat{r}}^h) - \phi_{\widehat{r}}^h \widehat{r}_t \tag{9}$$

$$\mathrm{var}_t\left[\widehat{r}_{t+h}\right] = \sigma_{\widehat{r}}^2 \sum_{j=1}^{h} \phi_{\widehat{r}}^{2(j-1)} \tag{10}$$

In an affine model, where $\widehat{r}_t = r_t$ in all states of the world, these equations also describe the conditional distribution of future short rates. To begin to get a sense of why the shadow-rate model delivers qualitatively different results than the affine model, note that, once the ELB is imposed, the short rate at any period in the future is distributed *truncated* normal if $\widehat{r}_{t+h} > b$ and is simply equal to b otherwise. Therefore, the mean and variance of r_{t+h}, conditional on information at time t, are given by

$$\mathrm{E}_t\left[r_{t+h}\right] = (1 - \Phi_t^{(h)}) \left(\mathrm{E}_t\left[\widehat{r}_{t+h}\right] + \varphi_t^{(h)} \sqrt{\mathrm{var}\left[\widehat{r}_{t+h}\right]}\right) + \Phi_t^{(h)} b \tag{11}$$

[10] Although the ELB is imposed *a priori* here, it is trivial to extend the model to endogenize it by allowing investors to hold an elastic supply of cash (paying zero nominal return) in addition to the risk-free asset. Alternative short-rate processes that impose the ELB also exist. For example, in Monfort, Pegoraro, Renne, and Roussellet (forthcoming), once the ELB is reached, the short rate stays there with some time-varying probability. At least qualitatively, such differences in specification are unimportant. The crucial features are that short-rate volatility is low at the ELB and that the ELB is "sticky," in the sense that the short rate tends to stay there for some time once it reaches it. Any model that generates these properties (which are amply evident in the data) will produce results along the lines of those presented below.

$$\begin{aligned}\text{var}_t\left[r_{t+h}\right] &= (1-\Phi_t^{(h)})\text{var}_t\left[\widehat{r}_{t+h}\right]\left[1-\frac{\varphi_t^{(h)}}{1-\Phi_t^{(h)}}\left(\frac{\text{E}_t\left[\widehat{r}_{t+h}\right]-b}{\sqrt{\text{var}_t\left[\widehat{r}_{t+h}\right]}}\right)-\left(\frac{\varphi_t^{(h)}}{1-\Phi_t^{(h)}}\right)^2\right] \\ &\quad +\Phi_t^{(h)}(\text{E}_t\left[r_{t+h}\right]-b)^2 \qquad (12)\end{aligned}$$

where $\varphi_t^{(h)}$ and $\Phi_t^{(h)}$ are the standard-normal PDF and CDF, respectively, evaluated at the point $\frac{b-\text{E}_t[\widehat{r}_{t+h}]}{\sqrt{\text{var}_t[\widehat{r}_{t+h}]}}$. It immediately follows that $\text{E}_t\left[r_{t+h}\right] \geq \text{E}_t\left[\widehat{r}_{t+h}\right]$ and $\text{var}_t\left[r_{t+h}\right] \leq \text{var}_t\left[\widehat{r}_{t+h}\right]$, with equality if and only if $b = -\infty$. In general, short-rate expectations are pushed up by the proximity of the ELB, and short-rate volatilities are damped.

Figure 1 depicts these conditional moments of the forward short rate, across different values of $\widehat{r}_t$. Keeping all other parameters the same, I compare the outcome when $b = 17$ basis points (solid lines) to the outcome when $b = -\infty$ (dashed lines).[11] The vertical lines indicate the location of the ELB in the first case. The forward moments are shown for horizons of 2, 5, 10, and 15 years.

As shown in Panel A, when $b = -\infty$, the forward expected short rate $\text{E}_t[r_{t+h}]$ is an affine function of $\widehat{r}_t$, with the slope of that function decreasing in h. When b is finite, $\text{E}_t[r_{t+h}]$ approaches this function as $\widehat{r}_t$ moves far above b. This is intuitive, since, as the ELB gets farther away, it should have less influence on asset prices and the model should behave approximately linearly. However, going in the other direction, expected future short rates asymptote to b as $\widehat{r}_t \to -\infty$. Because of this, when b is finite, the derivatives of $\text{E}_t[r_{t+h}]$ with respect to $\widehat{r}_t$ decrease and eventually go to zero. In other words shocks to $\widehat{r}_t$ have smaller effects on expected future short rates, particularly at relatively short horizons, when $\widehat{r}_t < b$.

Similarly, when $b = -\infty$, the conditional variance of the short rate is a constant value for all values of $\widehat{r}_t$ at any given horizons. This is depicted by the horizontal dashed lines in Panel B. Again, with finite b, the conditional short-rate variances approach these values as $\widehat{r}_t$ gets much larger than b. However, as $\widehat{r}_t$ falls below b, the conditional variance of the future short rate drops notably. The reason for this is intuitive—when the shadow rate is far below the ELB, the *actual* short rate will almost certainly be

[11] The other parameter values are those shown in the top row of Table 2 in Section 3; the details of the calibration are discussed there. I note that the comparison in Figure 1 is intended to isolate the consequences of imposing the ELB holding everything else fixed. When fitting such models to the data, of course, *all* of the parameters will generally differ between affine and shadow-rate specifications, so the quantitative discrepancies between the solid and the dashed lines could be different than those shown.

equal to the ELB for a long time and therefore will display little variation.

Returning to equation (5) with these observations in mind, we can see heuristically how the ELB will matter for the propagation of shocks. First, shocks to $\widehat{r}_t$ will have generally weaker effects on the expectations component of yields when $\widehat{r}_t < b$. Moreover, these shocks may have larger effects on the expectations component of medium or long-term yields than on shorter-term yields. This contrasts to an environment far above the ELB, where the effects of shocks to $\widehat{r}_t$ are always largest at the short end of the curve. Second, current and future short-rate volatilities are lower at the ELB, which will mean that the volatilities of *all* yields are lower, *ceteris paribus*. Thus the covariance terms that represent the multipliers on s_t in equation (5) will generally be smaller. This means that a given shift in the supply distribution will have a smaller effect on term premia at the ELB than it does away from the ELB (or in an affine model). Finally, the covariance terms that determine term premia are increasing in the level of $\widehat{r}_t$. As shown in Figure 1, at and near the ELB there is a positive relationship between the expected future short rate and its variance; all else equal, this will translate into a positive relationship between short-rate expectations and term premia. Because the conditional variance of r_t is constant when it follows an affine process, this channel does not exist in affine models.

2.3 Bond Supply

The arguments just sketched for the qualitative effects of the ELB made no reference to the way in which asset supply $s_t(\tau)$ was determined. Indeed, they hold for a variety of possible processes for bond supply. Nonetheless, to obtain a quantitative assessment, we must specify a particular process.

Since bond supply is continuous across maturities, the object $s_t(\tau)$ is an infinite-dimensional vector. Clearly, it is desirable to reduce this dimension. Previous literature specifies $s_t(\tau)$ as an affine function of a finite state vector β_t that follows a linear-Gaussian process:

$$s_t(\tau) = \zeta(\tau) + \theta(\tau)\beta_t \tag{13}$$

$$\beta_t = \phi_\beta \beta_{t-1} + e_t^\beta \qquad e_t^\beta \sim Niid(0, \sigma_\beta) \tag{14}$$

where $\zeta(\tau)$ and $\theta(\tau)$ are maturity-specific intercepts and coefficients. I follow this approach, taking the dimension of β_t to be 1. I further assume that the intercept is constant across maturities: $\zeta(\tau) = \zeta$. This involves only a small loss of generality,

since $\zeta(\tau)$ is integrated out in equation (5) and is thus only a level shifter. Similarly, from equation (5), the individual factor loadings $\theta(\tau)$ do not matter for yields; only the weighted sum $\int_0^T \tau' \theta_t(\tau') \text{cov}_t \left[y_{t+1}^{(\tau)}, y_{t+1}^{(\tau')} \right] d\tau'$ does. This suggests that the exact specification of the function $\theta(\tau)$ is not of first-order importance, so long as it can generate realistic behavior for overall portfolio duration.[12] For simplicity, I follow Greenwood, Hanson, and Vayanos (2015) by assuming that this function is linear in τ:

$$\theta(\tau) = \left(1 - \frac{2\tau}{T}\right) \tag{15}$$

This specification implies that the bond distribution behaves with a see-saw motion across maturities. Positive supply shocks reduce the amount of long-term bonds and increase the amount of short-term bonds in equal measure, with the fulcrum at $\tau = T/2$.

Two helpful summary measures of bond supply that are frequently used in the literature have direct counterparts in the model. The first measure is the weighted-average maturity (WAM) of the oustanding debt, which is given by

$$WAM_t = v \frac{\int_0^T \tau s_t(\tau) \, d\tau}{\int_0^T s_t(\tau)_t \, d\tau} \tag{16}$$

where v is the length of one period, expressed in years.[13] The second measure is the amount of "ten-year-equivalent" bonds outstanding. This variable is defined as the dollar value of ten-year bonds that would produce the same duration-weighted value that the actual portfolio of outstanding bonds has. (Thus, for example, a portfolio of 5-year bonds with a value of \$100 is worth \$50 in terms of ten-year equivalents.) Mathematically, the amount of ten-year equivalents is defined as

$$10YE_t = \frac{v}{10} \int_0^T \tau s_t(\tau) \, d\tau \tag{17}$$

The integrals in both of these equations can be evaluated analytically, given the assumed process for $s_t(\tau)$, providing convenient ways of translating real-world changes

[12] Malkhozov, Mueller, Vedolin, and Venter (2016) make a similar point.

[13] Note that, since all bonds are assumed to be zero-coupon, duration is simply equal to maturity.

in the outstanding bond distribution into the bond-supply shocks of the model.

3 Calibration and Solution

There are nine parameters in the model, which I set to match empirical moments of Treasury supply and the yield curve. Specifically, I use the Gurkaynak, Sack, and Wright (2007) zero-coupon yields available on the Federal Reserve Board's website and the Treasury security data available in CRSP.[14] I start the sample in August 1971 because at that time 10-year yields become available. (I also recalibrated using a sample excluding the volatile period of the 1970s without much effect on the results.) The sample ends in December 2015. The specific moments that I match are discussed below. The calibration is summarized in the top line of Table 2.[15]

A period is normalized to one calendar quarter (i.e., $v = 1/4$), and I take T, the maximum-maturity bond available to investors, to be 60 quarters. This matches the longest maturity bond that was continuously available over the period under consideration. It also happens to be close to the typical duration of a 30-year coupon bond, which is the longest bond issued by the Treasury at any point during this period. Alternative values for T make little difference, however, because they affect the volatility of term premia in a way that is largely neutralized by the calibration of a.

I calibrate the autoregressive coefficient on the supply factor ϕ_β to match the persistence of the weighted-average maturity of outstanding Treasury debt (the same series used in the regressions in Table 1). In the data, this variable is calculated as the value-weighted timing of all cash flows on all Treasury instruments held by the public. In the model, the WAM of the debt held by investors, expressed in years, is given by equation (16):

$$WAM_t = v\frac{\int_0^T \tau\left[\zeta + \left(1 - \frac{2\tau}{T}\right)\beta_t\right] d\tau}{\int_0^T \zeta + \left(1 - \frac{2\tau}{T}\right)\beta_t d\tau} = vT(\frac{1}{2} - \frac{1}{6\zeta}\beta_t) \tag{18}$$

[14] Center for Research in Security Prices, Booth School of Business, The University of Chicago. Used with permission. All rights reserved. crsp.uchicago.edu.

[15] Since the calibration matches selected moments of the yield curve, it can be viewed as a form of limited-information estimation. Because the model is solved numerically and involves state variables that are both unobserved and nonlinear, full-information maximum-likelihood estimation would be computationally quite challenging and is beyond the scope of this paper.

Since WAM_t is linear in β_t, it has the same persistence. Thus, I match the persistence of WAM in the data, using the four-quarter autocorrelation (0.92) to abstract from seasonal patterns in Treasury issuance. This gives $\phi_\beta = 0.98$.[16] The parameter σ_β determines the scale of the bond-supply factor. Since β_t is unitless, this parameter has no economic content and indeed is not separately identified. Without loss of generality, therefore, I set it such that the unconditional variance of β_t is normalized to 1.

I calibrate the remaining parameters jointly to match the long-run empirical features of the yield curve. The specific moments I match are the unconditional mean and standard deviation of the three-month Treasury yield (5.2% and 3.6% respectively), the unconditional mean and standard deviation of the ten-year yield (6.7% and 2.8%), and the correlation between the three-month and ten-year Treasury yield (0.91).[17] Finally, I match the average value of the three-month Treasury yield during the ELB period. Specifically, between December 2008 and December 2015, the three-month yield averaged 0.22%, with a maximum value of 0.68%. The calibration that achieves a mean short rate of 0.0022 conditional on $\widehat{r}_t < 0.0068$, given the other values of the short-rate parameters, is $b = 0.0017$.

For comparison, I consider two alternative models in which $b = -\infty$, i.e., models with an affine process for the short rate. In the first such model, shown in row 2 of Table 2, I set all parameters other than b equal to the same values as in the shadow-rate model in order to isolate the effects of imposing the ELB. In the second affine model, shown in row 3, I recalibrate the parameters to match the same set of unconditional yield-curve moments that the shadow-rate model matches. The parameters turn out to be fairly similar to those in the baseline model, with the primary differences being that $\mu_{\widehat{r}}$ is a bit higher and $\sigma_{\widehat{r}}$ is a bit lower due to the truncation effects shown in equations (11) and (12).

In general, the model does not have an analytical solution. I solve globally by discretizing the state space and iteratively (a) calculating state-contingent yields in equation (5) given conditional expectations, and (b) calculating conditional expectations given state-contingent prices using the transition densities implied by equations

[16] Of course, equating investors' bond holdings in the model with Treasury debt in the data might be taking the model too literally given that investors may also have duration exposure through other instruments. However, other choices for ϕ_β in the 0.8 to 1 range, holding the rest of the parameters constant, produce similar outcomes to those reported below.

[17] Note that the three-month yield used in these calculations is the fitted value of the Gurkaynak et al. curves, which are based on Treasury coupon-security data. It is not a Treasury bill rate. It thus avoids any premium associated with very liquid, "money-like" assets.

(8) and (14). Cubic interpolation between the discretized nodes is used for situations, such as model simulation, in which state values are required to be continuous. The details of the solution method are described in Appendix A.

4 Results

4.1 Model fit

Table 3 summarizes the properties of bond yields produced by the calibrations in the shadow-rate and affine models and compares these results to the data. The model-implied moments are calculated by drawing 1,000,000 times from the distributions of e_t^r and $e_t^\beta$and simulating the resulting paths of the state variables $\widehat{r}_t$ and β_t. To illustrate the importance of the ELB, I report the results conditional on the short rate being both below and above the value 0.68%. Again, the reason for choosing this threshold is that it was the maximum attained by the three-month Treasury yield in the data during the time that the Federal Reserve kept its policy rate in the $0 - 25$ basis point range.

The shadow-rate model matches the data quite well when the short rate is at its lower bound, coming within a few basis points of the means and standard deviations of all but the longest yields. In contrast, the affine model with the same calibration predicts a short rate that averages -1.3% (negative Treasury rates never actually appear in the data) and an average yield-curve slope that is dramatically steeper than what was observed. It also predicts slope volatilities that are somewhat farther from the data than those given by the shadow-rate model. The affine model that is recalibrated to match the unconditional yield curve moments does slightly better at the ELB, but it still significantly underperforms the shadow-rate model. It also underestimates the frequency with which the ELB binds by about 40%.

The shadow-rate model achieves its successes near the ELB without sacrificing performance relative to the affine models in other regions of the state space. As the bottom panel shows, all three models differ from each other by only a few basis points for all of the reported statistics when the short rate is greater than 0.68%. To be sure, the shadow-rate model's performance is not perfect—like most structural yield-curve models, it produces too little curvature on average and has trouble simultaneously matching volatilities at all maturities. But even in these respects the differences with the data are not dramatic, and it still performs at least as well as the affine models.

4.2 Yield determination in the shadow-rate model

I now consider the effects of the ELB on the yield curve and on the propagation of the two shocks in different regions of the state space. To facilitate this discussion, consider the first-order Taylor series expansion of the τ-maturity yield. For arbitrary state values, $\widehat{r}$ and β, we have

$$y_t^{(\tau)} \approx C_t^{(\tau)} + A_{\widehat{r},t}^{(\tau)}\widehat{r} + A_{\beta,t}^{(\tau)}\beta \tag{19}$$

where $A_{\widehat{r},t}^{(\tau)} \equiv \partial y_t^{(\tau)}/\partial \widehat{r}_t$ and $A_{\beta,t}^{(\tau)} \equiv \partial y_t^{(\tau)}/\partial \beta_t$, with both derivatives evaluated at the time-t values of the states, and $C_t^{(\tau)}$ is the corresponding intercept term. In general, these derivatives vary over values of the state variables; hence, their time subscripts. They thus have an interpretation as state-dependent factor loadings. In the case of the affine process for the short rate, yields themselves are affine in the states (as shown, for example, in Vayanos and Vila, 2009). Thus, in that case, equation (19) holds exactly, and the derivatives are constant.

The solid lines in Figure 2 depict the factor loadings in the shadow-rate model across a range of values for $\widehat{r}$, holding β fixed at its mean value of zero. (The derivatives are a byproduct of the solution algorithm.) The dashed lines depict the corresponding loadings in the affine model under the baseline calibration (line 2 of Table 2). Since yields are affine functions of the states in the affine model, those lines are always flat. The factor loadings in the shadow-rate model asymptote to those of the affine model as $\widehat{r}$ rises farther above the ELB.

The factor loadings in the shadow-rate model near and below the ELB follow from the nonlinear conditional moments of the short rate that were depicted in Figure 1. As was shown there, for lower values of $\widehat{r}_t$, both the conditional mean and the conditional variance of r_{t+h} in future periods falls. The lower mean leads to lower yields through the expectations component, and the lower variance leads to lower yields through the term premium. Thus, as shown in panel A of Figure 2, $A_{\widehat{r},t}^{(\tau)}$ is monotonically increasing in $\widehat{r}$. Furthermore, for $\widehat{r}$ low enough, $A_{\widehat{r},t}^{(\tau)} < A_{\widehat{r},t}^{(\tau')}$ when $\tau < \tau'$. Consequently, longer-term yields respond to shadow-rate shocks by more than shorter-term yields do. This pattern is the opposite of what we observe when $\widehat{r} > b$, and it is the opposite of what the affine model predicts.

Recall that increases in β_t reduce the duration exposure of investors and therefore have negative effects on term premia. As shown in panel B, in the shadow-rate model,

$A^{(\tau)}_{\beta,t}$ is monotonically decreasing (i.e., becoming more negative) in $\widehat{r}$. This is because $\text{var}_t[r_t]$ falls as $\widehat{r}$ moves below the ELB, causing the covariance terms in (5) to become smaller. Consequently, at the ELB—and particularly when $\widehat{r}$ is deeply negative—bond-supply shocks have smaller effects on yields than they do in the affine model.

Figure 5 shows the factor loadings plotted across a range of values for β, holding $\widehat{r}$ fixed at either the unconditional mean of the short rate (Panel A) or at a value of -2.7% (Panel B). I choose the latter value for illustration of the ELB environment because it is the average of the shadow rate estimated by Krippner (2012) over the period December 2008 to December 2015.[18] From this perspective, the differences between the shadow-rate and the affine model are evident even when the short rate is at its unconditional mean and are of first-order importance when the shadow rate is negative. In that region, two particularly noteworthy results stand out. First, $A^{(\tau)}_{\widehat{r},t}$ is not only strictly lower that it is when $\widehat{r}$ is positive, it is also decreasing in β. The reason is that, when β is positive, investors have relatively little exposure to long-term bonds. Consequently, when the shadow rate rises, the resulting increase in short-rate risk has a relatively small effect on term premia. When β is negative, in contrast, investors' bond exposures are greater, and increases in the shadow rate have a larger impact on term premia through their effects on short-rate volatility. Second, $A^{(\tau)}_{\beta,t}$ is increasing (becoming less negative) in β. Intuitively, higher levels of β reduce exposure to long-term bonds, making long-term yields less sensitive to the changes in short-rate risk induced by the shadow-rate. Consequently, positive shocks to β reduce the volatility of yields, making *further* shocks to β less potent. This result implies, for example, that the marginal effects of asset purchases decline as the central bank does more of them. It will also be important for analyzing the relative effectiveness of alternative policies in different environments in Section 6.

The state-dependent factor loadings explain the empirical patterns that were illustrated in Table 1 in the introduction. Recall that the regressions reported there showed that Treasury supply had smaller effects during the ELB period than in the pre-ELB sample. This is exactly the result predicted by panel B of Figure 2. Table 1 also showed that regression coefficients of long-term on short-term yields became larger at the ELB, in particular rising from less than 1 to greater than 1. That the model

[18] To interpret the meaning of $\widehat{r} = -.027$, simulations starting from this value produce a modal time of 6 quarters until the shadow rate moves above the ELB. This is roughly consistent with survey evidence on market participants' expectations and other evidence collected during much of the ELB period (see Femia, Friedman, and Sack (2013)).

reproduces this result can be seen by examining the factor loadings in the panel 2.A, and in particular how they cross in the sub-ELB region.

Table 4 makes these patterns clearer by computing the linearized relationships between long-term yields, shorter-term yields, and WAM_t in the model. Given equations (18) and (19), the response of the τ-period yield to a change in WAM_t is $-6\zeta A^{(\tau)}_{\beta,t}/vT$. Panel A reports these responses at various values of the shadow rate, for $\tau = 5$, 10, and 15 years, in the first group of columns. To make the comparison to Table 1 clearer, the second and third groups of columns report the sensitivity to WAM_t holding fixed the one-year yield or the two-year yield.[19] The sensitivities show a clear decline at the ELB, similar to that suggested by the Table 1 regression results.[20] Similarly, panel B shows state-contingent linearized coefficients of the τ-period yield on the one- and two-year yields, holding the supply factor fixed ($A^{(\tau)}_{\widehat{r},t}/A^{(4)}_{\widehat{r},t}$ and $A^{(\tau)}_{\widehat{r},t}/A^{(8)}_{\widehat{r},t}$). As in Table 1, the coefficients rise from less than 1 to greater than 1 at the ELB and switch from being most-sensitive to least-sensitive at the five-year maturity. These results further support the ability of the model to explain the behavior of the yield curve and its relationship to bond supply in the ELB environment.

4.3 The effects of shocks

To see the results in another way, Figure 4 shows the dynamic effects of shocks in the shadow-rate model. I consider independent one-standard-deviation shocks to the shadow rate and bond supply in directions that lower yields ($e^r_t = -.0078$ or $e^\beta_t = 0.20$). I consider two cases: one in which the short-rate begins at its unconditional mean of 5.2%, and one in which it begins at -2.7%, its ELB-period average according to the Krippner estimates. (In both cases, I let β_t start at its mean value of zero.) For each set of starting values, I simulate the model forward ten years, both with the shocks in the first period and without them. I compute impulse-response functions as the difference between those two simulations. In the figure, each IRF depicts the response of the entire yield curve over the ten-year period, with maturities on the lower-left axis and calendar time on the upper-left axis. Panel A shows the response of spot yields, while Panel B presents the same information in terms of forward rates, where

[19] These coefficients are given by $-6\zeta(A^{(\tau)}_{\beta,t} - A^{(\tau')}_{\beta,t}/A^{(\tau')}_{\widehat{r},t})/vT$ for $\tau' = 4$ or 8 quarters.

[20] Since there is no concept of GDP in the model, it is not possible to mimic the specifications in Table 1 that use the maturity-weighted debt-to-GDP ratio. However, if the model results are computed replacing WAM_t with the model-implied measure of ten-year-equivalent bonds the same patterns hold.

the patterns are somewhat easier to see.

When the short-rate starts from 5.2%, the shock to $\widehat{r}_t$ has a monotonic impact on the yield curve. The response to the bond-supply shock displays nearly the opposite pattern, with no effect on short-term yields and relatively large effects at the long end. The shadow-rate shock initially lowers the ten-year yield by 54 basis points, and the bond-supply shock lowers it by 23 basis points. The effects of both shocks decay monotonically over time, with the yield curve coming most of the way back to its starting position by the end of the ten years shown.

When the shadow rate starts below b, the outcomes are notably different. For one thing, the effects of both shocks are smaller; the shadow-rate shock now reduces the ten-year yield by just 28 basis points, while the bond-supply shock reduces it by 15 basis points. In addition, the shock to $\widehat{r}_t$ now produces a hump-shaped reaction across maturities in forward rates, as shown in panel B. This pattern will be critical for explaining the observed forward-curve response in event studies, considered later. Again, it occurs because shorter-term forwards cannot move much lower to begin with; the deeply negative shadow rate has already depressed them to near b. Meanwhile, long-term forward rates are still not much affected by shadow-rate shocks because they depend mostly on expectations of the short rate in the far future. Consequently, shadow-rate shocks have their largest effects on medium-term yields. This prediction of the model is consistent with the evidence presented by Swanson and Williams (2014), who show that responses of shorter-term yields to macroeconomic shocks were muted during the ELB period, and Carvalho, Hsu, and Nechio (2016), who show that Federal Reseve communications had their largest effects in the 2- to 10-year maturity range during the ELB period. The exact maturity of the peak of the hump (about 7 years, in the case shown in figure 4) depends on the size of the shock and how far below b the shadow rate initially is.

While the bond-supply shock operates entirely through the term premium, the shadow-rate shock has effects through both the expectations component of yields and the term premium. The the latter effect arises because of the way that the shadow rate affects short-rate volatility at the ELB, and it is more pronounced when the shadow rate is negative. This is illustrated in Figure 5, which plots the initial responses of spot and forward rates to the same shadow-rate shock shown in the previous figure, with those responses now decomposed into the expectations component, shown in pink, and the term premium, shown in blue. (The expectations component is given by integrating

equation (11) across maturities, and the term premium is just the difference between the total yield response and the expectations component.) When the shadow rate starts at 5.2%, the term premium barely moves in response to the shock, reflecting the remoteness of the ELB. In contrast, when the shadow rate starts at -2.7%, the reduction in the term premium in response to the shock accounts for about a third of the overall yield decline at longer maturities (e.g., 10 of the 28 basis points on the ten-year yield). Its largest effect on forward rates is at the 8-year maturity, contributing to the hump shape.

5 Assessing unconventional policy

5.1 Simulating unconventional policy

I now use the model to study the effects of unconventional monetary policy. The Federal Reserve implemented two main types of such policy: asset purchases (also known as "quantitative easing") and forward guidance about the future course of the short-term interest rate. Jointly, these policies can be mapped into the shadow-rate and bond-supply shocks of the model. However, one should avoid associating QE only with shocks to bond supply and associating shadow-rate shocks only with forward guidance. As a number of authors have noted, QE may have worked in part through a "signaling channel," serving as a commitment by the Fed to keep the short rate at the ELB for a longer time. (E.g., Woodford (2012); Bauer and Rudebusch (2014); Bhattarai, Eggertsson, and Gafarov (2015).) If so, then such policies involve shocks to both bond supply and the shadow rate. For this reason, I do not attempt to distinguish the effects of forward guidance and QE *per se* but rather model the joint effects of changes in the anticipated short rate and bond supply.

The strategy is to feed the model a set of shocks that approximate those associated with unconventional policy during the ELB period and calculate the yield-curve effects of those shocks. To conduct this exercise, one must translate the actions taken by the Federal Reserve into shocks that can be input into the model. Cumulatively, we know fairly precisely how large these shocks were. r_t remained at the ELB for seven years, and so the shadow-rate shocks in the simulation must keep r_t at b for exactly 28 periods. Over the same time, Greenwood, Hanson, Rudolph, and Summers (2015) report that the Fed removed approximately $2.7 trillion of ten-year-equivalent bonds from the market, including Treasuries, agency debt, and MBS. This was approximately

18% of the total 10-year equivalents outstanding in these markets as of December 2015, which from equation (17) is sufficient to pin down the cumulative size of the shocks to bond supply.[21]

While it is tempting to interpret the bond-supply shocks associated with QE events simply as realizations of e_t^β, Federal Reserve asset purchases likely differed in important ways from the other types of bond-supply fluctuations that dominate the long span of data. In particular, in the baseline model above the parameter ϕ_β was calibrated to a value of 0.98, implying a half-life of 8.5 years, to match the persistence of Treasury debt since 1971. But Fed asset purchases were almost certainly interpreted as less persistent than that. Carpenter, Ihrig, Klee, Quinn, and Boote (2013) inferred from surveys of market participants, conducted while the QE programs were taking place, that the size of the Fed's balance sheet was expected to normalize by August 2020. By that reckoning, the expansion of the Fed's balance sheet, which occurred between December 2008 and December 2014, had a perceived half-life of less than 4.5 years on average, substantially less than that of e_t^β shocks under the baseline calibration.[22]

To account for these differences, I extend equation (13) to allow for an additional supply factor Q_t representing bond-supply shocks due to changes in the Federal Reserve's balance sheet:

$$s_t(\tau) = \zeta + \theta(\tau)(\beta_t + Q_t) \tag{20}$$

where

$$Q_t = \phi_Q Q_{t-1} + e_t^Q \tag{21}$$

for "Fed balance-sheet" shocks e_t^Q. I set $\phi_Q = 0.96$, giving these shocks a half life of 4.25 years. When calculating bond yields in this model, I set the variance of e_t^Q to zero, so that the perceived risk associated with total bond supply is the same as in

[21] In December 2015, the CRSP Treasury data show ten-year-equivalent Treasury bonds of $11.6 trillion, while SIFMA data show $7.2 trillion of agency-backed MBS and CMOs and $1.3 trillion of long-term agency debt outstanding (http://www.sifma.org/research/statistics.aspx). Hanson (2014) shows that the average duration of a 30-year MBS is about 3.5 years, and I assume that the duration of long-term agency debt is 5 years. Under these approximations, ten-year equivalents outstanding totalled $14.8 trillion.

[22] Other evidence on the persistence of QE is mixed, but it does not suggest a coefficient as high as 0.98. Wright (2012) estimates a half-life of less than a year for the effects of unconventional monetary-policy shocks on yields in a VAR. Altavilla and Giannone (forthcoming) show that markets expected most of the effects of unconventional policy to persist for at least a year, but the survey data they use do not extend beyond that horizon. Similarly, Swanson (2017) finds that the persistence of most QE shocks was large, but he limits the estimation horizon to 180 business days.

the baseline model.[23] The process for β_t remains the same, and all other parameters continue to take the values shown in the top row of Table 2. Since Q_t does not add additional risk to the model, the conditional moments of yields and the factor loadings for $\widehat{r}_t$ and β_t also remain the same as above. I note that, although the addition of the balance-sheet factor adds realism to the model, the results presented below are largely unchanged if asset purchases are simply treated as ordinary shocks to β_t.

The nonlinearities induced by the ELB mean that both starting values and the trajectory of the shocks matter. I use starting values based on the configuration of the yield curve on the eve of unconventional policy. Letting $t = 0$ denote the period immediately before unconventional policies were enacted, I set $\widehat{r}_0 = 0.0017$, just at the ELB. I set $\beta_0 = -0.34$, which produces a ten-year yield-curve slope of 3.0%, the observed slope as of the FOMC meeting prior reaching the ELB. I initialize Q_0 to zero, since QE did not exist prior to the ELB.[24]

Since we cannot directly observe the trajectories of the shadow rate and the Fed balance-sheet factor in the data, I simulate a range of possible trajectories, with each trajectory being consistent with the observed outcomes of (1) a short rate that stays at zero for exactly 28 periods and (2) a cumulative net reduction in 10-year-equivalent bonds of 18%.[25] The details of these simulations are discussed in Appendix B. Each simulation i consists of a set of 28 shocks to both the shadow rate and the Fed's balance sheet $\{(e^{\widehat{r}}_{i,1}, e^{Q}_{i,1})$, ..., $(e^{\widehat{r}}_{i,28}, e^{Q}_{i,28})\}$, which accumulate into the state trajectories $\{(\widehat{r}_{i,0}, Q_{i,0})$, ..., $(\widehat{r}_{i,28}, Q_{i,28})\}$ via equations (8) and (21). The initial values $(\widehat{r}_{i,0}, Q_{i,0})$, which were just discussed, are the same in all simulations. The shocks to β_t are set to zero, so that variable simply decays back toward its mean over the period, following a

[23] This assumption is justified because QE purchases account for very little of the unconditional variation in the duration risk of investors' portfolios. Allowing for a positive variance of e^{Q}_{t} does not substantively change the results below, as long as it is less than the unconditional variance of e^{β}_{t}.

[24] The ELB was officially reached on December 16, 2008, when the FOMC cut the target federal funds rate from 1% to a range of 0 to 25 basis points. However, from the Treasury market's perspective the effective date may have been slightly earlier. The three-month yield declined 102 basis points over the intermeeting period leading up to December 16, in anticipation of the cut. In addition, the first announcement of asset purchases came on November 25. Using a starting value for bond supply based on the situation as of October 29, 2008, ensures that it does not include these pre-ELB influences of unconventional policy.

[25] One might use data on the Fed's holdings as an observed measure of Q_t, and such a path is in fact spanned by the set of trajectories I simulate. I allow for more ambiguity, however, because the term premium depends on expectations of future asset purchases, not just the amount the Fed currently holds. In the model here, $\mathrm{E}_t[Q_{t+h}]$ depends only on Q_t, making the distinction between current balances and expected future balances fuzzy. Greenwood, Hanson, and Vayanos (2015) explore this distinction in detail.

path that is identical across all simulations.

Panel A of Figure 6 shows the resulting distribution of the simulated trajectories for $\widehat{r}_t$. This distribution spans empirical estimates of the shadow-rate path during the ELB period, including those of Krippner (2012) and Wu and Xia (2016). Panel B shows the distribution of the Q_t, trajectories, converted to cumulative percentage changes in ten-year-equivalent bonds outstanding for ease of interpretation, using equation (17). It is more difficult to know what the "right" path of this variable ought to be (see footnote 25), but the distribution covers a fairly wide range of possibilities.

With the simulated distributions of the state-variable trajectories in hand, I use the model to extract the yield-curve responses. To report the results, for each period in each simulation I calculate how the yield curve changes, relative to how it would have changed if there had been no shock in that period. I then sum these differences across periods within each simulation. This proceedure is analagous to empirical event studies that attempt to isolate and accumulate the immediate impact of policy shocks without accounting for their dynamics. (Here, there is an "event" in every period.) Specifically, letting $y^{(\tau)}\left(\widehat{r}_t, \beta_t, Q_t\right)$ denote the τ-maturity yield as a function of the state variables, I calculate

$$D_i^{(\tau)} = \sum_{t=1}^{28} \left[y^{(\tau)}\left(\widehat{r}_{i,t}, \beta_t, Q_{i,t}\right) - y^{(\tau)}\left(\widehat{r}_{i,t} - e_{i,t}^r, \beta_t, Q_{i,t} - e_{i,t}^Q\right) \right] \tag{22}$$

The distribution of $D_i^{(\tau)}$ across simulations is shown in Figure 7.A, with the corresponding calculation for forward rates shown in 7.B.

The median decline in the ten-year yield produced by the simulated unconventional policy shocks is 202 basis points. This estimate does not differ much across simulations, with the middle 90% of the distribution spanning only the range of -205 to -194 basis points. It is worth noting that these are similar magnitudes to the effects that have been estimated in event studies of unconventional policy. For example, looking at 23 important policy announcements during the ELB period, and controlling for macroeconomic news, Altavilla and Giannone (forthcoming) find a net effect on the ten-year yield of -176 basis points. Comparisons between these kinds of results and the model cannot be made precise because empirical event studies necessarily capture only a subset of the relevant shocks, but they suggest that the model's results are quantitatively realistic.

Perhaps more importantly, the model reproduces another key stylized fact from

the event-study literature. Rogers, Scotti, and Wright (2014) and Greenwood, Hanson, and Vayanos (2015) show that unconventional policy annnouncements typically resulted in a hump-shaped reaction across the forward curve, with forward rates in the 5- to 7-year range moving the most. The model generates exactly this pattern (see Figure 7.B). As was evident in Figure 5, it does so because of the non-monotonic effects of shadow-rate shocks induced by the ELB, which operate through both the expectations and term-premium components of yields.

The consistency of the model with the empirical evidence on unconventional policy, together with the other results presented earlier, suggests that it does a good job of matching the behavior of Treasury yields and their relationship to bond supply in dimensions that are observable. Thus, one can have some confidence in what the model has to say about the *unobservable* aspects of unconventional-policy shocks. I examine two of these aspects. First, I decompose the cumulative contemporaneous yield-curve reaction shown in Figure 6 into various channels of unconventional policy. While this calculation captures the sources of the changes in yields in the periods when shocks occurred, it does not account for the dynamic effects of those shocks. Therefore, the second approach I take is to calculate a decomposition of the total model-implied variance in yields during the ELB period. In both exercises, the breakdown is calculated by computing what the change in yields would have been if only the shadow-rate or the Fed balance-sheet shocks had occurred (again, relative to a baseline case in which there are no shocks at all). In the case of the shadow-rate shocks, the response can be further decomposed into the expectations and term-premium components. Finally, because of the nonlinearities, the responses to the individual shocks do not sum exactly to the total response when both types of shocks occur simultaneously. An "interaction" term captures the residual.

Table 5 shows the results of these decompositions according to the medians across simulations, with the 5% and 95% quantiles reported in parentheses. Panel A presents the decomposition of the cumulative contemporaneous effects and panel B presents the total-variance decomposition. Looked at either way, the shadow-rate shocks are responsible for considerably more of the change in yields than the Fed balance-sheet shocks are. For example, as shown in column 4 of panel A, the model implies that the duration effects associated with QE lowered the ten-year yield by a total of just 34 basis points in the periods when they occurred. The shadow-rate shocks explain over 80% of the contemporaneous declines in yields at this maturity, and at maturities

of less than five years they explain nearly all of the change (columns 2 and 3). The expectations component constitues the bulk of the effects of the shadow-rate shocks, but the term-premium effects of such shocks are also significant. They account for 63 basis points of the decline in yields at the ten-year maturity—about twice the effects of the balance-sheet shocks.

In panel B, the nonlinear interactions in column 6 loom larger, so that the precise contributions of each factor to the overall variance of yields are somewhat less clear. Nonetheless, in the case with Fed balance-sheet shocks alone (column 4), yields beyond the two-year horizon have only about 1% of the variance that they have when both shocks are present. In contrast, in the case with shadow-rate shocks alone (columns 2 and 3), the variance of yields is similar to the variance when both shocks are present. At intermediate and long maturities, the term-premium effects of the shadow-rate shocks explain about the same fraction of the total variance in yields that the expectations component explains.

In summary, the model—which is consitent with the evidence on unconventional policy in most verifiable respects—suggests that shadow-rate shocks by themselves can explain most of the response of yields to unconventional policy. The duration channel is relatively unimportant. Furthermore, although the shadow rate's effect on rate expectations is the single most important driver of yields, its effect on term premia is also significant. Indeed, the results suggest that the term-premium effects of the reduced short-rate uncertainty at the ELB—a previously overlooked channel of unconventional policy—may be at least as important as the much-discussed duration channel of QE.

5.2 Policy options in alternative environments

Because they take different units, it is not meaningful to ask whether the shadow-rate shocks or the bond-supply shocks are "more powerful" in general. However, one can compare their *relative* effectiveness in different states of the world. One reason that such a comparison may be interesting is that policymakers, who presumably have some notion of the implict cost of implementing each type of policy, may favor one over the other depending on the circumstances.

To measure relative efficacy, I compute the size of the bond-supply shock that would be required to generate the same effect on the τ-period yield that a 25-basis-point decline in the shadow rate has. Specifically, again letting $y^{(\tau)}\left(\widehat{r}_t, \beta_t, Q_t\right)$ denote

the τ-maturity yield as a function of the state variables, I solve for $\Delta\beta$ such that

$$y^{(\tau)}\left(\widehat{r}_t - .0025, \beta_t, Q_t\right) = \ y^{(\tau)}\left(\widehat{r}_t, \beta_t + \Delta\beta, Q_t\right) \tag{23}$$

using a range of initial values ($\widehat{r}_t$, β_t). I repeat an analagous exercise to solve for the relative efficacy of Fed balance-sheet shocks, ΔQ.[26] In an affine model, the values $\Delta\beta$ and ΔQ that solve these equations are constant across the state space. In the shadow-rate model, as was evident in Figures 4 and 5, the elasticities of yields with respect to shadow-rate and bond-supply shocks differ in different areas of the state space, and therefore their relative efficacy also differs.

Figure 8 presents contour maps of relative efficacy for 10- and 15-year yields, with darker colors indicating bigger values—i.e., areas of the space in which bond-supply shocks have relatively large effects compared to those of shadow-rate shocks. Both β_t and Q_t achieve their greatest relative efficacy in the southwest quadrant of the maps, where both $\widehat{r}_t$ and β_t are deeply negative. As noted earlier, both bond-supply and shadow-rate shocks are attenuated when the shadow rate is below the ELB. However, when β_t is negative (i.e., more duration in the market), the attenuation of the shadow-rate shocks is greater than the attenuation of the bond-supply shocks.[27] Thus, for example, a shock to β_t of about 0.11 or a shock to Q_t of about 0.15 in this region is sufficient to lower the ten-year yield by the same amount that a 25-basis-point shock to $\widehat{r}_t$ would achieve. In contrast, at the unconditional means of the states, the respective sizes of the β_t and Q_t shocks required are closer to 0.16 and 0.20, respectively.

Interestingly, this high-relative-efficacy region for the bond-supply shocks is approximately the region of the space in which the Fed asset purchases were conducted in practice. The greatest removal of duration from the market occurred during the QE and maturity extension programs that mostly operated between 2011 and 2013. During that time, empirical shadow-rate term-structure models show $\widehat{r}_t$ near its nadir, with the Krippner (2012) estimate, for example, averaging -4.5% over those three years. Meanwhile, the Treasury was lengthening the maturities of its issuance, so that the average duration outstanding stood near the upper end of its historical range. Moreover, fiscal expansion increased the total quantity of Treasury debt outstanding, further boosting the amount of interest-rate risk held by investors.[28] Thus, one possible interpretation

[26] Note that the solution for $\Delta\beta$ is the same regardless of whether we use the baseline process (13) or include Q_t in the model as in (20). In the latter case, the initial value of $Q_t$is set to zero.

[27] This result can be seen to some extent in Figure 5, panel B.

[28] Over the 2011 - 2013 period, the maturity-weighted debt-to-GDP ratio averaged 4.4, compared

of the Fed's actions during this time is that it saw the cost-benefit calculations around its policy options changing. During normal times, the Fed has a revealed preference for not engaging in asset purchases. This preference may have shifted during the ELB period if the FOMC perceived that the marginal benefits of forward guidance declined sufficiently relative to those of asset purchases.

6 Conclusion

This paper has augmented a model of risk-averse arbitrage in the bond market to account for the effective lower bound on nominal interest rates. The model successfully reproduces the conditional moments of the yield curve, particularly near the ELB, as well as empirical evidence on the effects of bond supply on yields. When considering shocks that approximate the experience of unconventional monetary policy in the U.S., the main finding is that the majority of the effects of such policies come through the expectations component of yields. The term premium effects of changes in policy expectations—a channel that does not exist in affine models and has been ignored by previous literature—also plays a significant role. The duration effects of bond-supply shocks are relatively weak, accounting for less than one-fifth of the overall change in the ten-year yield and even less at shorter maturities.

The Fed bought about $3 trillion of longer-term bonds, and the model suggests that this may have reduced the ten-year yield through the duration channel by about 30 basis points. Of course, many empirical studies suggest that QE had much larger effects—Williams (2014), for example, cites a consenus estimate of 15 - 25 basis points on the ten-year yield per $600 billion of asset purchases. Such estimates generally come from two sources. One source is models (reduced-form and structural) that are often both linear and parameterized to data from the pre-ELB environment. The results presented here suggest that the predictions of such models could be quite misleading because they do a poor job of matching the behavior of yields near the ELB, where the relevant factor loadings change both quantiatively and qualitatively. The second type of evidence comes from event studies of unconventional policy. While such studies usually point to a sizeable impact of asset-purchase annonucements, they cannot easily distinguish duration effects from implict signaling effects or from the effects of explicit forward guidance that often was announced simultaneously with changes in balance-

to a pre-ELB average of 2.4, according to the CRSP data.

sheet policy. Indeed, as was shown above, the model in this paper replicates the general features of the event-study evidence without the need for large duration effects.[29]

From a policy perspective, the results presented here are a mixed bag. On the one hand, the weakness of the duration channel suggests that monetary policymakers finding themselves at the ELB might be hesitant to rely on asset purchases if they have other tools—such as forward guidance—at their disposal. On the other hand, the results of Section 5 suggest that, in states of the world in which the ELB is a serious constraint, asset purchases may become more attractive, particularly if the quantity of duration risk in the market is high. It may also be the case that asset purchases provide a commentment device and thus have effects on the shadow rate through the signaling channel. Determining the optimal mix of policy tools at the ELB requires careful evaluation of these various mechanism, and my hope is that the results presented above are a useful further step in that direction.

References

Altavilla, C., G. Carboni, and R. Motto (2015): "Asset Purchase Programmes and Financial Markets: Lessons from the Euro Area," European Central Bank Working Papers, no 1864.

Altavilla, C., and D. Giannone (forthcoming): "The Effectiveness of Non-Standard Monetary Policy Measures: Evidence from Survey Data," *Journal of Applied Econometrics*.

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

Bhattarai, S., G. B. Eggertsson, and B. Gafarov (2015): "Time Consistency and the Duration of Government Debt: A Signalling Theory of Quantitative Easing," NBER Working Paper 21336.

[29] It may also be that the event studies pick up channels beyond short-rate expectations and duration risk. For example, D'Amico and King (2013) and Cahill et al. (2015) find a large "local supply" effect of Treasury purchases that is orthogonal to the duration channel. Incorporating such mechanisms into strucutral models is an important direction for future research.

Cahill, M. E., S. D'Amico, C. Li, and J. S. Sears (2013): "Duration Risk versus Local Supply Channel in Treasury Yields: Evidence from the Federal Reserve's Asset Purchase Announcements," Federal Reserve Board FEDS working paper 2013-35.

Carpenter, S. B., J. E. Ihrig, E. C. Klee, D. W. Quinn, and A. H. Boote (2013): "The Federal Reserve's Balance Sheet and Earnings: A Primer and Projections," Federal Reserve Board FEDS Working paper 2013-1.

Carvalho, C., E. Hsu, and F. Nechio (2016): "Measuring the Effect of the Zero Lower Bound on Monetary Policy," Federal Reserve Bank of San Francisco Working Paper 2016-06.

D'Amico, S., W. English, D. Lopez-Salido, and E. Nelson (2012): "The Federal Reserve's Large-scale Asset Purchase Programmes: Rationale and Effects," *Economic Journal*, 112(564), F415–46.

D'Amico, S., and T. B. King (2013): "Flow and Stock Effects of Large-Scale Treasury Purchases: Evidence on the Importance of Local Supply," *Journal of Financial Economics*, 108(2), 425–448.

Femia, K., S. Friedman, and B. Sack (2013): "The Effects of Policy Guidance on Perceptions of the Fed's Reaction Function," Federal Reserve Bank of New York, Staff Report no.652.

Gagnon, J. E., M. Raskin, J. Remache, and B. P. Sack (2011): "The Financial Market Effects of the Federal Reserve's Large-Scale Asset Purchases," *International Journal of Central Banking*, 7(1), 3–43.

Gilchrist, S., D. Lopez-Salido, and E. Zakrajsek (2015): "Monetary Policy and Real Borrowing Costs at the Zero Lower Bound," *American Economic Journal: Macroeconomics*, 7(1), 77–109.

Greenwood, R., S. Hanson, J. S. Rudolph, and L. H. Summers (2015): "Chapter 1: The Optimal Maturity of Government Debt; Chapter 2: Debt Management Conflicts between the U.S. Treasury and the Federal Reserve," Book chapters in in The $13 Trillion Question: How America Manages Its Debt, edited by David Wessel, 43-89. Brookings Institution Press.

Greenwood, R., S. Hanson, and D. Vayanos (2015): "Forward Guidance in the Yield Curve: Short Rates versus Bond Supply," NBER Working Paper No. 21750.

Greenwood, R., and D. Vayanos (2014): "Bond Supply and Excess Bond Returns," *Review of Financial Studies*, 27(3), 663–713.

Gurkaynak, R. S., B. Sack, and J. H. Wright (2007): "The U.S. Treasury Yield Curve: 1961 to the Present," *Journal of Monetary Economics*, 54(8), 2291–304.

Haddad, V., and D. A. Sraer (2015): "The Banking View of Bond Risk Premia," Working Paper.

Hamilton, J. D., and J. C. Wu (2012): "The Effectiveness of Alternative Monetary Policy Tools in a Zero Lower Bound Environment," *Journal of Money, Credit and Banking*, 44(s1), 3–46.

Hanson, S. G. (2014): "Mortgage Convexity," *Journal of Financial Economics*, 113(2), 270–299.

Hattori, M., A. Schrimpf, and V. Sushko (2016): "The Response of Tail-Risk Perceptions to Unconventional Monetary Policy," *American Economic Journal: Macroeconomics*, 8(2), 111–36.

Hayashi, F. (2016): "Affine Term Structure Pricing with Bond Supply as Factors," FRB Atlanta CQER Working Paper 2016-1.

Joyce, M. A. S., A. Lasaosa, I. Stevens, and M. Tong (2011): "The Financial Market Impact of Quantitative Easing in the United Kingdom," *International Journal of Central Banking*, 7(3), 32–49.

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

King, T. B. (2015): "A Portfolio-Balance Approach to the Nominal Term Structure," FRB Chicago Working Paper 2013-18.

Krippner, L. (2012): "Modifying Gaussian Term Structure Models When Interest Rates Are Near the Zero Lower Bound," CAMA Working Paper 2012-05.

Krishnamurthy, A., and A. Vissing-Jorgensen (2011): "The Effects of Quantitative Easing on Interest Rates: Channels and Implications for Policy," *Brookings Papers on Economic Activity*, (2), 215–265.

Malkhozov, A., P. Mueller, A. Vedolin, and G. Venter (2016): "Mortgage Risk and the Yield Curve," *Review of Financial Studies*, 29(5), 1220–53.

Monfort, A., F. Pegoraro, J.-P. Renne, and G. Roussellet (forthcoming): "Staying at Zero with Affine Processes: An Application to Term Structure Modeling," *Journal of Econometrics*.

Newey, W. K., and K. D. West (1987): "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55(3), 703–8.

Rogers, J. H., C. Scotti, and J. H. Wright (2014): "Evaluating Asset-Market Effects of Unconventional Monetary Policy: A Cross-Country Comparison," Federal Reserve Board International Finance Discussion Paper 1101.

Swanson, E. T. (2011): "Let's Twist Again: A High-Frequency Event Study Analysis of Operation Twist and Its Implications for QE2," *Brookings Papers on Economic Activity*, (1), 151–88.

——— (2017): "Measuring the Effects of Federal Reserve Forward Guidance and Asset Purchases on Financial Markets," NBER Working Paper 21816.

Swanson, E. T., and J. C. Williams (2014): "Measuring the Effect of the Zero Lower Bound on Medium- and Longer-Term Interest Rates," *American Economic Review*, 104(10), 3154–3185.

Vayanos, D., and J.-L. Vila (2009): "A Preferred-Habitat Model of the Term Structure of Interest Rates," NBER Working Paper 15487.

Williams, J. C. (2014): "Monetary Policy at the Zero Lower Bound: Putting Theory into Practice," Hutchins Center on Fiscal and Monetary Policy at Brookings, January 16.

Woodford, M. (2012): "Methods of Policy Accommodation at the Interest-Rate Lower Bound," Federal Reserve Bank of Kansas City, Jackson Hole symposium, September 16.

Wright, J. (2012): "What Does Monetary Policy Do to Long-Term Interest Rates at the Zero Lower Bound?," *Economic Journal*, 122, F447–66.

Wu, J. C., and F. D. Xia (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. Solution Algorithm

Transform log prices from the time/maturity domain to the state-space domain by writing them as a function $p(.)$:

$$p_t^{(\tau)} = p\left(\tau, \widehat{r}_t, \beta_t\right) \tag{24}$$

We want to solve for this function.

Discretize the state and maturity space into $N_\tau \times N_{\widehat{r}} \times N_\beta$ nodes. Let τ^n, $\widehat{r}^n$, and β^n denote the values of the maturity and state variables at node n. Set $p\left(1, \widehat{r}_t, \beta_t\right) = -\max\left[\widehat{r}_t, b\right]$ for all β^n. Then, from equations (3), (13), and (15), for $\tau > 1$ we have

$$\begin{aligned} p\left(\tau^n, \widehat{r}^n, \beta^n\right) &= E\left[p\left(\tau^n - 1, \widehat{r}_{t+1}, \beta_{t+1}\right)|\widehat{r}^n, \beta^n\right] + \widehat{r}^n \\ &\quad + \{a\sum_{\tau'=1}^{N_\tau}\left[\zeta + \left(1 - \frac{2\tau'}{T}\right)\beta^n\right] \\ &\quad \times \operatorname{cov}\left[p\left(\tau', \widehat{r}_{t+1}, \beta_{t+1}\right), p\left(\tau^n - 1, \widehat{r}_{t+1}, \beta_{t+1}\right)|\widehat{r}^n, \beta^n\right]\} \end{aligned} \tag{25}$$

Rewrite the law of motion for the state vector as the conditional probability function

$$\pi\left(\widehat{r}_{t+1}, \beta_{t+1}|\widehat{r}_t, \beta_t\right) = \varphi\left(\widehat{r}_{t+1} - \mu^{\widehat{r}}(1 - \phi^{\widehat{r}}) - \phi^{\widehat{r}}\widehat{r}_t\right)\varphi\left(\beta_{t+1} - \phi^\beta \beta_t\right) \tag{26}$$

where $\varphi(.)$ is the standard-normal PDF (and the separability makes use of the independence of $\widehat{r}_t$ and β_t).

The algorithm proceeds as follows:

Step 0. Set $i = 0$. Begin with an initial guess of the pricing function $p^0(.)$. For example, choose $p^0(\tau, \widehat{r}_t, \beta_t) = -\max\left[\widehat{r}_t, b\right]$ for all τ, $\widehat{r}_t$, β_t.

Step 1. At each node n, evaluate the functions

$$\begin{aligned} m^i\left(\tau^n, \widehat{r}^n, \beta^n\right) &\equiv \int\int \pi\left(\widehat{r}, \beta|\widehat{r}^n, \beta^n\right) p^i\left(\tau - 1, \widehat{r}, \beta\right) d\widehat{r}d\beta \\ &= E\left[p^i\left(\tau^n - 1, \widehat{r}_{t+1}, \beta_{t+1}\right)|\widehat{r}^n, \beta^n\right] \end{aligned} \tag{27}$$

and

$$\begin{aligned}\omega^i\left(\tau^n,\tau',\widehat{r}^n,\beta^n\right) &\equiv \int\int\{\pi\left(\widehat{r},\beta|\widehat{r}^n,\beta^n\right)\left[p^i\left(\tau-1,\widehat{r},\beta\right)-m^i\left(\tau,\widehat{r}^n,\beta^n\right)\right] \\ &\quad\times\left[p^i\left(\tau',\widehat{r},\beta\right)-m^i\left(\tau',\widehat{r}^n,\beta^n\right)\right]\}d\widehat{r}d\beta \\ &= \operatorname{cov}\left[p^i\left(\tau',\widehat{r}_{t+1},\beta_{t+1}\right),p^i\left(\tau^n-1,\widehat{r}_{t+1},\beta_{t+1}\right)|\widehat{r}^n,\beta^n\right]\end{aligned} \tag{28}$$

with both functions set to zero when $\tau = 1$.

Step 2. Update the pricing function by calculating, at each node,

$$\begin{aligned}p^{i+1}\left(\tau^n,\widehat{r}^n,\beta^n\right) &= m^i\left(\tau^n,\widehat{r}^n,\beta^n\right)-\max\left[\widehat{r}_t,b\right] \\ &\quad+a\sum_{\tau'=1}^{N_\tau}\left[\zeta+\left(1-\frac{2\tau'}{T}\right)\beta^n\right]\omega^i\left(\tau^n,\tau',\widehat{r}^n,\beta^n\right)\end{aligned} \tag{29}$$

Set $i = i + 1$.

Repeat steps (1) and (2) to convergence.

The expectations in Step 1 are computed numerically using the probability function $\pi(.)$ and the pricing function $p^i(.)$. The integration is performed by quadrature and, to ensure accuracy, relies on a much finer grid than the price computation in Step 2 does. To obtain bond prices over this refinement of the space, the values of $p^i(.)$ are interpolated between each pair of nodes, at each iteration, using a cubic spline. At the edges of the discretized space, to avoid explosive behavior, prices are log-linearly extrapolated for the purposes of computing expectations. (So long as the edges are far away from the region of the space that is being considered, the conditional expectations used there have little influence on the results.)

In the baseline model of the paper, I use $N_\tau = 60$, $N_{\widehat{r}} = 101$, and $N_\beta = 25$, for a total of $151,500$ nodes, distributed uniformly in each dimension over the intervals $\tau = [1, 60]$, $\widehat{r} = [-0.25, 0.35]$, and $\beta = [-6.0, 6.0]$. Expanding the density of the nodes or their range beyond this point had no noticeable effect on the results reported in the paper. The algorithm converges to three significant digits in approximately 400 iterations.

Appendix B. Simulation details

This appendix describes the construction of the simulated distributions of the state variables used in the pseudo event study, as depicted in Figure 6.

Clearly, it had to be the case that the average shadow-rate shock during the ELB period was negative. Thus, for each simulation i, I draw a series of shadow-rate shocks $\{e^{\widehat{r}}_{i,1}, ..., e^{\widehat{r}}_{i,28}\}$ from the distribution $N[\mu_i, \sigma_{\widehat{r}}]$, where $\mu_i < 0$ is chosen to make the terminal value of the shadow rate $\widehat{r}_{i,28}$ exactly equal b. I reject any draw in which the simulated value of $\widehat{r}_t$ ever rises above b.

To determine the size of the Fed balance-sheet shocks, note that equation (17) implies that changes in the supply factors translate into percentage changes in 10-year-equivalent bonds as follows:

$$\begin{aligned} \%\Delta 10\text{YE}_{t+h} &= \frac{\int_0^T \tau \left[\left(1-\frac{2\tau}{T}\right)\left(\beta_{t+h}+Q_{t+h}-\beta_t-Q_t\right)\right] d\tau}{\int_0^T \tau \left[\zeta+\left(1-\frac{2\tau}{T}\right)\left(\beta_t+Q_t\right)\right] d\tau} \\ &= -\frac{\Delta\beta_{t+h}+\Delta Q_{t+h}}{3\zeta-\beta_t-Q_t} \end{aligned} \tag{30}$$

Since the e^{β}_t are taken to be zero in this exercise, plugging in the actual percentage change in 10-year equivalents that resulted from QE allows one to uniquely solve for ΔQ_{t+h}, given initial values. In particular, at the end of the simulation we must have $Q_{28} = 0.23$ in order to achieve a reduction in ten-year equivalents of 18% relative to a case in which the $Q_{28} = 0$.[30]

This calculation provides a value for the cumulative effect of the Fed balance-sheet shocks on bond supply, but it does not tell us about the individual values of those shocks. I take a conservative approach by considering the widest possible distribution for the shocks, while respecting the empirical facts that (1) QE balances never fell below their starting value of zero, and (2) QE attained its maximum value at the end of the ELB period. Specifically, for each simulation i, I take draws $\{\widetilde{e}_{i,1}, ..., \widetilde{e}_{i,28}\}$ from $N[0,1]$ and compute the balance-sheet shocks $\{e^{Q}_{i,1}, ..., e^{Q}_{i,28}\} = \sigma_i\{\widetilde{e}_{i,1}, ..., \widetilde{e}_{i,28}\}$, where σ_i is the the largest value that is consistent with $\min[\{Q_{i,1}, ..., Q_{i,28}\}] > 0$ and $\max[\{Q_{i,1}, ..., Q_{i,28}\}] = Q_{i,28}$.

[30] The difference between the path of ten-year equivalents in the simulation and the counterfactual case in which no QE occurs is $-\frac{\Delta Q_{t+h}}{3\zeta-\beta_t-Q_t} = -\frac{Q_{28}}{3(0.31)+0.21}$. Setting this equal to -0.18 gives $Q_{28} = 0.23$.

Table 1. Regressions of long-term yields on Treasury duration and short-term yields

	Independent variables						Adj. R^2
	WAM of Treas. debt			1y yield			
Dep. Var.	Pre-ELB	ELB	*Break t-stat*	Pre-ELB	ELB	*Break t-stat*	
5y yield	0.140 (0.095)	0.002 (0.101)	*-2.17*	0.842*** (0.050)	2.271*** (0.785)	*1.84*	0.951
10y yield	0.221* (0.121)	0.058 (0.116)	*-2.25*	0.736*** (0.060)	3.028** (1.203)	*1.92*	0.901
15y yield	0.261* (0.133)	0.110 (0.126)	*-2.05*	0.688*** (0.065)	2.966** (1.276)	*1.80*	0.870

	Independent variables						Adj. R^2
	WAM of Treas. debt			2y yield			
Dep. Var.	Pre-ELB	ELB	*Break t-stat*	Pre-ELB	ELB	*Break t-stat*	
5y yield	0.102* (0.373)	-0.002 (0.060)	*-2.57*	0.901*** (0.032)	1.910*** (0.217)	*4.74*	0.981
10y yield	0.187** (0.094)	0.053 (0.088)	*-2.25*	0.794*** (0.048)	2.328*** (0.429)	*3.61*	0.942
15y yield	0.227** (0.109)	0.113 (0.108)	*-1.62*	0.746*** (0.056)	2.167*** (0.537)	*2.68*	0.915

	Independent variables						Adj. R^2
	Maturity-weighted debt/GDP			1y yield			
Dep. Var.	Pre-ELB	ELB	*Break t-stat*	Pre-ELB	ELB	*Break t-stat*	
5y yield	0.179* (0.010)	-0.058 (0.083)	*-2.65*	0.850*** (0.049)	2.097*** (0.679)	*1.89*	0.952
10y yield	0.250* (0.129)	-0.056 (0.093)	*-2.81*	0.743*** (0.062)	2.940*** (1.043)	*2.15*	0.902
15y yield	0.282** (0.140)	-0.027 (0.101)	*-2.74*	0.696*** (0.070)	2.962*** (1.128)	*2.06*	0.871

	Independent variables						Adj. R^2
	Maturity-weighted debt/GDP			2y yield			
Dep. Var.	Pre-ELB	ELB	*Break t-stat*	Pre-ELB	ELB	*Break t-stat*	
5y yield	0.126* (0.067)	-0.044 (0.051)	*-2.86*	0.906*** (0.032)	1.818*** (0.205)	*4.77*	0.981
10y yield	0.207** (0.104)	-0.058 (0.076)	*-2.80*	0.800*** (0.050)	2.301*** (0.698)	*4.32*	0.943
15y yield	0.239** (0.118)	-0.013 (0.095)	*-2.43*	0.751*** (0.600)	2.218*** (0.463)	*3.34*	0.915

Notes: Each row in each table reports the estimates of a single regression, where the dependent variable is a longer-term Treasury yield, as indicated in the first column. Each regression uses two independent variables: either the weighted-average maturity of Treasury debt in public hands or the maturity-weighted Treasury-debt-to-GDP ratio and either the one- or two-year zero-coupon Treasury yield. In each regression, the coefficient on each variable is allowed to differ between the period when the ELB was not binding (prior to December 2008) and the period when it was binding (December 2008 through December 2015), with the break accomplished using interactive dummy variables. The samples begin in August 1971 for the 5- and 10-year maturities and in December 1971 for the 15-year maturity. Yield data are Gurkaynak et al. (2007) zero-coupon yields. Treasury debt variables are constructed from CRSP data, following Greenwood and Vayanos (2014). All data are monthly. Newey-West standard errors, using 36 lags, are reported in parentheses, and statistical significance at the 10% (*), 5% (**), and 1% (***) levels is indicated by asterisks. The *t* statistics, reported in italics, test the significance of the break in each of the two coefficients in each regression.

Table 2. Model calibration

	Bond supply				Short rate				Risk aversion
	T	ϕ_β	σ_β	ζ	$\mu_{\hat{r}}$	$\phi_{\hat{r}}$	$\sigma_{\hat{r}}$	b	a
[1] Shadow-rate model	60	0.98	0.20	0.31	5.0%	0.98	0.78%	0.17%	0.15
[2] Affine model – base calibration	60	0.98	0.20	0.31	5.0%	0.98	0.78%	-∞	0.15
[3] Affine model – recalibrated	60	0.98	0.20	0.32	5.2%	0.98	0.68%	-∞	0.17

Notes: The table shows the calibrated values of the parameters in the baseline shadow-rate model, as well as in two models with affine short-rate processes. In the first affine model, all parameters (except the ELB) are the same as in the shadow-rate model. In the second affine model, the parameters are recalibrated to match the same set of unconditional yield-curve moments that the shadow-rate model matches. Details of the calibration are described in the text.

Table 3. Conditional moments of yield curve in data vs. models

Short rate below 0.68%

			Slopes (to 3m)			
	% of obs.	3m rate	2Y	5Y	10Y	15Y
Conditional means						
Data	*16%*	*0.2%*	*0.3%*	*1.3%*	*2.5%*	*3.1%*
Shadow-rate model	13%	0.2%	0.4%	1.3%	2.6%	3.8%
Affine Model – base calibration	13%	-1.3%	0.7%	1.9%	3.6%	4.9%
Affine Model – recalibrated	10%	-1.0%	0.7%	1.8%	3.3%	4.7%
Conditional standard deviations						
Data		*0.1%*	*0.3%*	*0.6%*	*0.8%*	*0.8%*
Shadow-rate model		0.1%	0.3%	0.7%	1.1%	1.4%
Affine Model – base calibration		1.7%	0.3%	0.7%	1.3%	1.7%
Affine Model – recalibrated		1.5%	0.3%	0.7%	1.2%	1.6%

Short rate above 0.68%

			Slopes (to 3m)			
	% of obs.	3m rate	2Y	5Y	10Y	15Y
Conditional means						
Data	*84%*	*6.1%*	*0.5%*	*0.9%*	*1.3%*	*1.5%*
Shadow-rate model	87%	6.0%	0.3%	0.7%	1.4%	2.0%
Affine Model – base calibration	87%	6.0%	0.2%	0.7%	1.3%	1.9%
Affine Model – recalibrated	90%	5.9%	0.3%	0.7%	1.3%	1.9%
Conditional standard deviations						
Data		*3.1%*	*0.9%*	*1.3%*	*1.6%*	*1.7%*
Shadow-rate model		3.2%	0.4%	0.9%	1.5%	2.0%
Affine Model – base calibration		3.2%	0.4%	0.9%	1.3%	1.9%
Affine Model – recalibrated		3.0%	0.3%	0.8%	1.5%	1.9%

Notes: The table shows conditional moments of zero-coupon yields simulated from the shadow-rate and affine models, based on the calibrations shown in Table 2, together with the corresponding moments from the data. Model results are based on 1 million simulations of the state variables. Yield data are from the Gurkaynak et al. (2007) dataset and cover the period August 1971 – December 2015, except for the 15-year yields, which begin in December 1971.

Table 4. Model-implied relationships among long-term yields, shorter-term yields, and Treasury supply

A. Sensitivity of long-term yields to WAM

	Shadow rate	Sensitivity to WAM			Sensitivity to WAM, holding 1Y yield fixed			Sensitivity to WAM, holding 2Y yield fixed		
		5Y	10Y	15Y	5Y	10Y	15Y	5Y	10Y	15Y
	8%	0.09	0.15	0.20	0.07	0.14	0.19	0.06	0.13	0.17
	4%	0.09	0.15	0.19	0.07	0.14	0.18	0.05	0.12	0.17
	2%	0.08	0.14	0.19	0.07	0.13	0.18	0.05	0.12	0.17
ELB	1%	0.08	0.14	0.19	0.06	0.13	0.18	0.05	0.12	0.16
	0%	0.07	0.13	0.18	0.06	0.12	0.17	0.04	0.11	0.16
	-1%	0.06	0.12	0.17	0.05	0.11	0.16	0.03	0.10	0.15
	-2%	0.04	0.11	0.16	0.04	0.10	0.15	0.03	0.09	0.14
	-4%	0.02	0.08	0.14	0.02	0.08	0.13	0.01	0.07	0.12

Notes: The first group of columns reports the model-implied sensitivity of 5-, 10-, and 15-year yields to the weighted-average maturity of Treasury debt at various values of the shadow rate. The second and third groups of columns report these sensitivities, holding fixed the level of the one-year or the two-year yield. The calculations, which are given in the text, rely on the derivatives of yields with respect to the shadow rate and the bond supply factor ($A_{\hat{r}}^{(\tau)}$ and $A_{\beta}^{(\tau)}$), evaluated under the baseline calibration (line 1 of Table 2).

B. Sensitivity of long-term yields to shorter-term yields

	Shadow rate	Sensitivity to 1Y			Sensitivity to 2Y		
		5Y	10Y	15Y	5Y	10Y	15Y
	8%	0.9	0.7	0.6	0.9	0.7	0.6
	4%	0.9	0.7	0.6	0.9	0.7	0.6
	2%	0.8	0.7	0.6	0.9	0.8	0.6
ELB	1%	0.8	0.7	0.6	0.9	0.8	0.7
	0%	1.3	1.2	1.1	1.1	1.0	1.0
	-1%	2.7	3.0	2.8	1.5	1.7	1.6
	-2%	7.5	9.6	9.6	2.3	2.9	2.9
	-4%	140.1	259.0	291.5	6.6	12.2	13.8

Notes: The table reports the model-implied sensitivity of 5-, 10-, and 15-year yields to 1- and 2-year yields, holding the bond-supply factor fixed, at various values of the shadow rate. The coefficients are calculated as the ratio $A_{\hat{r}}^{(\tau)}/A_{\hat{r}}^{(4)}$ or $A_{\hat{r}}^{(\tau)}/A_{\hat{r}}^{(8)}$ where $A_{\hat{r}}^{(\tau)}$ is the derivative of the yield at maturity τ with respect to $\hat{r}$ under the baseline calibration (line 1 of Table 2).

Table 5. Decompositions of yield responses to unconventional policy shocks

A. Contemporaneous responses (bps)

Maturity [1]	Shadow-rate shocks: Expectations component [2]	Shadow-rate shocks: Term premium component [3]	Fed balance-sheet shocks: Term premium component [4]	Interaction [6]	Total [7]
2 years	-63	-12	-9	5	-79
	(-93, -36)	(-14, -10)	(-10, -8)	(4, 6)	(-109, -49)
5 years	-98	-45	-21	8	-156
	(-117, -73)	(-47, -42)	(-22, -18)	(6, 11)	(-174, -127)
10 years	-114	-63	-34	7	-202
	(-121, -101)	(-70, -54)	(-36, -29)	(5, 11)	(-205, -194)
15 years	-111	-63	-40	6	-208
	(-113, -105)	(-75, -53)	(-44, -35)	(4, 9)	(-213, -201)

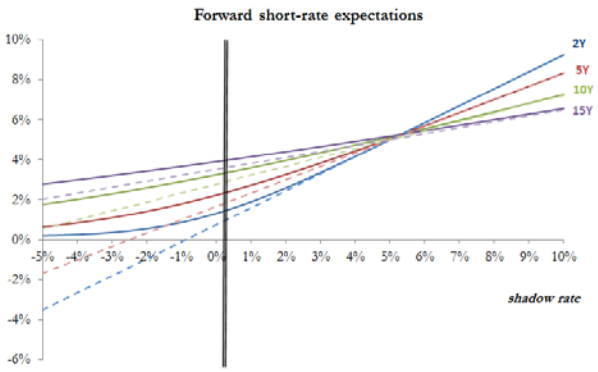
B. Total variance ($bps^2/100$)

Maturity [1]	Shadow-rate shocks: Expectations component [2]	Shadow-rate shocks: Term premium component [3]	Fed balance-sheet shocks: Term premium component [4]	Interaction [6]	Total [7]
2 years	20	7	0.0	-22	27
	(13, 30)	(4, 10)	(0.0, 0.0)	(-33,-12)	(18, 39)
5 years	21	13	0.2	-18	34
	(11, 34)	(6, 24)	(0.1, 0.2)	(-35, -0)	(17, 58)
10 years	19	18	0.4	-13	39
	(9, 35)	(7, 38)	(0.3, 0.5)	(-29, 3)	(18, 74)
15 years	16	17	0.6	-9	36
	(8, 31)	(6, 39)	(0.4, 0.7)	(-22, 2)	(16, 72)

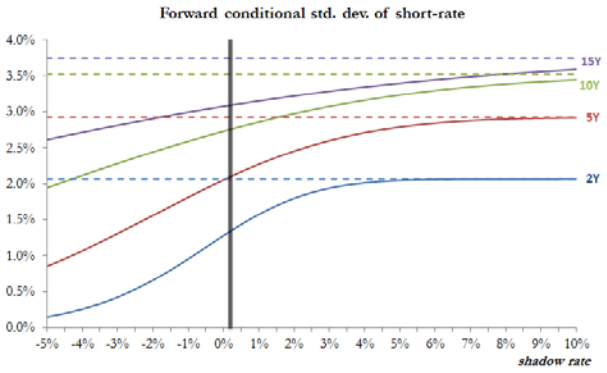
Notes: Panel A reports the cumulative response of the spot zero-coupon yield curve in model simulations based on the distribution of state-variable trajectories shown in Figure 6, summing the responses to the shocks in each period. Panel B reports the total variance in yields in the simulations, relative to a baseline scenario in which no shocks occur. In both cases, for each maturity, the median response is reported, with the 5% and 95% quantiles in parentheses below. The total effect on the yield of each maturity is shown in the last column. The bond-supply and shadow-rate shocks are simulated both separately and together to obtain the decomposition reported in the other columns. The "interaction" column represents the effect of nonlinearities that cause the sum of the two individual simulations to differ from that of the joint simulation. For the shadow-rate shocks, the change in the expectations component is calculated from equation (11), while the change in the term-premium component is calculated as the difference between the total change in yields and the change in the expectations component. By construction, the individual components sum to the totals in each simulation, but the median values in columns [2] through [6] may not sum to the values in column [7] because of the asymmetry in the distributions across simulations.

Figure 1. Conditional moments of short rate at various horizons

A.

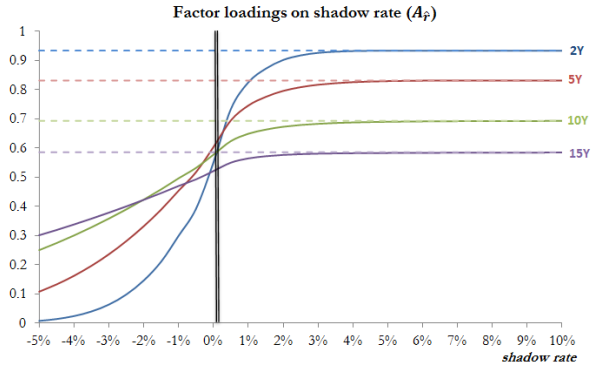


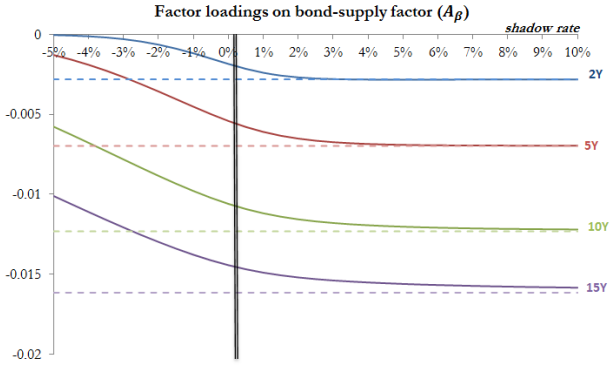
B.



Notes: The figure shows the conditional mean (panel A) and standard deviation (panel B) of the time $t+h$ short rate, conditional on the value of the shadow rate in time t, where h = 2, 5, 10, and 15 years. The solid lines show these moments in the shadow-rate model, under the calibration shown in the top line of Table 2. The dashed lines show the moments in an affine model with the same calibration but with the ELB removed.

Figure 2. Factor loadings across values of the shadow rate at $\beta = 0$

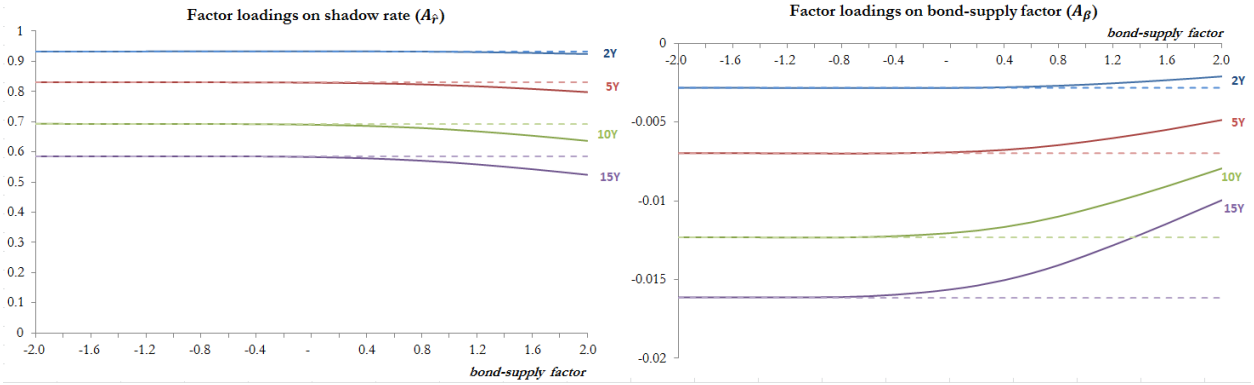




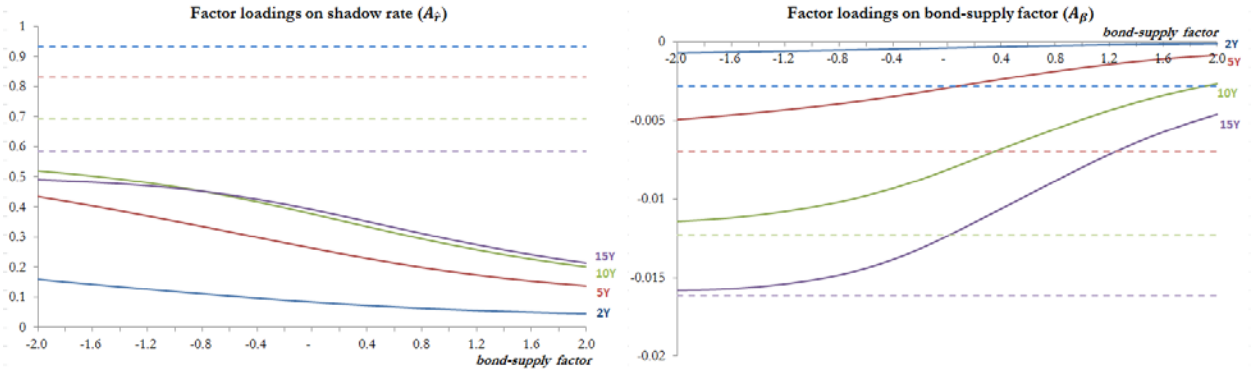
Notes: The figure shows the model-implied factor loadings for the τ-period yield, conditional on the time-t value of the shadow rate, where τ = 2, 5, 10, and 15 years. The solid lines show the loadings in the shadow-rate model, under the calibration shown in line 1 of Table 2. The dashed lines show the loadings in an affine model with the same calibration but with the ELB removed (line 2 of Table 2). The bond-supply factor β_t is held fixed at its mean of zero.

Figure 3. Factor loadings across values of the supply factor

A. $\hat{r}_t$ = 5.2%



B. $\hat{r}_t$ = -2.7%



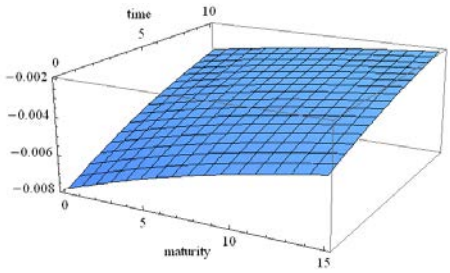
Notes: The figure shows the model-implied factor loadings for the τ-period yield, across time-t values of the bond-supply factor, where τ = 2, 5, 10, and 15 years. The solid lines show the loadings in the shadow-rate model, under the calibration shown in line 1 of Table 2. The dashed lines show the loadings in an affine model with the same calibration but with the ELB removed (line 2 of Table 2). In panel A, the shadow rate is held fixed at the mean value of the short rate (5.2%), while in panel B it is held fixed at -2.7%, its average value during the ELB period according to the Krippner (2012) estimates.

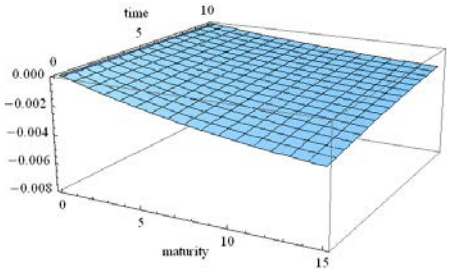
Figure 4. Impulse-response functions

A. Spot yield curve

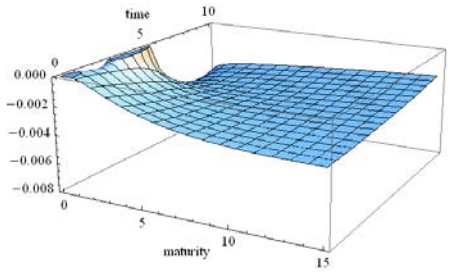
Shadow-rate shocks *Bond-supply shocks*

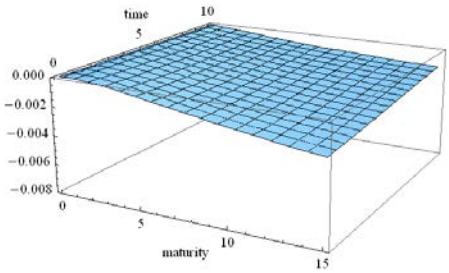
Shadow rate at 5.2%





Shadow rate at -2.7%

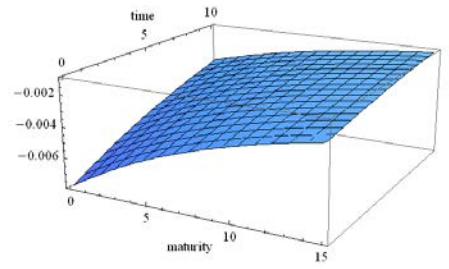


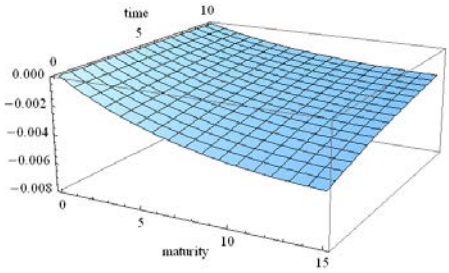


B. Forward rate curve

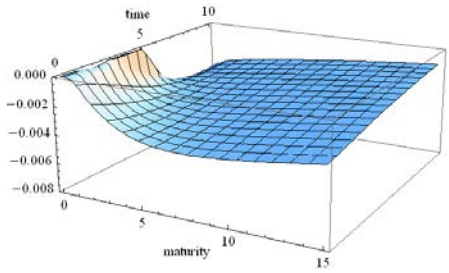
Shadow-rate shocks *Bond-supply shocks*

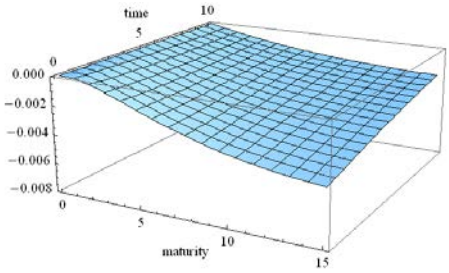
Shadow rate at 5.2%





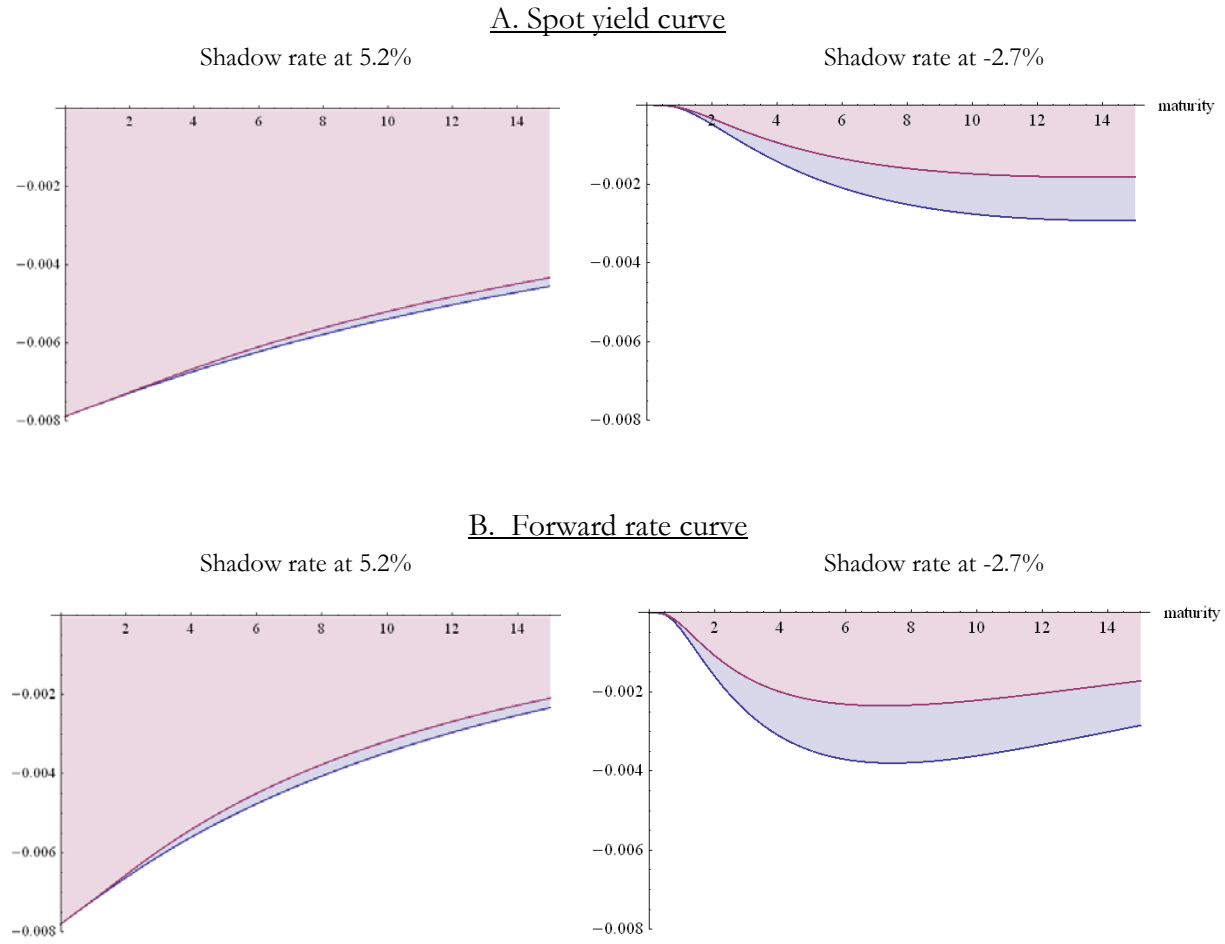
Shadow rate at -2.7%





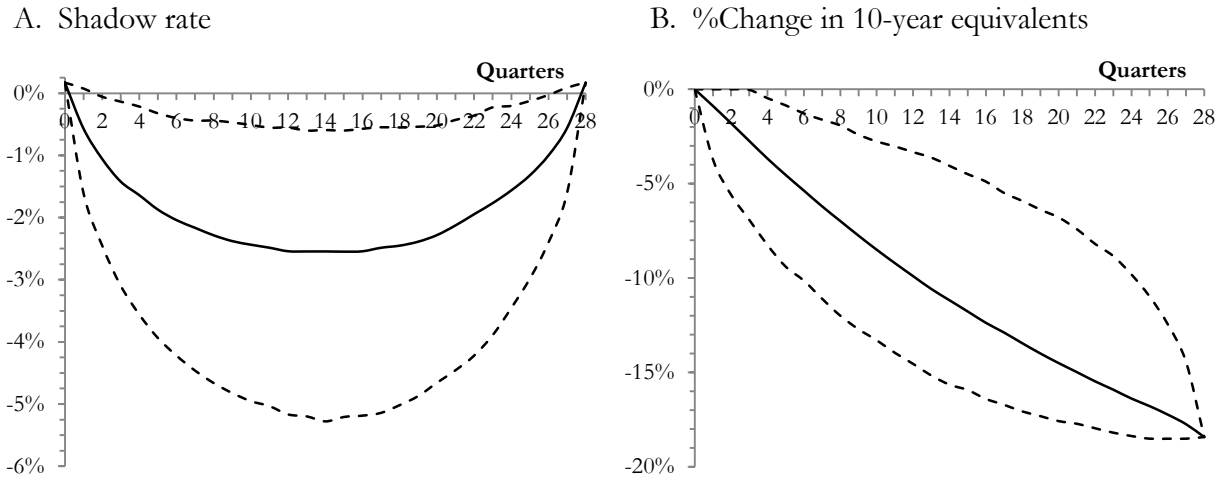
Notes: The figure shows the model-implied responses of yields (panel A) and forward rates (panel B) to one-standard-deviation shocks to each of the two factors over the subsequent 10 years. Maturity in years is shown on the lower-left axis in each graph, while calendar time is on the upper-left axis. Responses are evaluated starting both from a shadow rate at the mean value of the short rate (5.2%) and a value of -2.7%, its average during the ELB period according to the Krippner (2012) estimates. In both cases, the starting value of the bond-supply factor is zero.

Figure 5. Decomposition of initial response to shadow-rate shocks



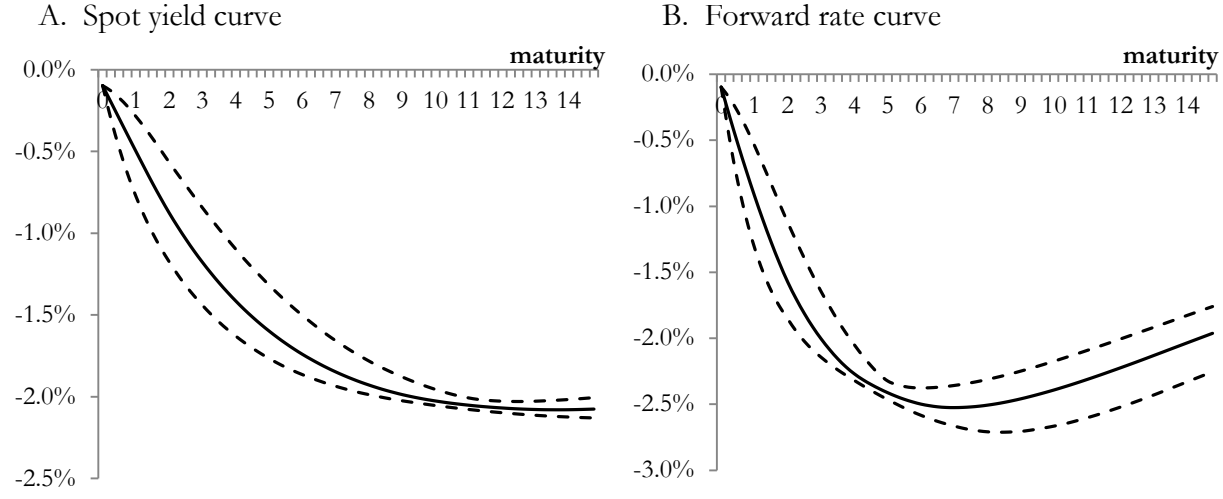
Notes: The figure shows the model-implied response of the yield curve (panel A) and the forward-rate curve (panel B) to one-standard-deviation shock to the shadow rate in the period when the shock occurs. The pink region shows the change in the expectations component of yields, while the blue region shows the change in the term premium. The change in the expectations component is calculated from equation (11), while the change in the term-premium component is calculated as the difference between the total change in yields and the change in the expectations component. Responses are evaluated starting both from a shadow rate at the mean value of the short rate (5.2%) and a value of -2.7%, its average during the ELB period according to the Krippner (2012) estimates. In both cases, the value of the bond-supply factor is constant at zero.

Figure 6. Distributions of state-variable trajectories in model simulations



Notes: The figures show the distributions of 100,000 simulated paths of the state variables during the ELB period, in terms of the pointwise medians (solid lines) and 5% and 95% quantiles (dotted lines). The simulations are constructed to exactly match the observed values of the short-term interest rate and the percentage of ten-year equivalent bonds held by the Fed as of December 2008 and December 2015, with the intermediate values simulated from the model as described in Appendix B. For the purposes of presentation, in panel B the QE state variable (Q_t) is converted to a cumulative percentage change in 10-year equivalent bonds held by investors, relative to the amount that would have been outstanding in the absence of shocks, using equation (17).

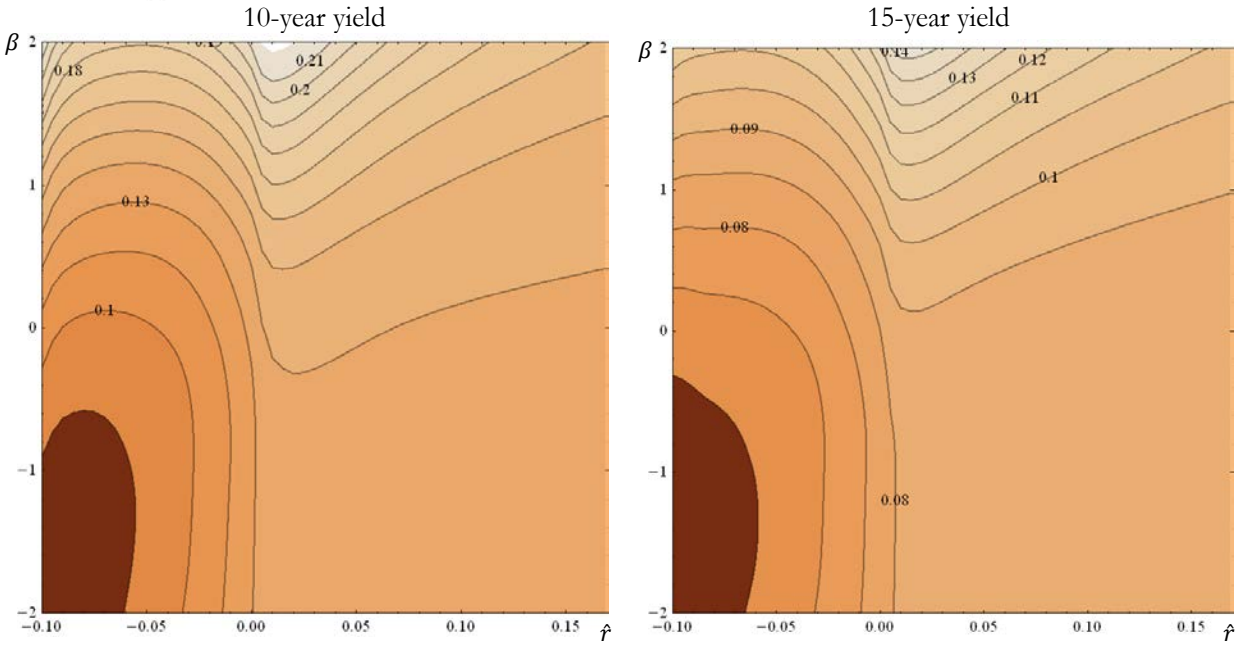
Figure 7. Cumulative yield-curve responses in model simulations



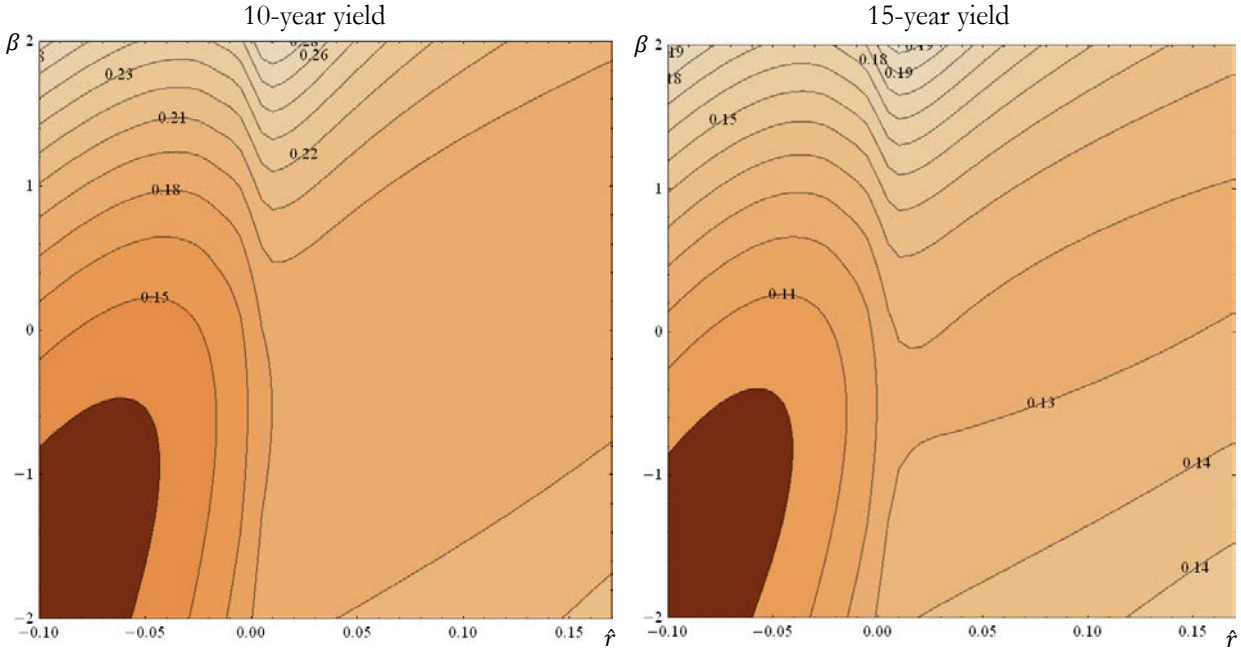
Notes: The figure shows the cumulative response of bond yields (panel A) and forward rates (panel B) in model simulations based on the distribution of state-variable trajectories shown in Figure 6. The figures sum the contemporaneous responses to the shocks in each of the 28 simulated periods. In each panel, the solid line shows the pointwise median and dashed lines show 5% and 95% quantiles.

Figure 8. Relative efficacy of bond-supply shocks across state values

A. Bond-supply factor (β_t)



B. Fed balance-sheet factor (Q_t)



Notes: The figure shows contour maps of the effects of bond-supply shocks on 10- and 15-year yields, relative to the effects of shadow-rate shocks. Relative efficacy is calculated, for each yield, as the size of the bond-supply shock that would be necessary to equal the effects of a -25-basis-point shock to the shadow rate. The values of this ratio are shown across different regions of the state space, with darker coloring indicating regions where the bond-supply shocks are relatively more powerful.