

A Portfolio-Balance Approach to the Nominal Term Structure

Thomas B. King*

August 26, 2014

Abstract

Explanations of why asset prices are affected by changes in the relative quantities of safe debt—including central bank asset purchases—often appeal informally to a “portfolio balance” channel. I show how this mechanism can be incorporated into structural, arbitrage-free models of the yield curve using a solution method that allows for a wide range of nonlinearities. I apply the approach to estimate a model in which investors have preferences over real returns, inflation is heteroskedastic, the short rate is bounded by zero, and the maturity structure of outstanding Treasury debt varies stochastically. This model fits yields and excess returns on nominal bonds well since 1971, and it suggests that the duration of Treasury supply explains a portion of the variation in term premia that is comparable to the portion explained by inflation risk. On the other hand, partly reflecting an attenuation of portfolio-balance effects when interest rates are near zero, the Federal Reserve’s asset purchase programs are estimated to have had a fairly small impact on the yield curve by removing duration from the market.

*Federal Reserve Bank of Chicago. Contact: thomas.king@chi.frb.org. I thank Stefania D’Amico, Jonathan Goldberg, Francisco Vazquez-Grande, and seminar participants at the Federal Reserve Bank of Chicago and the Federal Reserve Board for helpful comments and discussions. Roger Fan provided excellent research assistance. The views expressed here do not reflect official positions of the Federal Reserve.

1 Introduction

The effect of government liability structure on financial conditions is relevant for optimal debt-management policy, the potential role of safe assets in the economy, and the transmission of monetary and fiscal policy through financial markets. It also has implications for theoretical and empirical modeling of interest rates, asset prices, and macrofinancial dynamics. These issues have received particular attention following the recent efforts of several central banks to reduce long-term interest rates by purchasing large quantities of government debt and other securities. In the United States, this policy has taken the form of the Federal Reserve’s Large-Scale Asset Purchase (LSAP) programs, which have, at the time of this writing, removed over \$3 trillion of government-backed debt from the market.

Recent empirical work has been nearly universal in concluding that fluctuations in government debt do have significant effects on the term structure of interest rates and, most likely, on other asset prices.¹ While much of this evidence comes from studies of central-bank asset purchases themselves, enough of it derives from other episodes to demonstrate that the mechanism transmitting debt fluctuations to financial conditions is not unique to recent experience or, for that matter, to monetary-policy interventions. A general theme of this literature is that changes in Treasury debt structure that increase the interest-rate risk borne by investors seem to result in higher long-term yields and greater excess returns on long-term bonds. Yet, despite a few theoretical advances, the profession lacks a comprehensive framework that can explain these phenomena or provide quantitative guidance to policymakers on their relevance.

This paper considers a class of arbitrage-free, rational-expectations models in which the relative amounts of default-free assets held by the public can matter for the term structure of interest rates and other asset prices. The way that this happens is that prices adjust to make investors willing to hold whatever securities are outstanding in each period, given that they know that security prices will be determined in the same way in the next period. Investors do not care about quantities of particular securities, but they do care about the overall risk of their portfolios as reflected in the pricing kernel. Thus, any change in the security mix that increases the aggregate exposure to a given risk factor will raise the prices of securities that depend on that factor. Resuscitating a term that is now somewhat out of fashion, I refer to this mechanism as the “portfolio balance” channel.

The original portfolio-balance models, in the tradition of Tobin (1961), mostly had to do with substitution between money and bonds and have been shunned by more-recent authors because they modeled demand for those instruments in an arbitrary, reduced-form way (see for example, Woodford, 2012).

¹See Bernanke et al. (2004), Kuttner (2006), Gagnon et al. (2010), Greenwood and Vayanos (2010, 2014), Krishnamurthy and Vissing-Jorgensen (2011, 2013), Meaning and Zhu (2011), Swanson (2011), D’Amico et al., (2012), Hamilton and Wu (2012), Ihrig et al. (2012), Joyce et al. (2012), Li and Wei (2012), Rosa (2012), D’Amico and King (2013), Bauer and Rudebusch (2014), Cahill et al. (2014), Chabot and Herman (2014), and Rogers et al. (2014).

My approach retains the intuition that investors may have downward-sloping demand curves for particular classes of assets, but it develops this idea using the machinery of modern financial economics. Investors' willingness to substitute among government liabilities is derived from familiar optimizing behavior and no-arbitrage conditions. I do not assume that there is anything "special" about money or short-term debt, although (as discussed in an appendix) it is straightforward to add terms that allow for such specialness.

While asset-pricing models that rely on some version of the portfolio-balance mechanism are fairly common, most assume either a static environment in which the distribution of asset payoffs is fixed or investor expectations that are not fully rational. (Frankel (1985) and Piazzesi and Schneider (2008) are good examples that are discussed in some detail in the following section.) The reason for these shortcuts is that, when *current* prices depend on payoffs that are themselves *future* prices, solving for the state-contingent price distribution involves a nonlinear recursion that does not generally have an analytical solution. Indeed, the only paper to date that has successfully characterized the equilibrium in such a model of the term structure is Vayanos and Vila (2009). Yet even that model can only be solved in closed form in the extreme cases of zero or infinite risk aversion, despite an assumed linear-Gaussian factor structure that helps to simplify matters. As an alternative, I propose a computational method for solving portfolio-balance models. The method requires only weak conditions on the functional form of the pricing kernel and the dynamics of the state of the economy. Consequently, one can dispense with assumptions that, though perhaps analytically convenient, are likely to be unrealistic for the Treasury market. In particular, the approach allows for departures from linearity, normality, and homoscedasticity, and there are essentially no restrictions in modeling the dynamics of the Treasury-supply distribution itself.

I apply the approach to study and estimate a particular model, building on Greenwood and Vayanos (2014), in which mean-variance investors face risk from fluctuations in the short-term interest rate, the structure of Treasury debt, and inflation. The model includes two key nonlinearities: the variance of inflation is time-varying, and the short rate is bounded below by zero.² I include time-varying inflation risk because, as discussed in the survey of Gurkaynak and Wright (2012), previous studies have shown it to be an important driver of the nominal term premium. The potential importance of accounting for the zero lower bound (ZLB) on nominal interest rates, particularly when focusing on the LSAP period, has been emphasized in reduced-form term-structure models by Swanson and Williams (forthcoming), Priebsch (2013), and others.

Over the period 1971 – 2013, the model explains bond prices well, accounting for 90% of the variation in the yields on ten-year zero-coupon notes and 40% of the variation in their (hypothetical) excess returns. Consistent with other

²The Greenwood-Vayanos model is linear and does not include inflation. It is a special case of Vayanos and Vila (2009). In particular, in Vayanos-Vila, the supply distribution faced by arbitrageurs is endogenous, whereas in Greenwood-Vayanos this endogeneity is shut down and bond supply curves are vertical. Both models are members of the portfolio-balance class considered here.

empirical studies, it generates an estimate of the term premium that varies significantly both over the business cycle and at a lower frequency, rising by about 200 basis points from the early 1970s to the mid-1980s and then gradually receding. While the three sources of exogenous variation interact nonlinearly to compose the term premium, changes in inflation risk and changes in Treasury debt structure are of roughly equal importance over the full sample. In particular, an increase in Treasury duration is responsible for most of the early run-up in the term premium, and a moderation of inflation risk is responsible for most of its subsequent decline. Toward the end of the sample, some of the reduction in the term premium is also explained by lower short-rate risk as the short-rate distribution approaches the ZLB.

Despite the importance of Treasury duration for the term structure over the long run, the model implies that the Federal Reserve’s LSAP programs had only modest effects by reducing the interest-rate risk in the hands of investors—on the order of 20 basis points for the ten-year yield. Observers of LSAPs have frequently asserted that their effect on interest rates comes from “removing duration from the market.”³ It is precisely this effect that the portfolio-balance model formalizes and shows to be weak. The conclusion that duration removal may not be the primary mechanism involved in LSAPs is consistent with the recent event-study analyses of Krishnamurthy and Vissing-Jorgensen (2013) and Cahill et al. (2014). As those authors note, the effects of Treasury purchases might instead be explained through a scarcity channel (as in D’Amico and King, 2013) or through the signals that unconventional monetary policy sends about future short-term rates (as in Bauer and Rudebusch, 2014).

The primary reason that duration risk does not matter much in the case of LSAPs, even though it is important in general, has to do with the nonlinearity created by the ZLB. The truncation of the lower tail of the rate distribution at zero reduces volatility along the entire yield curve, and—since duration effects depend on the product of duration and volatility—render this channel less effective. (Doh, 2010, also makes this point.) The reduction in interest-rate volatility is even greater when the central bank broadcasts its medium-term intentions for short-term interest rates through so-called “forward guidance.” Thus, forward guidance and LSAPs, two tools of monetary policy that are often viewed as complements at the ZLB, actually offset each other somewhat in terms of their effects on longer-term yields. More broadly, the structure of government debt matters most for interest rates when interest-rate volatility is high. This observation should inspire caution if the Federal Reserve needs to sell assets at some point, since, to the extent that the path of policy is likely to be less certain in that environment, the duration effect could be asymmetrically large relative to the period of asset purchases. Similar considerations may also be worth policymakers’ attention when designing the pattern of Treasury debt issuance.

The paper proceeds as follows. Section 2 lays out the general properties

³For example, Gagnon et al. (2010) appeal strongly to this idea. Federal Reserve officials who have advocated a similar view include Sack (2009) and Bernanke (2010).

of portfolio-balance models and explains the conditions under which changing quantities can affect asset prices under no-arbitrage. It also provides some examples of such models in the literature. Generally speaking, those models have not been solved under rational expectations and previous authors in this literature have instead relied on shortcuts to try to estimate portfolio-balance effects. Section 2.3 sketches the solution algorithm I use to solve such models, with the details contained in an appendix. Section 3 shows how portfolio-balance applies specifically to the term structure of interest rates and illustrates how it results in "duration effects" in a calibrated one-factor model, both with and without the ZLB imposed. Section 4 extends this model to the more realistic case with inflation risk and stochastic supply and estimates the model on U.S. data. Section 5 applies the model to study the LSAPs, and Section 6 concludes.

2 No-Arbitrage Portfolio Balance

2.1 General Asset-Pricing Relations

I use the term “portfolio balance” to encompass a broad set of models in which the equilibrium prices of financial assets are related to the quantities of those assets that investors must hold and in which the quantities can be considered to have an exogenous or policy-dependent component. In particular, suppose there exists a finite number of assets N , with time- t prices $\mathbf{p}_t = (p_{1t}, \dots, p_{Nt})$, and par values $\mathbf{X}_t = (X_{1t}, \dots, X_{Nt})$. In equilibrium, investor wealth must be equal to the value of all assets: $W_t = \mathbf{X}_t' \mathbf{p}_t$. The idea behind portfolio balance is that fluctuations in the state of the economy that change the value of \mathbf{X}_t will change \mathbf{p}_t because expected returns—and therefore current prices—must adjust to make investors willing to hold the outstanding supply of securities at each point in time. In the most straightforward cases, the supply of each security is treated as an exogenous quantity issued by a price-insensitive entity, such as the government. More generally, \mathbf{X}_t may itself depend on prices.

The absence of equilibrium arbitrage opportunities implies the existence of a stochastic discount factor (SDF) $M_{t,t+s}$ that prices all assets in the economy. In particular, the price of any asset at time t is given by

$$p_{nt} = E_t [M_{t,t+s} q_{nt+s}] \quad \forall n \tag{1}$$

where q_{nt+s} is the asset’s payoff s periods hence and E_t indicates the expectation conditioned information at time t . This condition must hold for all horizons $s > 0$. It can be rewritten as

$$p_{nt} = \exp[-sr_t] E_t [q_{nt+s}] + \text{cov}_t [M_{t,t+s} q_{nt+s}] \tag{2}$$

where r_t is the time- t instantaneous risk-free rate of interest.

While $M_{t,t+s}$ could in principle depend on \mathbf{X}_t in a number of ways, I restrict attention to cases in which \mathbf{X}_t does not enter M_t directly. That is, investors

do not care about the quantities of the particular securities that they hold *per se*. This rules out, for example, models with convenience yields, monetary services, or other special benefits that might attach to certain assets beyond their pecuniary returns.⁴ Instead, I consider models in which $M_{t,t+s}$ can be written as a function of the return on wealth $R_{t,t+s}$ and, possibly, of the state of the economy, which is summarized by the vector \mathbf{s}_t :

$$M_{t,t+s} = M(R_{t,t+s}, \mathbf{s}_t, \mathbf{s}_{t+s}) \quad (3)$$

By definition,

$$R_{t,t+s} = \frac{\mathbf{X}'_t \mathbf{q}_{t+s}}{\mathbf{X}'_t \mathbf{p}_t} \quad (4)$$

The response of the price of asset n to an incremental change in the quantity outstanding of asset m is

$$\begin{aligned} \frac{\partial p_{nt}}{\partial X_{mt}} = & \mathbf{E}_t \left[\frac{\partial q_{nt+1}}{\partial X_{mt}} M_{t,t+s} \right] \\ & + \mathbf{E}_t \left[\frac{q_{nt+1}}{W_t} \frac{\partial M}{\partial R_{t,t+s}} \left(q_{mt+s} - R_{t,t+s} p_{mt} + \sum_{k=1}^N X_{kt} \left(\frac{\partial q_{kt+s}}{\partial X_{mt}} - R_{t+1} \frac{\partial p_{kt}}{\partial X_{mt}} \right) \right) \right] \end{aligned} \quad (5)$$

While it is not possible to sign this reaction without specifying how payoffs are determined and the additional properties of the pricing kernel, it is clear that the derivative in equation (5) will not be zero in general. Indeed, a sufficient condition for quantities to matter for prices is that the covariance of $M_{t,t+s}$ and q_{kt+s} is nonzero for some k . This is true even if we assume that asset payoffs do not depend on the current portfolio allocation \mathbf{X}_t . In that case, it is straightforward to show that $dp_{nt}/dX_{mt} = 0$ for all n if and only if $q_{mt+s}/p_{mt} = R_{t,t+s}$ —that is, unless the return on asset m is equal to the return on the market portfolio in all states of the world, changing the quantity of m will affect all prices.

Readers familiar with Eggertsson and Woodford (2003) may wonder how the above results relate to their widely cited "neutrality proposition," which states, in essence, that $\frac{\partial p_{nt}}{\partial X_{mt}} = 0$ in certain models so long as assets are valued only for pecuniary returns. The answer is that Eggertsson and Woodford's proposition is derived under the assumption that investor utility is a time-separable function of consumption. Consequently, $M_{t,t+s}$ depends only on consumption in periods t and $t+s$; it does not involve the market return $R_{t,t+s}$. Thus, condition (3) does not hold, and there are no portfolio-balance effects.⁵ Yet, while time-separable preferences are clearly an important case to consider, there are equally important models in which (3) *does* hold. The following subsection considers some of these models that have been applied by previous studies.

⁴Models with nonpecuniary returns are considered in Appendix A.

⁵Eggertsson and Woodford describe the logic of their result as follows. In their model, "the marginal utility to the representative household of additional income in a given state of the world depends on the household's consumption in that state, not on the aggregate payoff of its asset portfolio in that state. And changes in the composition of the securities in the hands of the public don't change the state-contingent consumption of the representative household. . . ." [Emphasis in original.]

2.2 Examples from the Literature

Although they are easy to write down, models of the above type are challenging to solve. Fundamentally, the reason is that prices in these models depend on the distribution of market returns, which is itself a function of future prices. The difficulties associated with this nonlinear recursion generally rule out analytical solutions under rational expectations. This is true even in special cases that restrict the form of M or the process for \mathbf{X}_t in convenient ways.

One simplification that does allow a model like this to be solved analytically is to suppose that the distribution of asset payoffs is fixed and that wealth is predetermined. In this case, equation (5) reduces to

$$\frac{\partial p_{nt}}{\partial X_{mt}} = \frac{1}{W_t} E_t \left[q_{nt+s} q_{mt+s} \frac{\partial M}{\partial R_{t,t+s}} \right] \quad (6)$$

Thus, for example, if the SDF is linear in $R_{t,t+s}$ with slope coefficient θ and the short rate is exogenous we have⁶

$$\Delta p_{nt} = \theta \text{cov}_t [q_{nt+s} q_{mt+s}] \frac{\Delta X_{mt}}{W_t} \quad (7)$$

Since θ is typically negative (states of the world with higher returns are discounted by more), raising the quantity of asset m lowers the price of asset n whenever the two assets have positively correlated payoffs.

In a classic application of this type of model, Frankel (1985) considered a case in which agents solve a Markowitz portfolio choice problem over a variety of asset types:

$$\max_{\mathbf{w}_t} E_t [\mathbf{R}'_{t+1} \mathbf{w}_t] - \frac{a}{2} \text{var}_t [\mathbf{R}'_{t+1} \mathbf{w}_t] \quad (8)$$

where $\mathbf{R}_{t+1} = (R_{1t,t+1} \dots R_{Nt,t+1})$ is the vector of asset returns, a is the coefficient of relative risk aversion, and \mathbf{w}_t is the vector of dollar values allocated to each asset. The maximization is subject to $\mathbf{w}'_t \mathbf{i} < W_t$. Mean-variance problems of this type result in an SDF that is linear in the market return. Consequently, one can write:

$$E_t [\mathbf{R}_{t+1}] = \exp[r_t] + a \boldsymbol{\Sigma}_t \mathbf{x}_t \quad (9)$$

where $\boldsymbol{\Sigma}_t$ is the covariance matrix of \mathbf{R}_{t+1} conditional on time- t information and $\mathbf{x}_t = (x_{1t} \dots x_{Nt})$ is the vector of asset shares. (I.e., $x_{nt} \equiv X_{nt} / \sum_{m=1}^N X_{mt}$.) Since (9) is just an equilibrium condition, it must be true for any value of \mathbf{x}_t . Thus, if a policymaker can set \mathbf{x}_t to an arbitrary value, he can determine the expected return on any given asset $E_t [R_{nt+1}]$. (In the case of bonds, this is sufficient to determine period- t prices.) Friedman (1986), Engle et al. (1995), and Reinhart and Sack (2000) are among the subsequent studies to exploit this

⁶The assumption that the short rate is exogenous implies $\partial E_t [M] / \partial R_{t,t+s} = 0$, which allows the product term in equation (6) to be reduced to the covariance term in equation (7).

relationship. More recently, Gromb and Vayanos (2010) illustrate the effects of quantities on prices in version of this problem in which there are only two assets, and Neeley (2010) and Joyce et al. (2011) apply it specifically to the study of LSAPs.

While this approach is useful for studying certain problems, its main difficulty from a logical perspective is that it treats the distribution of future asset prices—in particular the covariance matrix Σ_t —as exogenous. For example, when taking this model to the data, Frankel (1985) and subsequent authors calibrate Σ_t to the unconditional covariance matrix of the historical returns data. But the assumption of a fixed payoff distribution cannot be correct in general. For multi-period assets, payoffs tomorrow depend on prices tomorrow, and those should be determined in the same way as prices today. In terms of equation (5), there is no reason to expect $\frac{\partial q_{m,t+1}}{\partial X_{m,t}}$ to be equal to zero, particularly if there is persistence in the value of \mathbf{X}_t across periods. If prices depend on quantities, then the variance of prices must also, in general, depend on quantities. Yet, as noted above, it is not generally possible to characterize this dependence analytically.

Another important class of models in which portfolio-balance effects are potentially operative are those involving recursive preferences. In particular, consider a representative investor with Epstein-Zin-Weil utility over real consumption. As Epstein and Zin (1989) showed, the SDF in this case can be written as

$$M_{t,t+s} = \left(\beta G_{t,t+s}^{-\frac{1}{\psi}} \right)^\theta (R_{t,t+s} - \pi_{t,t+s})^{1-\theta} \quad (10)$$

where $G_{t,t+s}$ is consumption growth, $\pi_{t,t+s}$ is inflation, and θ and ψ are utility parameters. This SDF has the form of (3), and so the portfolio-balance channel applies. Piazzesi and Schneider (2008) study the relationship between bond quantities and prices in this type of model. To solve the model, Piazzesi and Schneider assume that investors have adaptive expectations over the asset-return distribution, which they estimate using survey data and VARs. As they argue, there may well be good reasons to believe that investors form expectations in this way. However, it is also of interest to consider the case in which expectations are set rationally. For a given vector of asset shares \mathbf{x}_t , it is evidently not possible to solve analytically for the state-contingent price vector that satisfies both equation (4) and equation (10) in all states of the world.

2.3 Solution Method

As illustrated in the cases above, it is not generally possible to solve portfolio-balance models analytically for asset prices as functions of quantities under rational expectations. The special cases in which it is possible (for example, exogenous payoffs or risk neutrality) are typically not of practical interest. Here, I propose a numerical method. This approach has the added advantage that it places very few constraints on either the functional form of the pricing kernel or the dynamics of the state vector. Consequently, it is straightforward to consider

models with potentially important nonlinearities, such as the zero lower bound. This section provides a thumbnail sketch of the solution algorithm; the details, along with convergence results, are discussed in Appendix B.

As suggested above, the central difficulty is that the solution for prices involves the moments of future prices, and, under rational expectations, these moments themselves must be endogenous. While it is common in asset-pricing models for today’s asset prices to depend on the distribution of tomorrow’s asset prices, the particular difficulty with portfolio-balance models is that the SDF itself depends upon equilibrium prices.

However, note that, since we take the law of motion for the states as known, it is possible to calculate these moments for any *given* mapping of \mathbf{s}_t to \mathbf{p}_t . This observation suggests the following algorithm:

1. Guess a function $\mathbf{p}^i(\cdot)$ such that $\mathbf{p}_t = \mathbf{p}^i(\mathbf{s}_t)$.
2. Based on this function and the known law of motion for \mathbf{s}_t , compute the distribution of $\mathbf{R}_{t,t+1}^w$, including its covariance with \mathbf{s}_t and \mathbf{p}_t .
3. Using that distribution, solve for the updated function $\mathbf{p}^{i+1}(\mathbf{s}_t)$ via equation (1) and return to step 2.

The moments in step 3 cannot generally be computed analytically, and so a quadrature scheme is used. The procedure is similar to that of Tauchen and Hussey (1991).

3 Portfolio Balance in the Term Structure

3.1 Preliminaries

For the vector of payoffs \mathbf{q}_t , suppose that we impose the restrictions $q_{1t} = 1$ and $q_{n-1t} = p_{nt-1}$ for $n > 0$. Evidently, the assets now represent zero-coupon, default-free bonds, where the index n indicates the bond’s maturity at time t . By definition, $p_{1t} = \exp[-r_t]$ is the price of the one-period bond, and equation (2) determines the rest of the yield curve. In particular, perfect certainty corresponds to the “strong form” of the expectations hypothesis—long-term interest rates are the average of expected short-term interest rates (up to a Jensen’s inequality term). Otherwise, the covariance term in equation (2) adds a risk premium to expected returns and a term premium to bond yields. The SDF M only appears in this covariance; thus, quantity changes operate by affecting term premia in the presence of uncertainty.

Restricting attention to bonds in this way, portfolio-balance effects take on a special meaning—bond quantities affect bond prices through a *duration channel*. That is, tilting the distribution of outstanding assets toward lower maturities alters the term structure by “removing duration from the market,” precisely the effect that is often pointed to, in less formal terms, as the primary mechanism through which LSAPs operate. (See, for example, Gagnon et al., 2011.) To

date, the only paper to capture such duration effects in a theoretical model that features both no-arbitrage restrictions and rational expectations is Vayanos and Vila (2009). In that model, arbitrage investors solve a problem like (8), but they interact with “preferred-habitat” agents who have downward-sloping demand curves for assets of particular maturities. Thus, the portfolio shares \mathbf{x}_t that they hold are endogenous. Assuming a linear-Gaussian factor structure, Vayanos and Vila obtain an affine solution in which demand shocks that remove long-term debt from the hands of arbitrageurs push down long-term interest rates.

As the only model to incorporate supply effects into an otherwise standard representation of the term structure, Vayanos-Vila has been highly influential in the way that economists have designed and interpreted recent empirical studies.⁷ However, the presence of the preferred-habitat agents and the endogeneity of \mathbf{x}_t in that model can obscure the fundamentally simpler relationship between prices and quantities. The example below strips this mechanism to its essentials to show how duration effects work. (Appendix A shows how Vayanos and Vila’s model fits into the broader portfolio-balance framework.)

3.2 Example: A linear, one-factor model

Consider a model in which the short rate r_t is the only source of stochastic variation and in which investors solve the mean-variance portfolio-choice problem (8), where the available assets consist only of Treasury bonds with maturities $1, \dots, N$. The relative supply of bonds is fixed over time, $\mathbf{x}_t = \mathbf{x}$. Bond prices then solve the system of quadratic equations

$$\mathbf{p}_t = \exp[-r_t] \left(E_t[\mathbf{q}_{t+1}] - a \frac{\boldsymbol{\Omega}_t \mathbf{x}}{\mathbf{p}_t' \mathbf{x}} \right) \quad (11)$$

at each point in time, where $\boldsymbol{\Omega}_t$ is the conditional covariance matrix of the asset payoffs \mathbf{q}_{t+1} . (This is the equivalent of equation (2) for the Markowitz model.) Since the payoff on a one-period bond is always 1, the first row and column of $\boldsymbol{\Omega}_t$ are zeros. The conditional variance of the price of a one period bond, denoted ω_{1t}^2 , is the second diagonal element of $\boldsymbol{\Omega}_t$. Since $p_{1t+1} = -\log r_{t+1}$, this variance is a deterministic function of the short-rate process and is a known and exogenous quantity.

It is straightforward to show that the expectations hypothesis holds if $\mathbf{x} = (1 \ 0 \ \dots \ 0)$. That is, eliminating all duration from investors’ portfolios reduces the term premium to zero. On the other hand, as long as $a > 0$ and $x_1 < 1$, we have $p_{nt} < p_{1t} E_t[p_{nt+1}]$, for $n > 1$, so that all risky assets require a discount and all risky returns include a risk premium. Indeed, since r_t is the only source of variation, the term premium in this example reflects *only* the compensation for bearing duration risk. Specifically, there is a single market

⁷See, for example, Doh (2010), Hamilton and Wu (2011), and Li and Wei (2012).

price of risk given by the Sharpe ratio

$$\lambda_t = \frac{E_t[R_{nt,t+1}] - R_{1t,t+1}}{\sigma_{nt}} \quad (12)$$

where σ_{nt} is the standard deviation of the return on asset n . This, in turn, implies that all asset prices and returns are perfectly correlated, regardless of the value of \mathbf{x} or the dynamics of r_t . In particular, there exist coefficients A_{nt} and B_{nt} such that we can write $R_{nt,t+1} = A_{nt} + B_{nt}p_{1t+1}$ for any n .⁸ The coefficient B_{nt} represents the response of the return on a risky asset to the risk-free rate—the analogue of the “spot-rate duration” that appears in continuous-time term-structure models. The standard deviation of the return on the market portfolio can be expressed as $\sigma_t^w = \omega_{1t}\mathbf{x}'\mathbf{B}_t$, where \mathbf{B}_t is the vector of B_{nt} coefficients. Combining this with equation (12), we have

$$\lambda_t = a\omega_{1t}\mathbf{x}'\mathbf{B}_t \quad (13)$$

analogous to a result produced by Vayanos and Vila (2009). Reducing the quantity $\mathbf{x}'\mathbf{B}_t$ is the model’s version of “removing duration from the market.” Doing so will lower both the total risk of the Treasury portfolio and the price of that risk.⁹

To implement the model numerically, take periods to be one year in length and suppose that there are $N = 30$ bonds. Suppose that the short rate follows a linear process

$$r_t = \phi_0 + \phi_1 r_{t-1} + \varepsilon_t \quad (14)$$

where ε_t has variance σ^2 . For parsimony, I approximate the maturity structure of government debt with an exponential distribution:

$$x \propto \exp[-z] \quad (15)$$

where the parameter z is the average maturity outstanding. The exponential approximation is likely to be a good one because the distribution of government liabilities typically consists overwhelmingly of shorter maturities. For example, as shown in Figure 1, liabilities with less than five years of duration (including currency and reserves) account, on average, for about 80 percent of total face value. The exponential functional form implies that, as in the data, the distribution is highly skewed toward securities with very low duration. However, as discussed further below, the exact shape of the distribution beyond its first moment is actually of little importance.

⁸Specifically, $B_{nt} \equiv \sigma_{nt}/\omega_{1t}$ and $A_{nt} \equiv (1 - p_{2t}B_{nt})/p_{1t}$. Because r_t follows an AR(1) process, the function B_{nt} has a similar form across n at each value of the short rate as in the Vasicek (1977) model.

⁹Since (11) is the result of a Markowitz portfolio problem it also has a conditional CAPM representation. Specifically, since $E_t[R_{t,t+1}^w] = E_t[\mathbf{q}_{t+1}']\mathbf{x}/\mathbf{p}_t'\mathbf{x}$ is the expected return on wealth, we can write

$$E_t[\mathbf{R}_{t,t+1}] - r_t\mathbf{i} = \beta_t (E_t[R_{t,t+1}^w] - R_{1t,t+1})$$

where $\beta_t = \mathbf{\Omega}_t\mathbf{x}/\mathbf{x}'\mathbf{\Omega}_t\mathbf{x}$. The expected excess return on the market portfolio is $E_t[\mathbf{R}_{t+1}] - r_t = a\mathbf{x}'\mathbf{\Omega}_t\mathbf{x}$.

Let $\phi_0 = 0.003$, $\phi_1 = 0.95$, $\sigma = 0.015$, and $a = 8$. These values are calibrated roughly to match the dynamics of the short-rate and the average value of 15-year yield in the data. The black line in Figure 2 demonstrates how the duration of the Treasury portfolio affects the behavior of asset prices in the model by plotting the equilibrium values of the (inverse) price of risk λ_t across possible values of z in this model. The relationship between duration and risk prices is monotonic but nonlinear. Near the historical mean duration of 2.7 years (see the top panel of Figure 1), increasing duration by one year raises the price of risk by about 50%.

Figure 3 illustrates how these risk prices translate into yields. I compare two cases: $z = 2.7$ years and $z = 2.0$ years. As discussed in Section 5, the difference of -0.7 years is approximately equivalent to the effects of the cumulative asset purchases by the Fed through 2012 (the first and second round of LSAPs and the Maturity Extension Program). The difference in the maturity distribution of securities is illustrated by the red bars in the top panel of the figure. The effect on the yield curve, evaluated at the sample average value of the short rate, is shown in the middle panel. For the calibration used here, the ten-year yield is 56 basis points lower under the smaller-duration distribution. Again, this is due to reductions both in the price of risk illustrated in the previous figure and in risk itself. The latter is shown in the bottom panel of Figure 3, which plots the standard deviation of one-year returns (the square-root of the diagonal of Σ_t) under the two supply distributions.

Finally, we can demonstrate the earlier claim that the shape of the distribution \mathbf{x} is of little importance for the shape of the yield curve. Analytically, equations (11) and (13) show that the individual asset share x_n does not matter for the individual asset price p_{nt} . Only the weighted sum of asset shares $\mathbf{x}'\mathbf{B}_t$ is relevant, and it affects all prices in the same way. Consequently, this model cannot produce local-supply effects from large quantity gluts or shortages in particular sectors of the market. To illustrate this, consider a stark alternative to the exponential shape used for the maturity distribution above. In particular, suppose that the distribution was nearly degenerate around the same mean of $z = 2.7$, as shown in the top panel of Figure 4. This adjustment does slightly increase the duration risk of the portfolio, primarily because it completely eliminates the risk-free one-year bonds from investors' portfolios. However, after recalibrating a to match the average slope of the previous yield curve, there is virtually no difference between the term structures generated under the two distributions. This can be seen in the bottom panel of the figure, where the solid blue line is the same yield curve shown in Figure 3, and the dashed red line is the (recalibrated) yield curve at the same value of the short rate under the degenerate supply distribution. Similar results obtain for other possible choices of the shape of x .

3.3 Effect of the Zero Lower Bound

Since risk prices in portfolio-balance models are themselves a function of the quantity of risk held by investors, heteroskedasticity can cause risk prices—and

therefore term premiums—to differ significantly across states of the world. A particularly important case of this is the zero lower bound on the nominal short rate. The presence of this bound, all else equal, implies that there is less uncertainty about short-term interest rates in the near future when the current value of those rates is near zero. Consequently, investors in all bond maturities face less near-term interest-rate risk and therefore demand lower risk premiums. Furthermore, since equation (13) shows that the effect on risk prices of removing duration from the market (in the sense of reducing the weighted average $\mathbf{x}'\mathbf{B}_t$) is directly proportional to the variance of the one-period bond ω_{1t} . Thus, the reduction in the volatility of short-term interest rates induced by the ZLB will also dampen duration effects.

To illustrate these outcomes in the simple one-factor model above, suppose that, rather than allowing the short rate to follow an unrestricted AR process, we append the condition

$$\Pr[\varepsilon_t < 0] = 0 \tag{16}$$

to equation (14). That is, let the distribution of the short-rate shock ε_t be a truncated normal, where the truncation point varies with r_t in such a way that it always enforces the zero lower bound. The conditional distribution of the short rate under this process is illustrated in Figure 5. The reduction in volatility near zero is evident.

Referring back to Figure 2, the blue line illustrates risk prices calculated at the ZLB for different values of duration; on average, they are roughly half of the unrestricted case considered previously. Furthermore, the slope of the blue line is considerably flatter than that of the black line, indicating that a given change in duration has a smaller effect on λ_t when the ZLB binds. To see these results in terms of the yield curve, the solid lines in the top panel Figure 6 correspond to the same maturity distributions that were used in Figure 3. (2.7 years for the blue line and 2.0 years for the red.) The same LSAP-type experiment that caused a 56 basis point reduction when the short rate was at 5.8% causes only a 29 basis point reduction when the short rate is at 0% with a binding ZLB.

The reduction in duration effects becomes even more pronounced if the short rate is anticipated to stay near zero with a high degree of certainty for multiple periods, because this results in a larger reduction in the variance of rates across the term structure. This type of thought experiment is relevant because the forward-guidance language that has been included in most FOMC statements since 2008 is widely viewed as having essentially committed the Fed to keeping rates near zero for a significant time into the future. To illustrate, the dashed lines in the figure show the model's outcomes when r_t is constrained to equal 0 for both periods t and $t+1$ (and investors know this with certainty at time t) and only then follows its truncated AR(1) process. The forward guidance itself has a large effect on long-term rates; in this example, it reduces the 10-year yield by 52 basis points (the difference between the solid and dashed blue lines). However, with the forward guidance in place, the effect of the reduction in duration on the ten-year yield is only 23 basis points, rather than 29. The reason is that, as illustrated in the bottom panel, forward guidance in the one-

factor model implies that the returns on all bonds in period $t+1$ are known with certainty. Consequently, the risk premium for the period over which the forward guidance is in place is zero, regardless of the amount of duration outstanding—i.e., ω_{1t} is zero, and so, by equation (13), the price of risk λ_t is zero. Thus, although they both serve to reduce longer-term yields, forward guidance and asset purchases are somewhat offsetting, at least if duration effects are the only channel through which the supply distribution matters for prices.

4 A Three-Factor Model

While the one-factor model is useful for understanding the mechanics of portfolio-balance effects, it lacks certain realistic features that would make it quantitatively credible. This section develops a somewhat more-sophisticated model of portfolio-balance effects in the yield curve and brings that model to the data. The aspects in which the model is more realistic than the one-factor case above are that (1) duration itself is treated as a separate stochastic factor, (2) inflation and the variance of inflation are also allowed to vary stochastically, and (3) investors are assumed to value real, rather than nominal, returns. In addition, I assume that the ZLB is in effect. As noted in the introduction, this model can be viewed as an extension of Greenwood and Vayanos (2014) that incorporates the nonlinearities associated with the ZLB and inflation. The importance of accounting for the ZLB, particularly when studying the LSAP experience, was suggested by the results of the previous section. The addition of inflation to the model is potentially important because inflation—or, more precisely, time-varying inflation *risk*—has been shown to be an important determinant of nominal term premiums. (See the literature surveyed by Gurkaynak and Wright, 2012.)¹⁰

4.1 Model Setup

Investors solve the problem

$$\max_{\mathbf{w}_t} \mathbb{E}_t [\mathbf{R}'_{t,t+1} \mathbf{w}_t - \pi_{t,t+1}] - \frac{a}{2} \text{var}_t [\mathbf{R}'_{t,t+1} \mathbf{w}_t - \pi_{t,t+1}] \quad (17)$$

where π_{t+1} represents the gross rate of inflation between periods t and $t+1$. It is straightforward to show that prices solve

$$\mathbf{p}_t = \exp[-r_t] \left(\mathbb{E}_t [\mathbf{q}_{t+1}] - a \frac{\boldsymbol{\Omega}_t \mathbf{x}_t}{\mathbf{p}_t' \mathbf{x}_t} - \text{cov}_t [\mathbf{q}_{t+1}, \pi_{t+1}] \right) \quad (18)$$

¹⁰There are three other, minor differences between this model and Greenwood-Vayanos. First, the Greenwood-Vayanos model is in continuous time. Second, I use the exponential form in (15) for the supply distribution, whereas they use a linear-factor structure. Finally, they assume that market-value weights w_{nt} are exogenous, while I assume that the par values x_{nt} are exogenous. None of these differences in approach is likely to have substantial effects on the model outcomes.

and that the nominal pricing kernel takes the form

$$M_{t,t+1} = \theta_{0t} + \theta_{1t} (R_{t,t+1}^w - \pi_{t,t+1})$$

Note that the portfolio weights \mathbf{x}_t now have time subscripts, as investors will anticipate them to vary stochastically from period to period. Nonetheless, the pricing equation (18) must hold at each point in time. Also note that changes in \mathbf{x}_t have no direct effect on inflation-risk premia through portfolio rebalancing. While they may change the covariance of prices with inflation through their effect on the overall volatility of asset prices, these effects will generally be second-order.

4.2 Estimation

I assume that inflation and the short rate are jointly determined by the process

$$r_t = \phi_0 + \phi_1 r_{t-1} + \phi_2 \pi_{t,t+1} + \varepsilon_t^r \quad (19)$$

$$\pi_{t,t+1} = \gamma_0 + \gamma_1 \pi_{t-1,t} + \varepsilon_t^\pi \quad (20)$$

subject to

$$\Pr [r_t < 0] = 0 \quad (21)$$

where ε_t^r is a truncated-normal disturbance with shape parameter σ_r , and ε_t^π is distributed normally with time-varying variance:

$$\sigma_{\pi t}^2 = \rho_0 + \rho_1 |\pi_{t-1,t}| \quad (22)$$

where ρ_0 and ρ_1 are both non-negative. This system produces behavior of the short rate similar to that depicted in Figure 5, except that now the configuration of the curves in that figure depends on the level of inflation. The dependence of the short rate on contemporaneous inflation, together with the heteroscedasticity of inflation shocks, provides a simple way of capturing time-varying covariance between r_t and $\pi_{t,t+1}$.

z_t is assumed to follow an AR(1) process, independent of r_t and $\pi_{t,t+1}$. This process is estimated using the data on aggregate duration of government liabilities (see Figure 1), giving

$$z_t = 0.11 + 0.97z_{t-1} + \varepsilon_t^z$$

with the standard deviation of ε_t^z equal to 0.16. Note that, although z_t has the interpretation of the average duration of debt outstanding, the entire distribution x_t is still the input in the model, and the shape of this distribution at each point in time depends on z_t through equation (15).

Estimation of the remaining parameters proceeds by fitting the measurement equations

$$y_t^n = \hat{y}_t^n + e_t^n \quad (23)$$

where y_t^n is the n -period yield in the data, \hat{y}_t^n is the corresponding yield generated by the model, and e_t^n is an iid normally distributed measurement error, for $n = \{5, 10, 15\}$. Estimation of this model without further restrictions indicated that it was not well identified from the data on longer-term Treasury yields alone. This is perhaps unsurprising, given that yields are highly serially and cross-sectionally correlated and that several of the model parameters (e.g., risk and risk aversion) are likely to have very similar effects. I therefore take two measures to aid identification. First, as is common in the term-structure literature, I restrict the covariance matrix of the measurement errors to be diagonal. (This is equivalent to minimizing the weighted mean-squared errors over the three long-term yields.) Second, I bring in information from the observable dynamics of the short rate and inflation themselves, rather than just the information contained in the Treasury bond yields. Because the data are observed monthly while the periods in the model are taken to be annual, joint estimation is not straightforward, but the mixed frequencies are easily dealt with using Bayesian methods. Specifically, I first estimate the system (19) - (22) by maximum likelihood on non-overlapping annual samples, using only the one-year Treasury rate and core PCE inflation. The results are reported in the top line of Table 1. I then use these estimates to form a prior distribution which I combine with the likelihood generated by the measurement equations (23). I put weak priors on the risk-aversion parameter a , which is not identified from the first-stage estimation. All of the parameters were then estimated using 50,000 Metropolis-Hastings draws from the posterior distribution, with a burn-in sample of 10,000 discarded. Because the forward guidance issued by the Federal Reserve after the short rate hit the ZLB affected expectations of the path of rates in an uncertain way, I exclude this period from the sample and use only the data from 1971 - 2008.

4.3 Results

4.3.1 Model fit

The parameter estimates (means and standard deviations of the posterior distributions) over the 1971 - 2008 period are shown in the second row of Table 1. Since $\phi_2 > 0$, inflation is correlated with the short rate. In addition, the variance of inflation varies significantly with its level ($\rho_1 > 0$). This implies a time-varying inflation risk for all bonds. However, since inflation itself is not highly persistent ($\gamma_1 = 0.71$), this risk is greatest for short- and medium-term securities, adding curvature to the yield curve.

Table 2 reports various statistics summarizing where the model succeeds and fails in fitting the data, when evaluated at the posterior mode. I assess the fit over spot yields (which were used in the estimation), forward rates, and excess returns. The one-year forward rate ending n years ahead is calculated from bond prices, in both the model and the data, as $f_n = \log(p_{n-1t}/p_{nt})$. Excess returns are computed annually as $exret_{nt} = \log(p_{n-1t}/p_{nt-1}) - R_{1t-1,t}$. (To avoid overlapping data, excess returns are only calculated using the December

data in each year.) I report statistics for 5-, 10-, and 15-year bonds over the entire sample and for 30-year bonds over the post-1985 sample when these data are available. Since the 30-year bonds were not used in the estimation, they provide a sort of out-of-sample check on the results.

Given the tight structural restrictions of the model and the relatively parsimonious parameterization, we would not expect it to fit the data as well as, say, a reduced-form multifactor affine model. Nonetheless, it generally does well, coming close to matching most of the first and second moments of yields within-sample. It predicts a bit too little volatility and a bit too much correlation of longer-term yields with the short rate. However, the model explains about 90% of the variation in yields of up to 15 years, and it even explains 73% of the variation in the out-of-sample 30-year yield. Perhaps most remarkably, the model generally matches the features of excess bond returns through the 15-year maturity, with R^2 s of 0.33 to 0.63. Although it misses substantially on the overall features of the 30-year bond returns, it still manages to explain 14% of their variation as well.

4.3.2 Impulse-response functions

Figure 7 examines the effects of shocks in the model. I consider one-standard-deviation shocks in a direction that leads to positive yield responses. Since the model is nonlinear, initial conditions matter, and I consider responses both at the sample mean (left column) and at a set of values that reflect conditions during the LSAP period (right column); specifically, the latter is characterized by $r_0 = 0.003$, $\pi_0 = 0.014$, and $z_0 = 2.8$.

Roughly speaking, short-rate shocks and inflation shocks have similar effects on long-term yields. Short-rate shocks have larger initial impact, but they die out rather quickly. The inflation shocks have more-persistent effects on short rates and also increase inflation uncertainty, both of which serve to amplify their impact on the long end of the curve. The dynamics of the system impart a hump shape to the short-rate response to inflation shocks over time, and this is reflected in the shape of the yield curve reaction to such shocks across maturities. There is not much qualitative difference between the responses at the steady state and during the LSAP period.

A shock to average duration has almost the opposite effect across maturities as a shock to the other variables: short rates do not move at all, and 30-year yields move by 15 to 20 basis points. The initial response is qualitatively similar to the comparative-statics exercises depicted in Figures 3 and 6. In the dynamic model here, these responses decay over time, as the shock to z_t dies out. The responses to duration shocks are about one-third smaller near the ZLB than at the sample mean, again reflecting the lower volatility of interest rates in that state of the world. Having stochastic supply in the model adds a source of variation to longer-term interest rates and thus, all else equal, increases their volatility. Nonetheless, as the figures suggest, the effect of the typical duration shock on, say, the 10-year yield is only about one-tenth of the effect of the typical short-rate shock. Thus, at least in this model, duration shocks account

for only a modest fraction of the overall variation in longer-term yields.

4.3.3 Term premium decomposition

The top panel of Figure 8 shows how the model decomposes the actual time series of ten-year yields into an expectations component, a term premium, and a model-error term. (Since only the period 1971 – 2008 was used in the estimation, the last four years in the figure are out-of-sample estimates.) The expectations component is calculated as the value of the ten-year yield that would be implied by the model at each point in time if a were equal to zero. The term premium is then just the difference between the model-implied yield on each date and the expectations component. As noted above, the model generally tracks the behavior of longer-term interest rates well over this time. The most notable exceptions are the early 1980s, during the Volcker disinflation when the market-perceived inflation risk may have been higher than the level captured by the model, and the period from about 2005 – 2007, corresponding to the so-called “conundrum” in long-term interest rates (see Gurkaynak and Wright, 2012). The latter is a well known anomaly in the yield curve that is not likely to be reproduced by the factors considered explicitly here. However, one possible explanation is that a large quantity of foreign investment increased the demand for longer-term Treasuries—the “global savings glut” introduced by Chairman Bernanke. Conceivably, this type of mechanism could be incorporated into the portfolio-balance model, perhaps by adding exogenous fluctuations in demand from preferred-habitat agents, similar to the model considered in Appendix A.

The second panel shows the term premium. In broad strokes, it follows a trajectory that is similar to the term premium produced by reduced-form models, such as Kim and Wright (2005), rising through the 1970s and early 1980s, and falling since the late 1980s. (Rudebusch et al., 2007, estimate the Kim-Wright term premium for a sample that covers this entire period.) Because of nonlinearities, the decomposition of the term premium into its structural contributing factors is not uniquely defined. However, to get a sense of the relative contributions, I calculate counterfactual yields holding z_t and π_t equal to their sample means in all periods and under the assumption that investors know that they will always be equal to these values. The latter assumption serves to eliminate the contribution of the uncertainty about each variable to the term premium. A non-trivial fraction of the movement in the term premium since the late 1990s is attributed to the reduced interest rate risk associated with the short rate approaching the ZLB. This can be seen by the dashed line in the middle panel, which isolates the contribution of the short rate to the term premium by setting both duration and inflation to their sample averages. It is also clear from this figure that these two factors have jointly contributed to significant variation in the term premium relative to what would have been implied by variation in the short rate alone.

The bottom panel shows the contributions of inflation and Treasury supply to the term premium, measured by subtracting the counterfactual yield in which both series are held constant from the counterfactual yield in which only one or

the other is held constant. Overall, fluctuations in inflation are estimated to have moved term premiums within a range of about 100 basis points over the course of the sample. Consistent with previous research (e.g., Campbell et al., 2009), these fluctuations have mostly been in the form of a downward trajectory since the early 1980s. On the other hand, the Treasury supply distribution shifted significantly to longer maturities during the period from about 1975 to 1985, and the model suggests that this resulted in an upward movement in term premiums of about 150 basis points. Since the 1980s, supply-related movements have been relatively small because shifts in Treasury supply itself have been small. Broadly speaking, the results are consistent with the finding of Li and Wei (2012) that measures of Treasury supply do affect term premiums over time. The net result of all of these effects, returning to the middle panel, has been an increase in the overall term premium from just over 100 basis points at the beginning of the sample to a peak of almost 300 basis points around 1990, followed by a decline to near its original levels by the end of the sample. The early run-up was due primarily to an increase in duration risk, while the subsequent decline was due to a decrease in inflation risk, together with the approach of the ZLB.

Table 3 summarizes these findings by comparing the monthly change in the model-implied ten-year yield and term premium to the changes that would have occurred in the counterfactual scenarios. By this measure, inflation accounts for 39% of the overall variance in term premiums and fluctuations in Treasury supply account for 70%. The two factors are correlated in the sample, so excluding them both reduces the volatility of the term premium by 94%. Although these fractions are large, it is important to bear in mind that most interest-rate fluctuations in the model are not due to term premiums but rather to expected short rates. Indeed, as shown in the top row of the Table, holding both inflation and average duration at their sample averages reduces the variance of 10-year yield changes by just 6%.

5 Assessment of the LSAP programs

In this section, I consider the Federal Reserve’s LSAP programs using the three-factor model presented above. The relevant aspects of the programs are summarized in Table 4. A large component of the Federal Reserve’s purchases consisted of agency mortgage-backed securities. Throughout the paper I have ignored agency mortgage-backed securities when calculating the duration of government liabilities, but this exclusion is likely inappropriate because these securities have typically been perceived to carry an implicit government guarantee. Two technical measurement problems arise when trying to extend the analysis to agency MBS. The first is that comprehensive data on the maturity structure of outstanding MBS are not available. The second is that, unlike the case of zero-coupon Treasuries, duration for MBS depends on economic conditions—in particular, it depends on the level of interest rates through the negative con-

vexity induced by the prepayment option. However, according to the Barclays MBS index, the average duration of MBS since 1989 is about three years, only slightly lower than that of Treasury debt. (See Hanson, 2012.) In performing the calculations below, I continue to use the approximation of an exponential distribution for the duration distribution. Moreover, I assume that asset purchases did not materially change the average duration of MBS outstanding, but I do incorporate the amount of MBS outstanding, so that Fed purchases swap their duration for zero-duration reserves.

The net effect of the LSAP programs, as of December 2012, was to reduce the outstanding supply of Treasury and MBS securities by \$2 trillion and increase reserves by a similar amount. Moreover, the Treasury securities removed from the market were, after the duration twist induced by the Maturity Extension Program (MEP), almost all of maturity greater than five years. Consequently, the average duration of Treasury securities in the hands of the public (not counting reserves) was about 0.5 years lower after the completion of the programs than it otherwise would have been. Taking all of these facts together, we can estimate how investors' duration changed as a result of the programs. Specifically, as of Q4 2008, Treasury debt held by the U.S. public was \$5.3 trillion, agency MBS was \$5.0 trillion, and the monetary base was \$1.4 trillion. The average duration of Treasuries (excluding the monetary base) in the hands of the public was 3.1 years as of the beginning of the program. Therefore, assuming that MBS have an average duration of three years (and again counting the base as zero duration), the weighted average of Treasuries, MBS, and the monetary base in the hands of the public was 2.7 years when the first LSAP was announced. Holding all else constant, this value would have fallen to 2.0 years as a result of the programs.

The final column of the table presents the range of values for the empirical effects of the LSAPs on the ten-year term premium, culled from the literature that has examined this question. In particular, this range draws on estimates from Gagnon et al. (2010), Krishnamurthy and Vissing-Jorgensen (2011), Ihrig et al (2012), D'Amico et al. (2012), and Rosa (2013). (In some cases, additional minor calculations were required to make the results comparable.) Of course, all of these estimates are subject to a high degree of uncertainty surrounding both parameter values and specification. Furthermore, since most of these papers employ an event-study methodology using program-announcement dates, the estimates may be understated. Nonetheless, they provide a rough guide to what we should expect for the combined effects of the programs—likely on the order of 100 to 250 basis points altogether. To be clear, this is (according to the authors of the studies) only the effect on the *term premium*, controlling for changes in expectations of the short rate, and it only reflects the initial impact of the programs, ignoring any subsequent dynamic effects.

How do these estimates compare to the results of the portfolio-balance model? The black line in Figure 9 shows the effect of removing 0.7 years of duration from the hands of investors, starting from a level of duration that is equal to what was observed prior to the LSAP's introduction (2.7 years) and a configuration of short rates and inflation that approximates the situation on

average since that time. Specifically, inflation is set at a level of 1.4 percent (its average over the period 2009 – 2012), and the short-term interest rate is taken to be near the zero lower bound with two years of forward guidance in place. Although the market’s perception of the length of time that the ZLB would bind surely fluctuated over this period, the assumption of two years is likely close to the average that prevailed, perhaps erring on the conservative side (see Femia et al., 2013). As a reality check, the first row of Table 5 shows that the slope of the yield curve produced by the model under these assumptions over the LSAP period was reasonably close to the average slope of the yield curve in the data during this time. (Recall that this part of the sample was not used in the estimation, and that the model was estimated under the assumption of no forward guidance.) The effect of the LSAPs on the ten-year yield of the LSAP shock in this environment is a mere 18 basis points. This is the initial impact—given the estimated dynamics of z_t , the effect would decay to zero over time, similarly to the response shown in Figure 7.

The other lines in the figure check the robustness of this result to the most likely sources of measurement error. First, although the two-year period assumed for the forward guidance may be a reasonable approximation in terms of the time interval involved, the assumption that the Fed committed to the level of the short rate with absolute certainty for this entire time could be too strong, given that the forward guidance issued by the FOMC has always left open the possibility that rates could rise sooner than expected if economic conditions warranted. Indeed, empirical measures of short-term uncertainty about the short rate have consistently remained significantly above zero while forward guidance has been in place. The green line examines the sensitivity in this dimension by instead imposing that investors believe that the ZLB will bind with certainty for only one year following the LSAP shock. This has the effect of raising the standard deviation of the one-year-ahead short rate from zero to 1.2%, and it brings the 10-year slope into even closer alignment with the data. However, it increases the effect of the LSAPs only to 24 basis points. Second, given that the calculation of the size of the 0.7-year duration shock was somewhat back-of-the-envelope, one might worry that it is less than the true amount of duration removed by the LSAPs, perhaps understating the duration outstanding at the beginning of the programs and overstating the amount left at the end. The orange line shows the effect of a shock that is twice as large as the baseline, going from an average duration of 3.05 years to 1.65 years; the effect on the ten-year yield is about double the previous estimate, but it is still only about a third of the low end of the empirical range of LSAP effects.

One might also object that, during much of the LSAP period, the average level of risk aversion was higher than usual. The second row of Table 5 shows the effect of doubling the estimated risk-aversion coefficient to a value of 25.4. The impact of the LSAP shock in this case is still just 29 basis points under the assumption of two-year forward guidance. This value rises to 60 basis points—perhaps half of the empirical value—if one assumes shorter forward guidance in addition to higher risk aversion, but that scenario does not seem likely to be apposite, as it implies a counterfactually steep slope for the yield curve (420 bp,

rather than 268 bp, on average).

6 Conclusion

This paper has presented a new method for studying the term structure of interest rates and, in particular, the ways in which the term structure is affected by supply fluctuations. The type of model considered is a rational-expectations version of portfolio-balance models that have been in use, with varying degrees of formality, for decades. It may be viewed as a generalization of the preferred-habitat model of Vayanos and Vila (2009), with investors that potentially have a broader class of objective functions and in which nonlinearities may be important. It is hoped that this approach may be of use in studying other asset-pricing phenomena as well.

One feature of the particular set of models to which I have applied the approach is that they likely overstate the effect that supply fluctuations could have through a duration channel because they assume that the investor portfolios consist only of Treasury bonds and that they cannot adjust their consumption or other behavior to offset the utility effects of losses of wealth. Pursuing these extensions is left for future research. The reasonably strong fit of the model considered here suggests that the extent to which it exaggerates portfolio-balance effects may not be large, although it does call for some caution when thinking about results like the term-premium decomposition. However, the upper-bound nature of the models can be viewed as an advantage when trying to gauge the likely effects of asset-purchase programs, since even under their potentially strong assumptions the models suggest that those effects are fairly small. I interpret these results, not as evidence that the LSAPs were ineffective, but as evidence that they probably had their effects primarily through mechanisms other than the removal of duration risk. It seems possible that phenomena not easily captured by a no-arbitrage model, such as market dislocations, liquidity shocks, and capital constraints, were important during and after the financial crisis and that LSAPs had additional effects through those channels. This reading is broadly consistent with the empirical conclusions of Krishnamurthy and Vissing-Jorgensen (2011), Cahill et al. (2013), and D'Amico and King (2013). More structural modeling of the behavior driving such scarcity effects is needed to determine whether they would significantly alter the results presented above.

Appendices

A. Nonpecuniary Returns

Imperfect substitutability

In this appendix, I consider ways of introducing imperfect substitutability among assets by allowing them to have "convenience yields." Such a formulation could be motivated, for example, by a planning horizon longer than one period, by a desire to match long-duration liabilities, or by supposing that short-term assets provide liquidity services. (Cox et al., 1981, Krishnamurthy and Vissing-Jorgenson, 2012, and Greenwood et al., 2013, provide models in which this can occur in different forms.) This is potentially important because these phenomena can create demand for assets in excess of what would be implied by their risk and return characteristics under no-arbitrage. Consequently they can interact with the portfolio-balance effects developed in the main text in interesting ways. I illustrate these effects in the one-factor version of the model, in which the supply distribution is static, although it is straightforward to include them in multifactor models as well.

To begin, re-define the return on asset n as

$$R_{nt,t+1} = \frac{q_{nt+1} + b_n(\mathbf{s}_t, \mathbf{p}_t)}{p_{nt}}$$

where $b_n(\cdot)$ is a security-specific benefit that may depend on the state of the economy and on the entire vector of security prices. Though a variety of possibilities exist for b_n , I consider two that seem particularly relevant. First, it may be that investors receive a benefit from holding securities of particular maturities. In particular, suppose that the return on an individual security is given not just by its price appreciation but also by an unobserved benefit that depends on its maturity. For parsimony, I assume that this dependence is linear in the end-of-period maturity, $n-1$:

$$b_n^{\text{maturity}} = b(n - 1)$$

for some parameter b , which could take either sign depending on whether benefits are greater for short- or long-term securities. Second, as suggested by Hanson and Stein (2012), investors may have preferences over yields, in addition to preferences over returns. (They point to this mechanism, reminiscent of "reach for yield"-type behavior, as a potential explanation for the excess volatility of longer-term yields over the business cycle.) To incorporate this idea, I allow investors to receive benefits on bonds based on their current yields, in excess of the short-term yield:

$$b_n^{\text{yield}} = c \left(p_{nt}^{-1/n} - \exp[-r_t] \right)$$

where $c > 0$ is a parameter reflecting the additional benefit generated by the current yield.

I combine the yield and maturity preference specifications into the single convenience-yield function $b_n = p_{nt}(b_n^{\text{maturity}} + b_n^{\text{yield}})$, where the multiplication by the price of bond n imposes that the benefit is proportional to the market value, rather than the face value, of the security held in the portfolio at time t . This gives the following specification for returns:

$$R_{nt+1} = \frac{p_{n-1t+1}}{p_{nt}} + b(n-1) + c \left(p_{nt}^{-1/n} - \exp[-r_t] \right)$$

For illustration in the one-factor mean-variance model, I calibrate the three parameters a , b , and c to match the average 15-year slope of the yield curve, the average curvature of the yield curve (as measured by the difference between the 1-15 and 15-30-year slopes over the more limited sample for which the 30-year yield is available), and the sample variance of the 15-year yield. This gives $a = 29.1$, $b = 0.0056$, and $c = 0.97$.

Figure A1 reconsiders the LSAP-type comparative statics using the model with the imperfect-substitutability terms as calibrated above. I examine both the case in which the short rate is at its sample-average value (top) and in which it is at the ZLB with one year of forward guidance (bottom). The solid lines show the effect on the yield curve of reducing z by 0.7 years in the baseline model without imperfect substitutability (the same as in Figures 3 and 6). The dashed lines show the corresponding yield curves using the imperfect-substitutability model. The inclusion of these terms increases the curvature of the yield curve in both panels. It also increases the sensitivity of long-term yields to the level of the short rate, as can be seen by comparing the blue lines in the bottom panel. The sensitivity of yields to the supply distribution is also two to three times larger under this calibration than in the model with $b = c = 0$.

This result must be interpreted with caution, however. Almost the entirety of the difference between the effect of the LSAP in this model and the baseline is due to the much larger value of risk aversion used here, not to the effects of the convenience-yield terms themselves. If risk aversion is set to its baseline value of $a = 8$, maintaining the above values of b and c , the LSAP effect is almost identical to the effect in the model in which b and c were zero. Thus, the primary effect of the convenience yield is to change the average shape of the yield curve and its response to short-rate shocks, without having any major consequences for its sensitivity to the supply distribution. In the three-factor model of Section 5, where inflation and supply dynamics naturally contribute to the curvature and volatility of longer-term rates, the convenience-yield terms generally add little, and estimates of that model including those terms turn out to produce negligible values for b and c .

Preferred habitat

In the models considered in the text, the supply of debt that must be held by investors is exogenous. More generally, this distribution itself could depend on interest rates, as in the model of Vayanos and Vila (2009). Apart from its affine structure, the essential feature of Vayanos-Vila that differs from the model considered above is that investors (“arbitraguers” in their terminology)

face an elastic supply curve at each maturity. This supply curve is assumed to arise from the presence of preferred habitat agents, each of whom deals in debt of only one maturity. Specifically, for each bond, preferred-habitat agents demand a (market-value) quantity that is a function h of maturity and price:

$$p_{nt}\xi_{nt} = h(n, p_{nt}) \quad (\text{A1})$$

where ξ_{nt} is the par value demanded. The par value left in the hands of the investors is simply $x_{nt} - \xi_{nt}$, up to the constraints that no one can hold negative quantities. The models in the text effectively assumed $h(n, p_{nt}) = 0$ for all n . Equation (A1) can be substituted directly into (4) to give the return on the investors' portfolio:

$$R_t^w = \frac{\sum_{n=1}^N \bar{x}_{nt}(p_{nt})p_{n-1t+1}}{\sum_{n=1}^N \bar{x}_{nt}(p_{nt})p_{nt}}$$

where

$$\bar{x}_{nt}(p_{nt}) \equiv \min \left[x_{nt}, \max \left[0, x_{nt} - \frac{h(n, p_{nt})}{p_{nt}} \right] \right]$$

For a given demand function h , this model can be solved numerically using the solution algorithm described in the text.

B. Solution Algorithm

Models satisfying the conditions discussed in Section 2 can be solved numerically for the time- t vector of asset prices \mathbf{p}_t using the following iterative, discrete-state approximation method. I first make explicit that prices and quantities depend on the state of the economy. Namely, let $x_n(\mathbf{s}_t)$ describe how the quantity of asset n depends on the state. Let \mathbf{s}_t be Markov on the support \mathbf{S} with transition density $\tau(\mathbf{s}_{t+1}|\mathbf{s}_t)$. It is assumed that the form of the pricing kernel in equation (3), the laws of motion for the states, and the dependence of quantities on the states are known—that is, we (and investors) have knowledge of the functions $\tau(\mathbf{s}_{t+1}|\mathbf{s}_t)$, $M(\mathbf{s}_t, \mathbf{s}_{t+1}, R_{t,t+1}^w)$, and $x_n(\mathbf{s}_t)$. We seek a vector-valued function $\mathbf{p}(\mathbf{s}_t) = (p_1(\mathbf{s}_t) \dots p_N(\mathbf{s}_t))$ that describes how all asset prices depend on \mathbf{s}_t .

The price of asset n is given by

$$p_n(s_t) = \int_{\mathbf{S}} \tau(\mathbf{s}'|\mathbf{s}_t) M(\mathbf{s}_t, \mathbf{s}', R_{t,t+1}^w) q_n(s_{t+1}) d\mathbf{s}' \quad (\text{B1})$$

where

$$q_n(s_t) \equiv \begin{cases} p_{n-1}(\mathbf{s}_t) & \text{for } n > 1 \\ 1 & \text{for } n = 1 \end{cases}$$

and the integral is taken over all dimensions of the state. (The extension to cases with nonpecuniary returns is straightforward.) Given a distribution for

R_{t+1}^w , this is a system of linear Fredholm equations of the first kind, which in principle can be solved by quadrature in one step. However, the fact that $R_{t,t+1}^w$ is defined as in (4) requires us to iterate by, first, solving (B1) using a given distribution of $R_{t,t+1}^w$; second, given the resulting pricing functions finding the updated distribution of $R_{t,t+1}^w$; and repeating these steps to convergence.

Specifically, let $\mathbf{p}_d^i(\mathbf{s}_t)$ be a proposal for the pricing function on a discretization of the state space $\mathbf{D} = (\mathbf{d}_1, \dots, \mathbf{d}_G) \in \mathbf{S}^G$, where G is the number of nodes and $i = 0, \dots, I$ indexes iterations, and let \mathbf{q}_d^i be the corresponding discretization of $\mathbf{q}(\mathbf{s}_t)$. Suppose that the nodes are uniformly distributed over the state space, so that the conditional transition probability from node j to node h can be approximated by

$$\hat{\tau}(\mathbf{d}_h|\mathbf{d}_j) \equiv \tau(\mathbf{d}_h|\mathbf{d}_j) \left[\sum_{g=1}^G \tau(\mathbf{d}_g|\mathbf{d}_j) \right]^{-1}$$

\mathbf{p}_d^i and $\hat{\tau}$ are used to generate a proposal for the joint distribution of period $t+1$ states and prices. That distribution, in turn, generates an updated pricing function $\mathbf{p}_d^{i+1}(\mathbf{s}_t)$ through the analogue to equation (B1)—that is, by solving

$$p_{dn}^{i+1}(\mathbf{d}_j) = \sum_{g=1}^G \hat{\tau}(\mathbf{d}_g|\mathbf{d}_j) M \left(\mathbf{d}_j, \mathbf{d}_g, \frac{\mathbf{x}(\mathbf{d}_j)' \mathbf{q}_d^i(\mathbf{d}_j)}{\mathbf{x}(\mathbf{d}_j)' \mathbf{p}_d^i(\mathbf{d}_j)} \right) p_{dn-1}^{i+1}(\mathbf{d}_g) \quad (\text{B2})$$

for the vector $\mathbf{p}_d^{i+1}(\mathbf{d}_j)$ at each node $j = 1, \dots, G$.

This procedure converges in G and I , so long as the moments of the pricing kernel are well behaved. In particular, sufficient conditions for equation (B2) to constitute a contraction mapping on p_{dn} are that (i) the short-term interest rate is always non-negative and (ii) $M(\cdot)$ is linear in R_t^w . The Banach Theorem then guarantees for any given discretization \mathbf{D} , $\mathbf{p}_d^i(\mathbf{d}_j) \rightarrow \mathbf{p}_d(\mathbf{d}_j) \forall \mathbf{d}_j \in \mathbf{D}$, where \mathbf{p}_d is the (unique) pricing function that obtains if $\hat{\tau}$ is the data-generating process. But continuity of τ ensures that, for any node j ,

$$\lim_{G \rightarrow \infty} p_{dn}(\mathbf{d}_j) = E_t \left[p_{n-1}(\mathbf{s}_{t+1}) M \left(\mathbf{d}_j, \mathbf{s}_{t+1}, \frac{\mathbf{x}(\mathbf{d}_j)' \mathbf{q}_d^i(\mathbf{s}_{t+1})}{\mathbf{x}(\mathbf{d}_j)' \mathbf{p}_d^i(\mathbf{d}_j)} \right) \right]$$

i.e., in the limit, the pricing function solves the no-arbitrage condition (1).

Finally, by construction, if the algorithm converges, any point of convergence is a rational-expectations equilibrium. This follows immediately, since convergence is *defined as* the fixed point at which the joint distribution of \mathbf{p}_{t+1} and M_{t+1} is consistent with the vector \mathbf{p}_t , for each point in the state space.

It is important to note that, although the algorithm only solves for the vector of prices at G points in the state space, once these solutions are in hand it is straightforward to calculate equilibrium prices at any point through the Nystrom extension. In particular, take an arbitrary state value \mathbf{s}_t . For G large enough, we have

$$p_n(\mathbf{s}_t) \approx \left[\sum_{g=1}^G \tau(\mathbf{d}_g|\mathbf{s}_t) M \left(\mathbf{s}_t, \mathbf{d}_g, \frac{\mathbf{x}(\mathbf{s}_t)' \mathbf{q}_d^i(\mathbf{d}_g)}{\mathbf{x}(\mathbf{s}_t)' \mathbf{p}_d^i(\mathbf{s}_t)} \right) p_{n-1}(\mathbf{d}_g) \right] \left[\sum_{g=1}^G \tau(\mathbf{d}_g|\mathbf{s}_t) \right]^{-1}$$

Once the algorithm has converged, the quantities on the right-hand side are all known. Thus, securities can be priced in at any point in \mathbf{S} .

Figure B1 displays some results on the convergence of the solution algorithm for the one-factor model discussed in Section 3. The top panel shows the computed 5-, 10-, 15-, and 30-year yields, shown for a short rate at its average value of 5.8%, across the first 50 iterations ($i = 1, \dots, 50$). The algorithm is initialized at a price vector $\mathbf{p}_d^0(\mathbf{d}_j) = (1, \dots, 1)$ for all values of \mathbf{d}_j and uses $G = 65$ nodes across the state space. It is evident from this figure that, for each maturity n , the solution converges very quickly once $i > n$.

The middle panel shows convergence in the number of gridpoints by displaying the computed yield curve (after $I = 50$ iterations), again using $r_t = 0.058$ for illustration. Yield curves are shown for $G = 5, 9, 17, 33,$ and 65 , in each case spaced equally across possible values of p_{1t} . The space is assumed bounded between 0 and 0.80, so this partitioning corresponds roughly to increments of between 38 basis points and 5 percentage points. While 5 nodes is clearly too few to achieve convergence, the solutions using 17 or more nodes are indistinguishable from each other.

For brevity, these results were shown for the average value of the short rate. Similar convergence results obtain for other points in the state space, although solutions will not be accurate near the bounds if the underlying state process itself is not actually bounded. For example, in the above case, we would not expect the procedure to generate correct solutions near $p_{1t} = 0.80$ (corresponding to a risk-free return of 25%). However, so long as the bound on the state space is imposed far enough away from the values of the states that are actually realized in practice, this limitation has a negligible effect on the results. The bottom panel of the figure illustrates this claim by comparing the yield curve computed above with the yield curve computed when the grid for p_{1t} is extended over the entire range $(0, 1)$, again at the average value of the short rate. (The latter computation used $G = 100$.) The two curves are virtually identical, differing by less than 1 basis point across maturities.

References

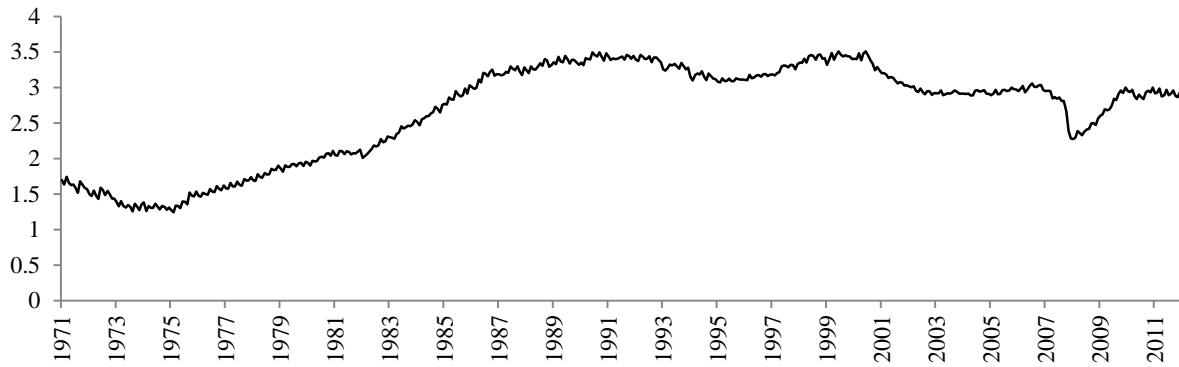
- Bauer, M. D. and Rudebusch, G. D., forthcoming. “The Signaling Channel for Federal Reserve Bond Purchases.” *International Journal of Central Banking*.
- Bernanke, B. S.; Reinhart, V., R.; and Sack, B. P., 2004. “Monetary Policy Alternatives at the Zero Bound: An Empirical Assessment.” *Brookings Papers on Economic Activity* (2): 1 – 78.
- Bernanke, B., 2010. *The Economic Outlook and Monetary Policy*. Remarks delivered at Jackson Hole, Wyoming, August.
- Cahill, M.; D’Amico, S.; Li, C; and Sears, J., 2013. “Duration Risk versus Local Supply Channel in Treasury Yields: Evidence from the Federal Reserve’s Asset Purchase Announcements.” Working paper (April).
- Campbell, J. Y.; Sunderam, A.; and Viceira, L. M., 2009. “Inflation Bets or Deflation Hedges? The Changing Risks of Nominal Bonds.” Working paper, Harvard University, January.
- Chabot, B. and Herman, G., forthcoming. “A History of Large-Scale Asset Purchases before the Fed.” FRB Chicago Economic Perspectives.
- Cox, J. C.; Ingersoll, J. E.; and Ross, S. A., 1985. “A Theory of the Term Structure of Interest Rates.” *Econometrica* 53(2): 385-408.
- Cox, J. C.; Ingersoll, J. E.; and Ross, S. A., 1981. “A Re-Examination of Traditional Hypotheses about the Term Structure of Interest Rates.” *Journal of Finance* 36(4): 769-99.
- Culbertson, J. M., 1957. “The Term Structure of Interest Rates.” *Quarterly Journal of Economics* 71(4): 485-517.
- D’Amico, S.; English, W.; Lopez-Salido, D.; and Nelson, E., 2012. “The Federal Reserve’s Large-Scale Asset Purchase Programmes: Rationale and Effects.” *Economic Journal* 122 (564): F415-46.
- D’Amico, S. and King, T. B., 2013. “Flow and Stock Effects of Large-Scale Treasury Purchases: Evidence on the Importance of Local Supply.” *Journal of Financial Economics* 108(2): 425-48.
- Doh, T., 2010. “The Efficacy of Large-Scale Asset Purchases at the Zero Lower Bound.” FRB Kansas City Economic Review (Q2): 5-34.
- Eggertsson, G. and Woodford, M., 2003. “The Zero Bound on Interest Rates and Optimal Monetary Policy.” *Brookings Papers on Economic Activity* (1): 139-211.
- Epstein, L. G. and Zin, S. E., 1989. “Substitution, Risk Aversion, and the Temporal Behavior of Consumption Growth and Asset Returns I: A Theoretical Framework.” *Econometrica* 57(4): 937-69.
- Femia, K.; Friedman, S.; and Sack, B., 2013. “The Effects of Policy Guidance on Perceptions of the Fed’s Reaction Function.” FRB New York Staff Report 652 (November).
- Frankel, J. A., 1985. “Portfolio Crowding-out, Empirically Estimated.” *Quarterly Journal of Economics* 100: 1041-65.
- Friedman, B. M., 1980. “The Effect of Shifting Wealth Ownership on the Term Structure of Interest Rates: The Case of Pensions.” *Quarterly Journal of Economics* 94(3): 567-90.

- Gagnon, J.; Raskin, M.; Remache, J.; and Sack, B., 2010. "Large-Scale Asset Purchases by the Federal Reserve: Did They Work?" *International Journal of Central Banking*.
- Greenwood, R.; Hanson, S. G.; and Stein, J. C., 2013. "A Comparative Advantage Approach to Government Debt Maturity." Working paper, September.
- Greenwood, R. and Vayanos, D., 2010. "Price Pressure in the Government Bond Market." *American Economic Review* 90(2): 585-90.
- Greenwood, R. and Vayanos, D., forthcoming. "Bond Supply and Excess Bond Returns." *Review of Financial Studies*.
- Gurkaynak, R. S.; Sack, B.; and Wright, J. H., 2006. "The U.S. Treasury Yield Curve: 1961 to the Present." *Journal of Monetary Economics* 54(8): 2291-394.
- Gurkaynak, R. S. and Wright, J. H., 2012. "Macroeconomics and the Term Structure." *Journal of Economic Literature* 50(2): 331-67.
- Hamilton, J. D. and Wu, J. C., 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., 2012. "Mortgage Refinancing, the Quantity of Risk, and Bond Risk Premia." Harvard Business School, Working paper (September).
- Hanson, S. G. and Stein, J. C., 2012. "Monetary Policy and Long-Term Real Rates." FEDS Working Paper 2012-46.
- Ihrig, J.; Klee, E.; Li, C.; Schulte, B.; and Wei, M., 2012. "Expectations about the Federal Reserve's Balance Sheet and the Term Structure of Interest Rates." FEDS working paper 2012-57.
- Kim, D. H. and Wright, J. H., 2005. "An Arbitrage-Free Three-Factor Term Structure Model and the Recent Behavior of Long-Term Yields and Distant-Horizon Forward Rates." FEDS paper 2005-33.
- Krishnamurthy, A. and Vissing-Jorgensen, A., 2011. "The Effects of Quantitative Easing on Interest Rates: Channels and Implications for Policy." *Brookings Papers on Economic Activity* (2): 215-65.
- Krishnamurthy, A. and Vissing-Jorgensen, A., 2012. "The Aggregate Demand for Treasury Debt." *Journal of Political Economy* 120(2): 233-67.
- Kuttner, K., 2006. "Can Central Banks Target Bond Prices?" NBER Working Paper 12454.
- Li, C. and Wei, M., 2012. "Term Structure Modeling with Supply Factors and the Federal Reserve's Large-Scale Asset Purchase Programs." *International Journal of Central Banking*.
- Markowitz, H. M., 1952. "Portfolio Selection." *Journal of Finance* 7(1): 77-91.
- Meaning, J. and Zhu, F., 2011. "The Impact of Recent Central Bank Asset Purchase Programmes." *BIS Quarterly Review*, Dec.
- Piazzesi, M. and Schneider, M., 2007. "Asset Prices and Asset Quantities." *Journal of the European Economic Association* 5(2-3): 380-89.
- Piazzesi, M. and Schneider, M., 2008. "Bond Positions, Expectations, and the Yield Curve." FRB Atlanta working paper 2008-2.

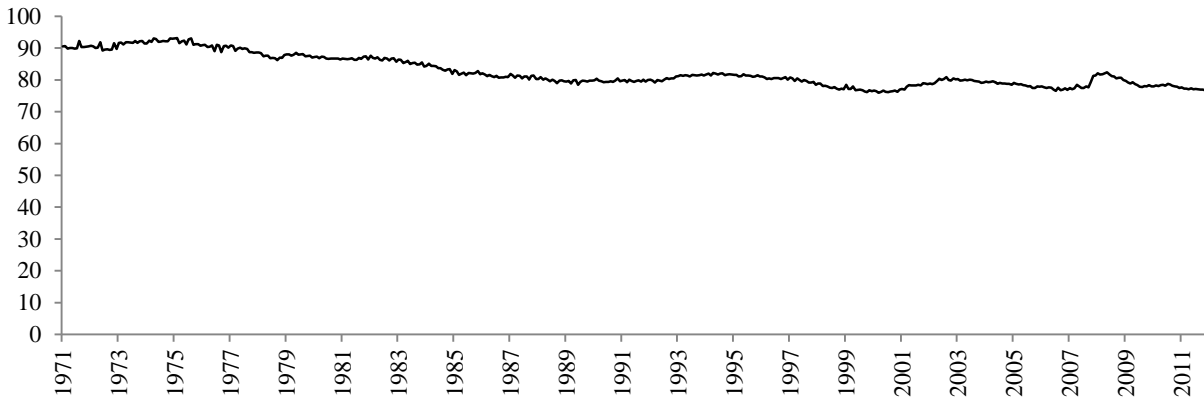
- Roley, V. V., 1982. "The Effect of Federal Debt Management Policy on Corporate Bond and Equity Yields." *Quarterly Journal of Economics*
- Roll, R., 1977. "A Critique of the Asset Pricing Theory's Tests. Part I: On Past and Potential Testability of the Theory." *Journal of Financial Economics* 4(2): 129-76.
- Rosa, C., 2012. "How 'Unconventional' Are Large-Scale Asset Purchases? The Impact of Monetary Policy on Asset Prices." FRB NY Staff Report 560, May.
- Rudebusch, G. D.; Sack, B. P.; and Swanson, E. T., 2007. "Macroeconomic Implications of Changes in the Term Premium." *FRB St. Louis Review* 89(4): 241-69.
- Sack, B., 2009. Remarks at the Money Marketeers of New York University, New York City, December 2: <http://www.newyorkfed.org/newsevents/speeches/2009/sac091202.html>.
- Svensson, L. E. O., 1994. "Estimating and Interpreting Forward Rates: Sweden 1992-4." NBER Working Paper 4871.
- Vasicek, O., 1977. "An Equilibrium Characterisation of the Term Structure." *Journal of Financial Economics* 5(2): 177-88.
- Vayanos, D. and Vila, J.-L.. 2009. "A Preferred-Habitat Model of the Term Structure of Interest Rates." Working paper.
- Wallace, N., 1981. "A Modigliani-Miller Theorem for Open-Market Operations." *American Economic Review* 71(3): 267-74.
- Weil, P., 1990. "Non-Expected Utility in Macroeconomics." *Quarterly Journal of Economics* 105(1): 29-42.
- Wright, J., 2012. "What Does Monetary Policy Do to Long-Term Interest Rates at the Zero Lower Bound?" Working paper (May).

Figure 1. Characteristics of U.S. government liabilities in public hands

A. Average duration

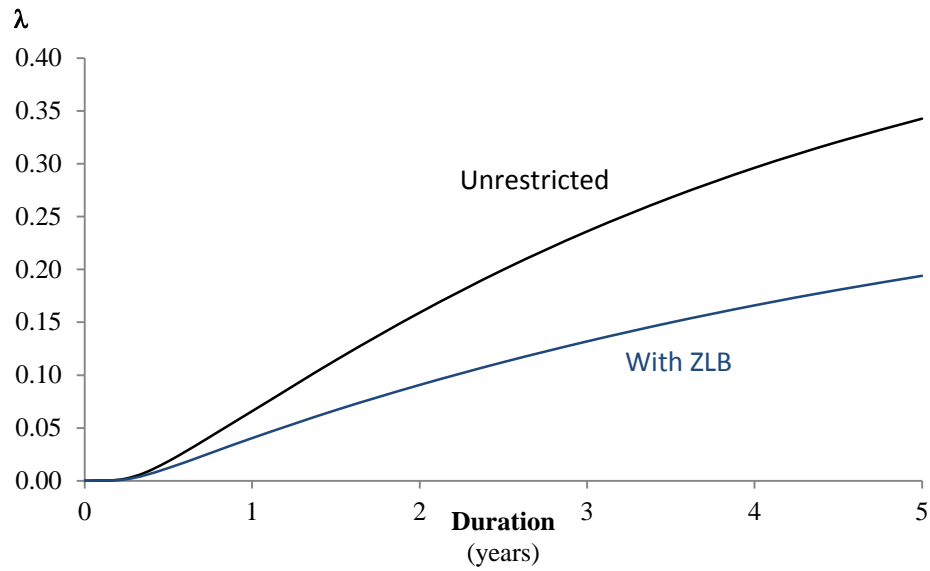


B. Percentage with duration < 5 years



Notes: Includes coupon notes, bonds, and bills issued by the U.S. Treasury, less the amount held in the Federal Reserve's SOMA portfolio, plus reserves and currency in circulation, which are assumed to have a duration of zero.
Sources: CRSP, Federal Reserve.

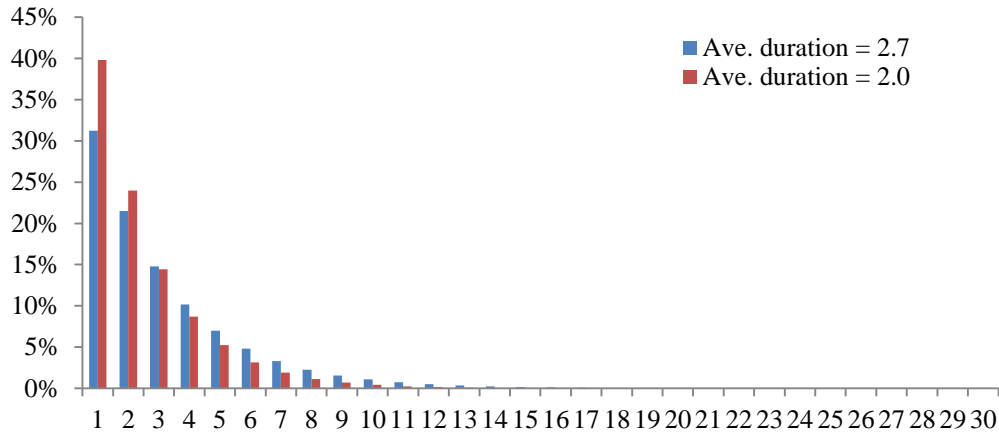
Figure 2. Risk prices in the one-factor model



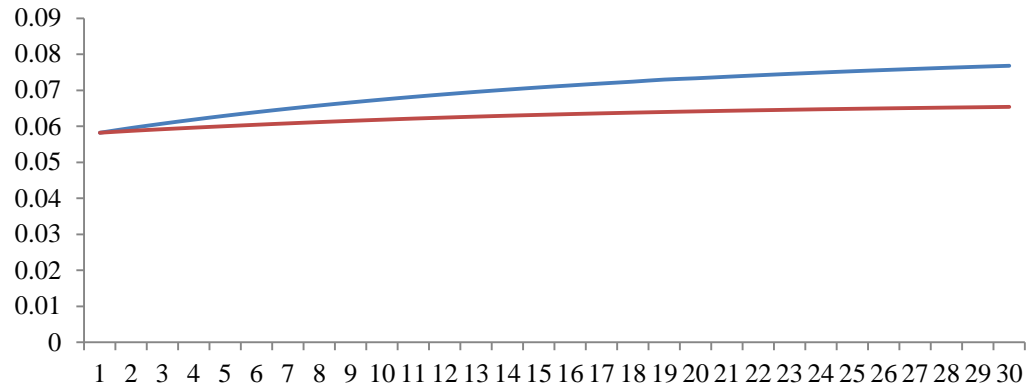
Note: The figure shows the price of short-rate risk (expected excess returns divided by their standard deviation) in the one-factor model for various values of aggregate duration (z). The black line corresponds to a model in which the short rate follows a simple AR(1) process. The blue line corresponds to a model in which that process is truncated at zero.

Figure 3. An LSAP-type shift in the one-factor model

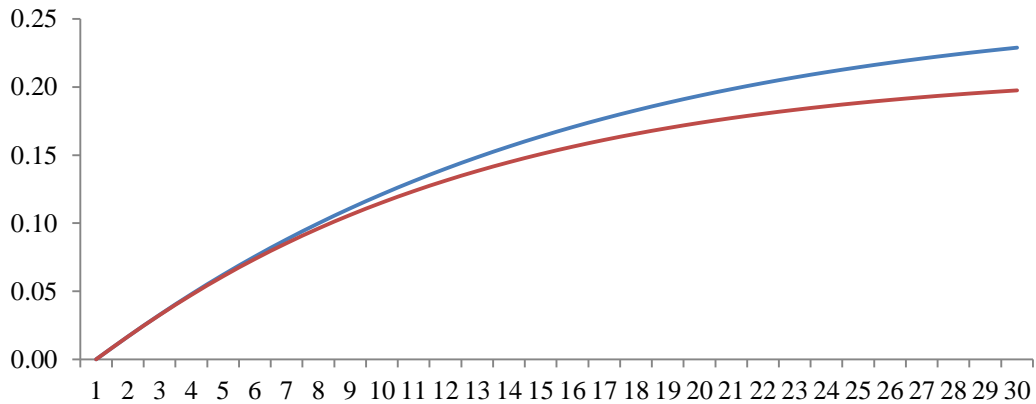
Duration distribution of outstanding debt



Yield curve



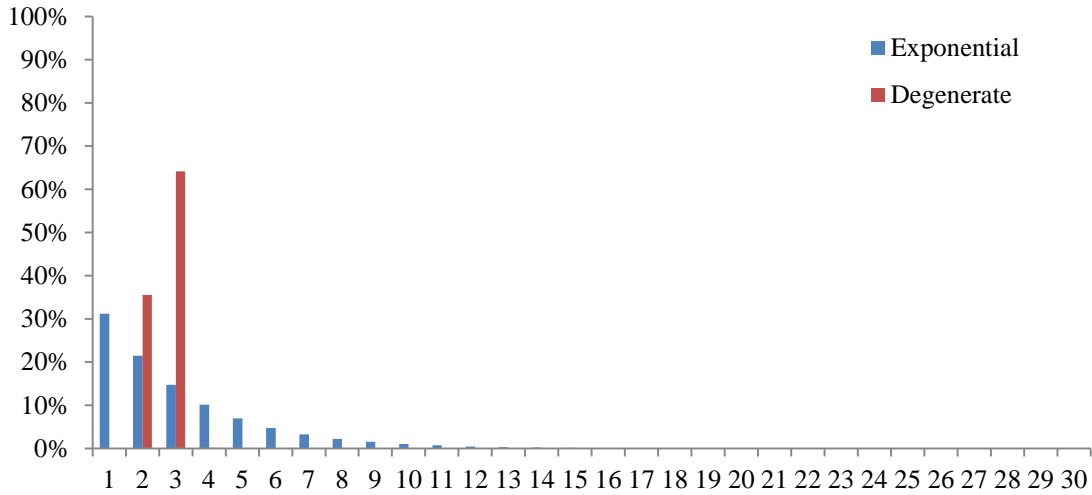
Return Volatility (One-Year)



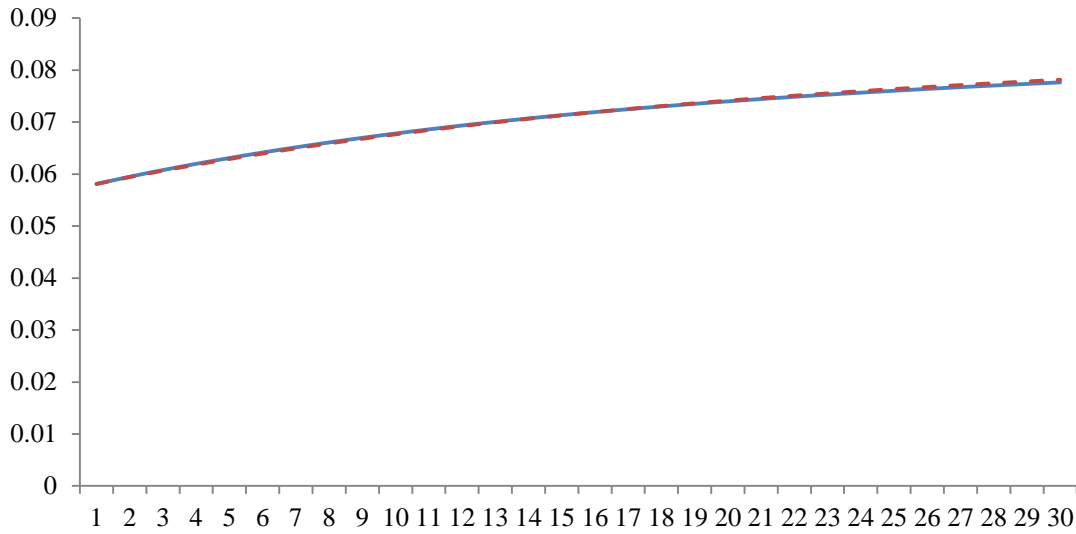
Notes: The figure shows the effect of moving from an average duration of 2.7 to 2.0 years in the one-factor model, evaluated at a short rate of 5.8% (the sample average).

Figure 4. Effect of the maturity-distribution shape in the one-factor model

Duration distribution of outstanding debt

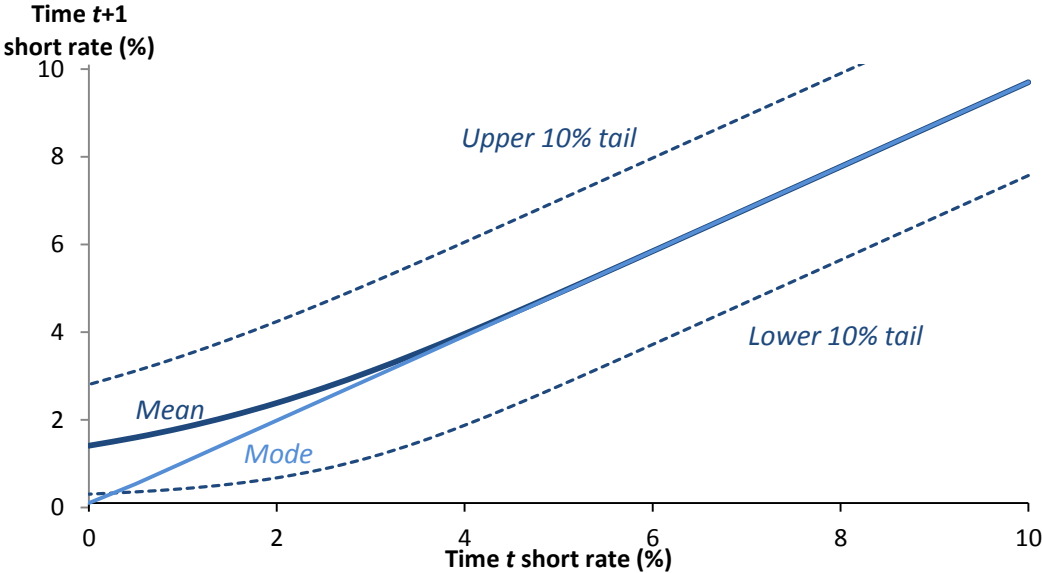


Yield curve



Notes: The top panel shows two possible distributions for the maturity of outstanding Treasury debt, one exponential and one nearly degenerate, both with means of 2.7 years. The bottom panel shows the corresponding yield curves generated by the one-factor model when the short rate is at its average value of 5.8%. In both cases, the risk-aversion coefficient is calibrated to match the average value of the 15-year yield.

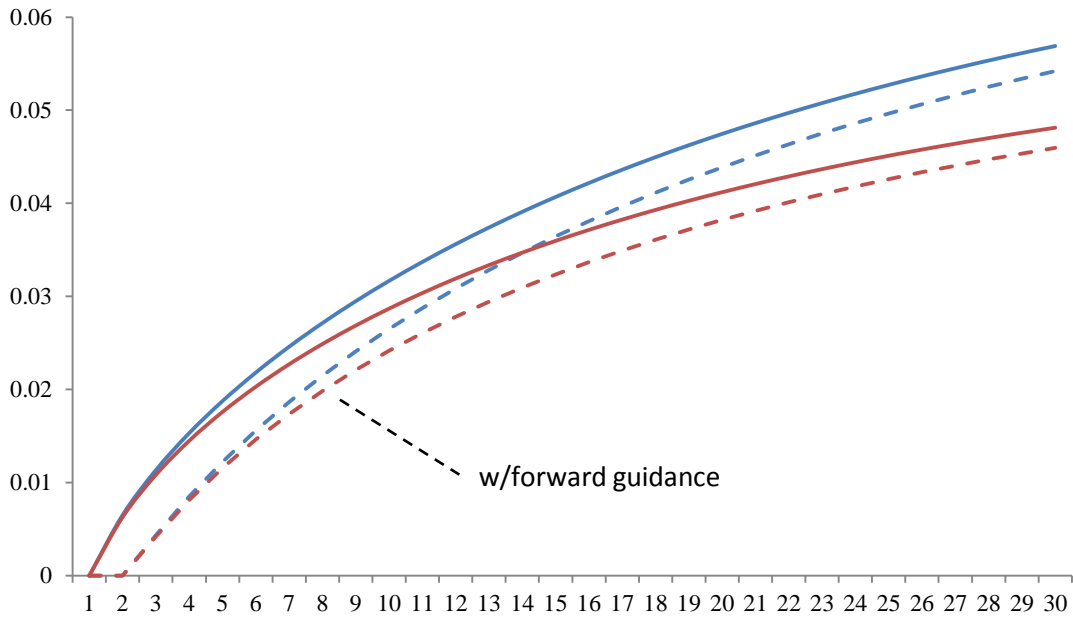
Figure 5. Estimated short-rate processes



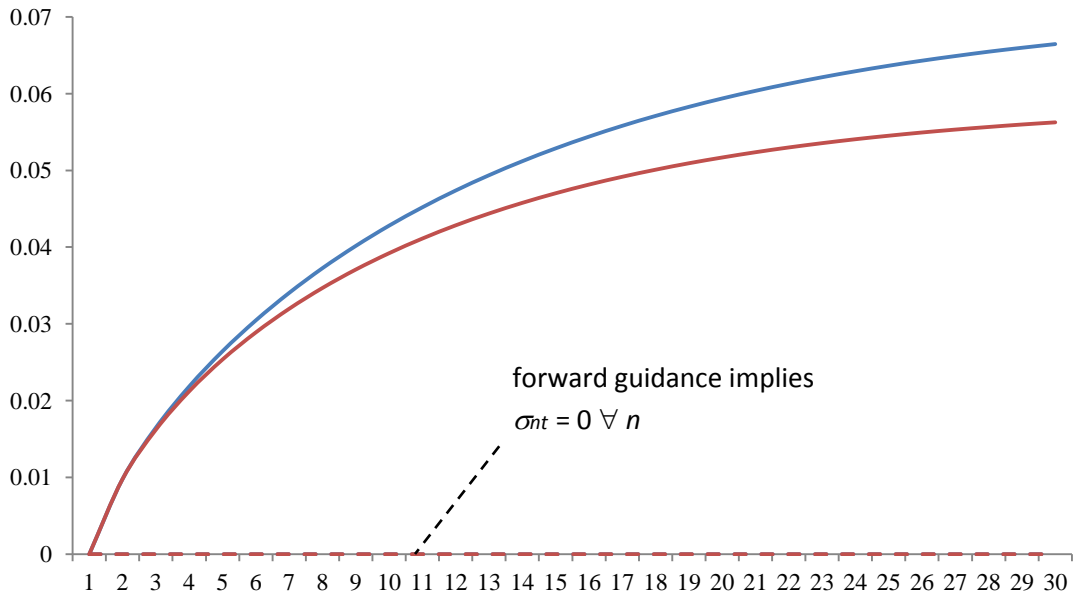
Notes: The figure shows the features of the conditional distribution of r_{t+1}^f , as a function of r_t^f near the ZLB, given by the truncated autoregressive process described by equations (XX) and (XX). Frequency is annual.

Figure 6. An LSAP-type shift in the one-factor model at the zero lower bound

Yield curve



Return Volatility (One-Year)



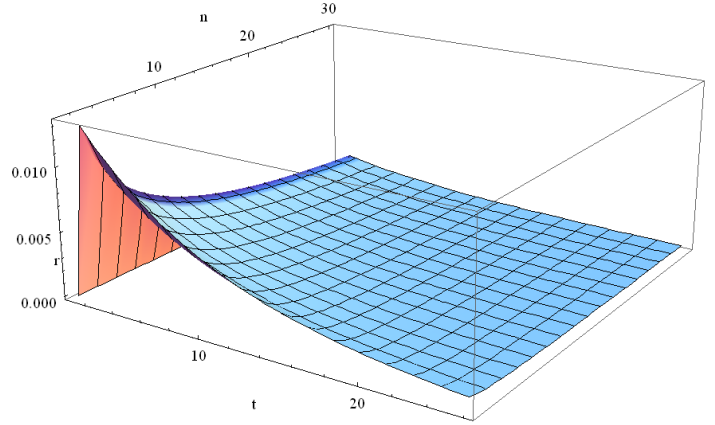
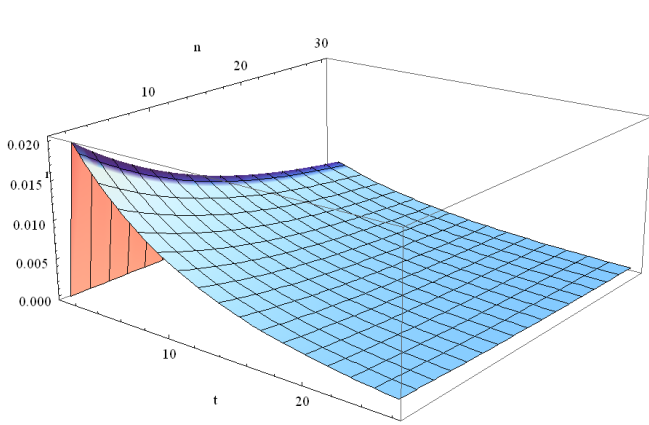
Notes: The figure shows the effect of moving from an average duration of 2.7 years (blue) to 2.0 years (red) in the one-factor model, evaluated at a short rate of 0%. The solid lines represent the case in which the short rate is expected to evolve according to its usual truncated-AR(1) process. The dashed lines (“forward guidance”) represent the case in which the short rate is anticipated to remain at 0% with certainty for one year and only then to follow its usual truncated-AR(1) process.

Figure 7. Impulse-response functions for the estimated three-factor model

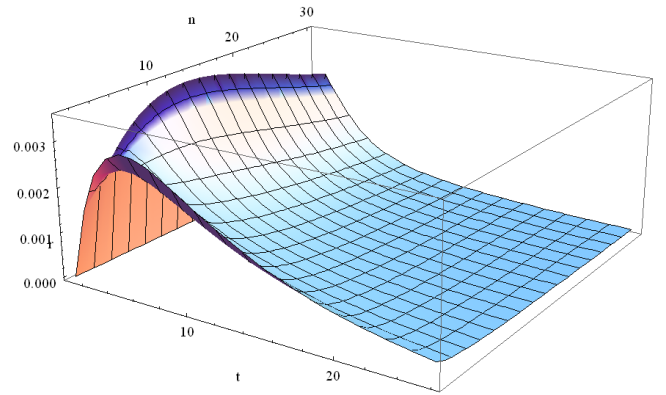
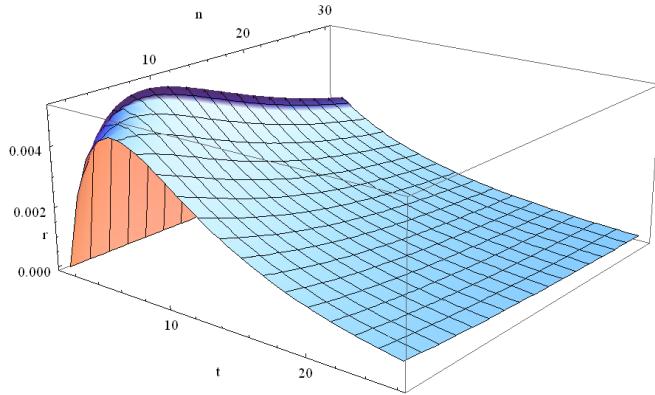
At sample means

At LSAP-period values

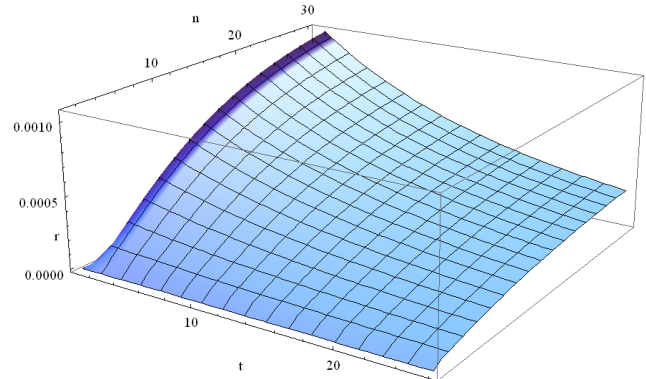
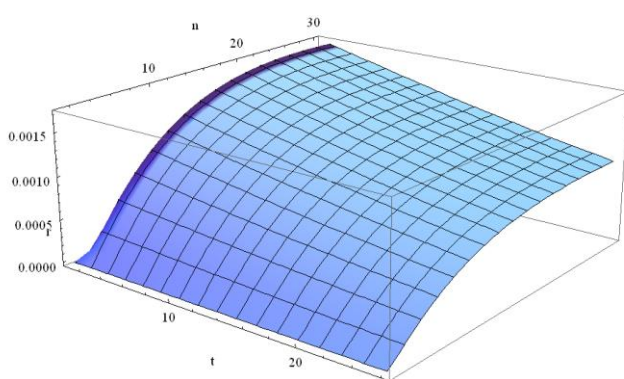
Response to short-rate shock



Response to inflation shock



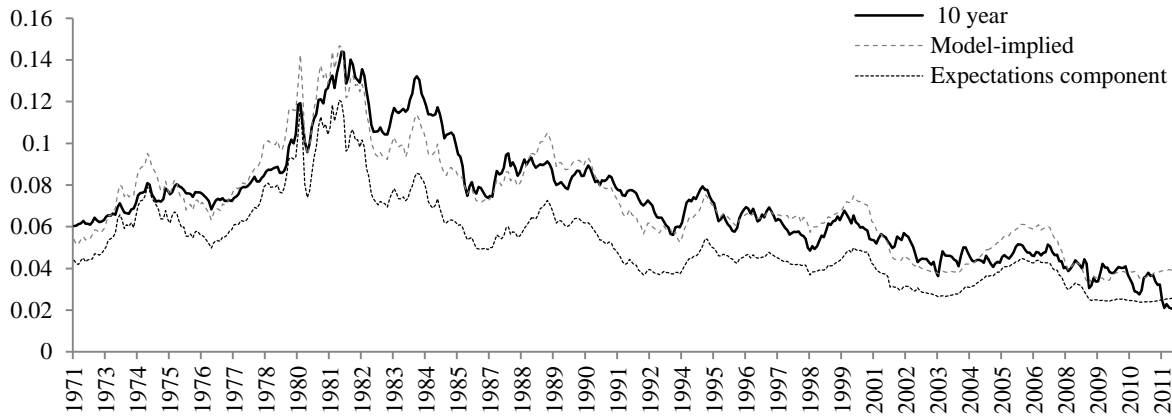
Response to supply shock



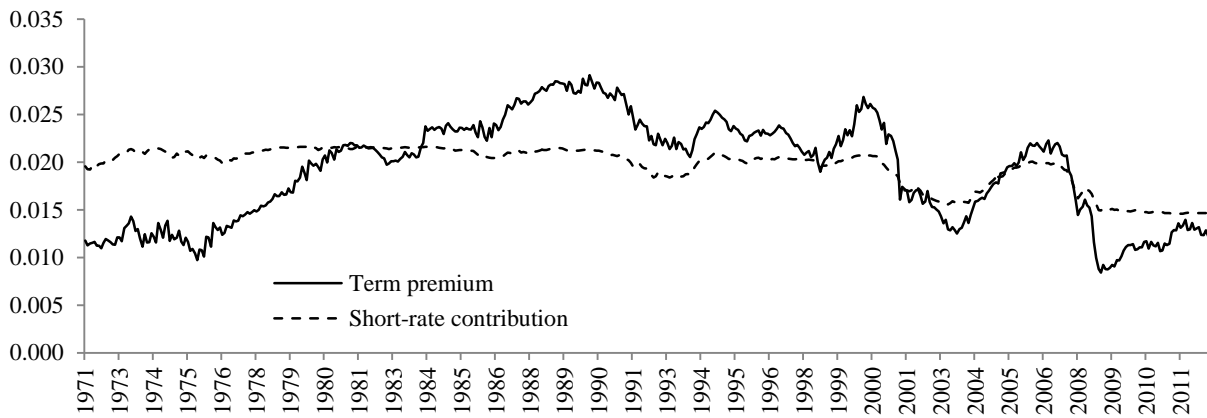
Notes: The figures show responses of the yield curve to one standard deviation positive shocks to the short-term interest rate, core PCE inflation, and average Treasury duration in the three-factor model, initialized at the 1971 - 2008 average values of the state variables (left panels) and at a set of state values representative of the 2008-2012 period (right panels). Maturities are plotted along the upper-left axis, and the lower-left axis represents calendar time (in years) after the shock.

Figure 8. Decomposition of the 10-year yield in the three-factor model

Yield components

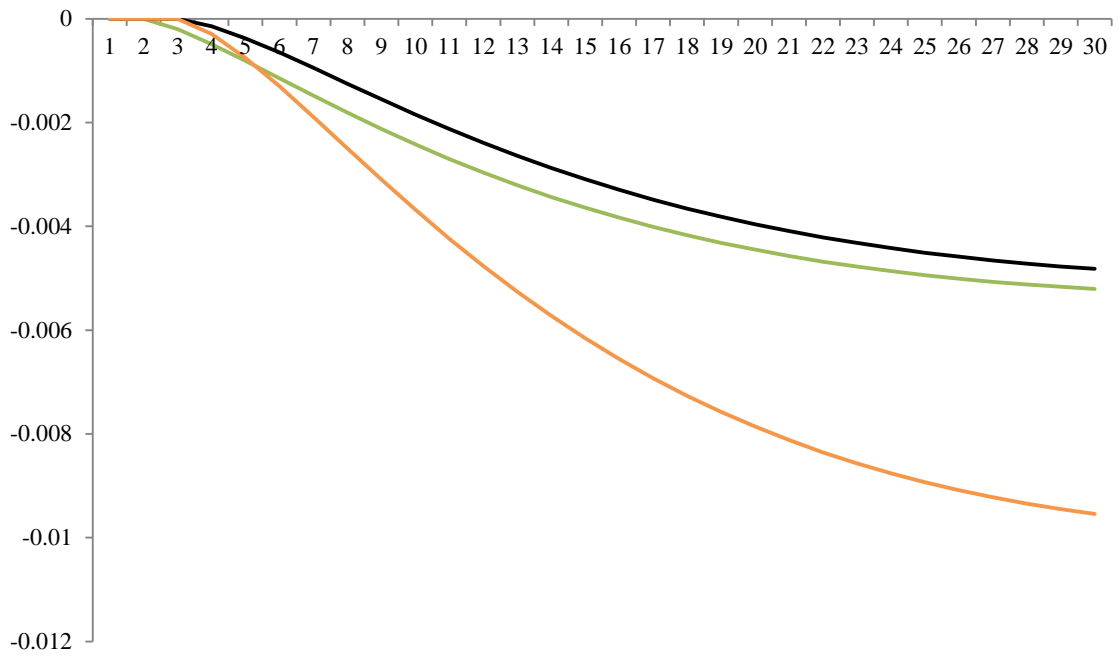


Term premium components



Notes: The model is estimated over the period Dec. 1971 – Dec. 2008. The term premium is calculated as the difference between the model-implied rate and the expectations component. Term premium shares are calculated as the difference between the term premium implied in the full model and that implied in counterfactual models in which inflation or Treasury supply are always equal to their sample-average values with certainty. The “short-rate contribution” to the term premium is calculated holding both Treasury supply and inflation equal to their average.

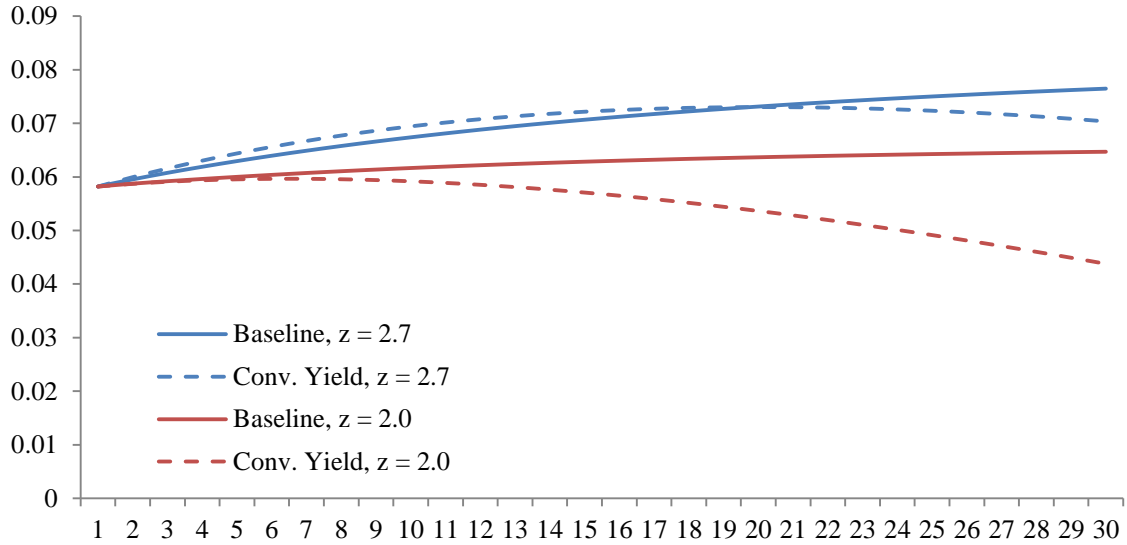
Figure 9. The effect of the LSAPs in the three-factor model



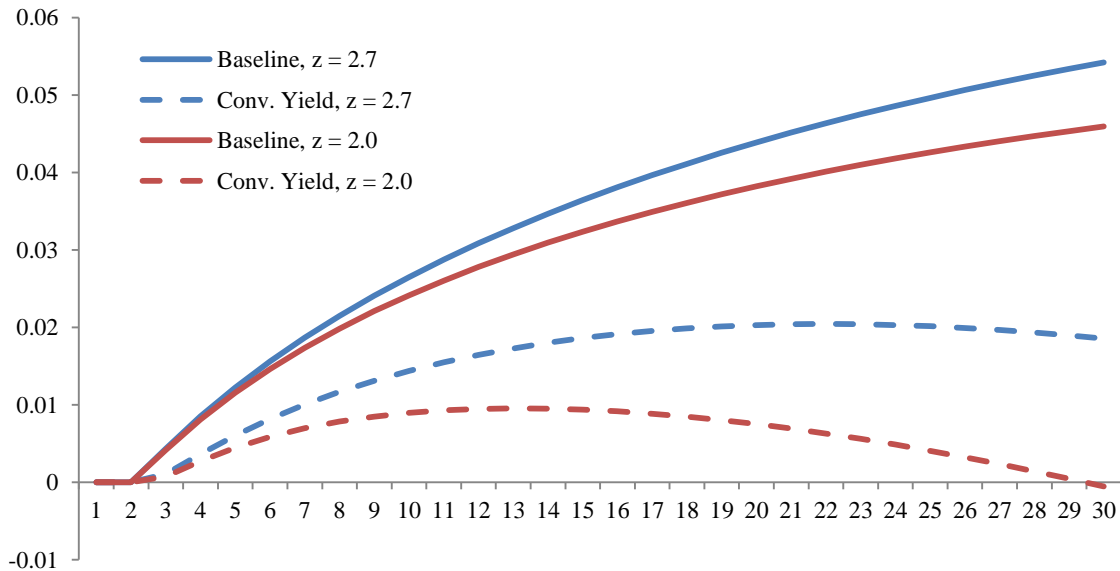
Notes: The black line shows the initial impact of a duration shock on the scale of the LSAP programs in the estimated three-factor model, evaluated at initial values approximating conditions experienced during the LSAP period, assuming that forward guidance leads market participants to expect short rates to remain at zero for two years. The green line shows the effect of the same shock when the forward guidance is for one year, rather than two. The orange line shows the effect of a duration shock of double the size.

Figure A1. Supply effects in the model with convenience yield

At the sample mean

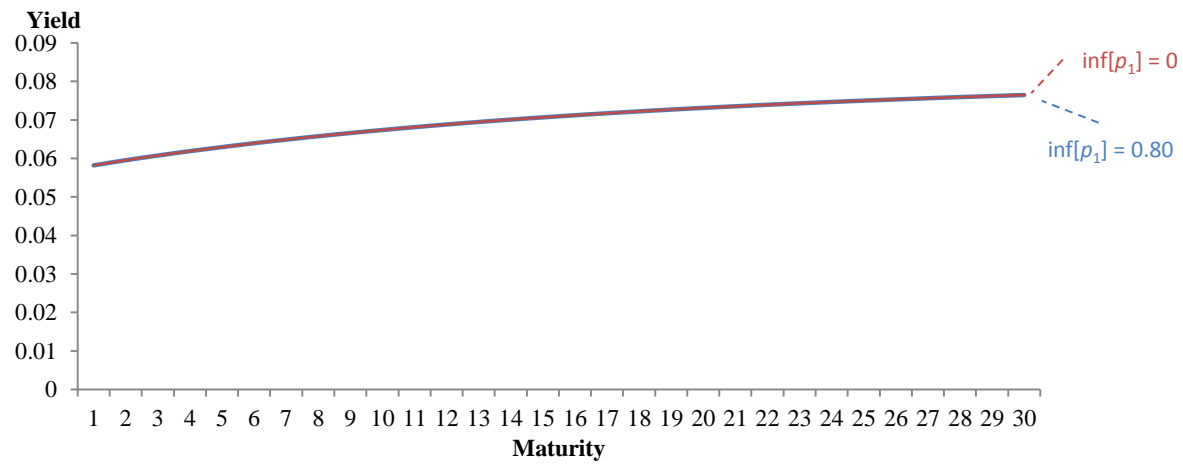
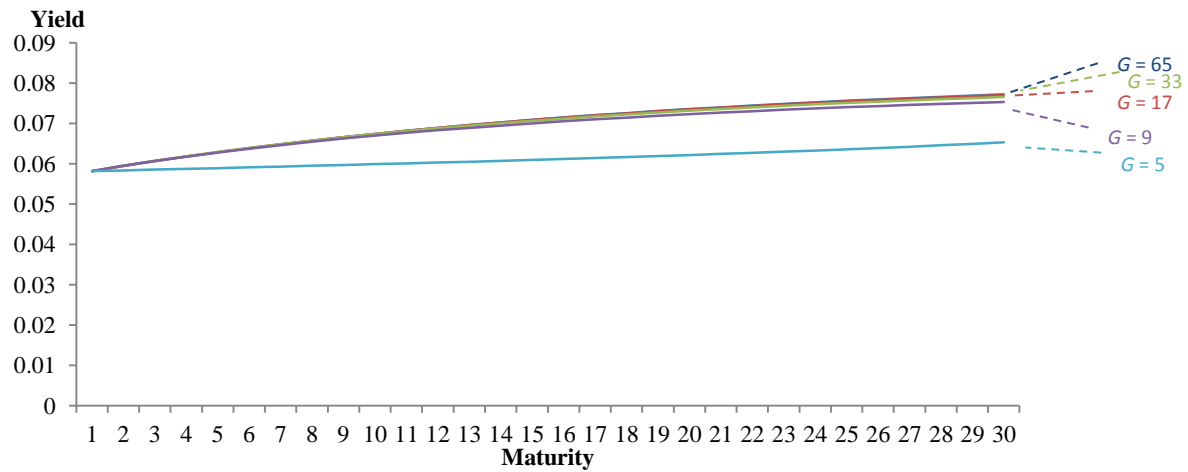
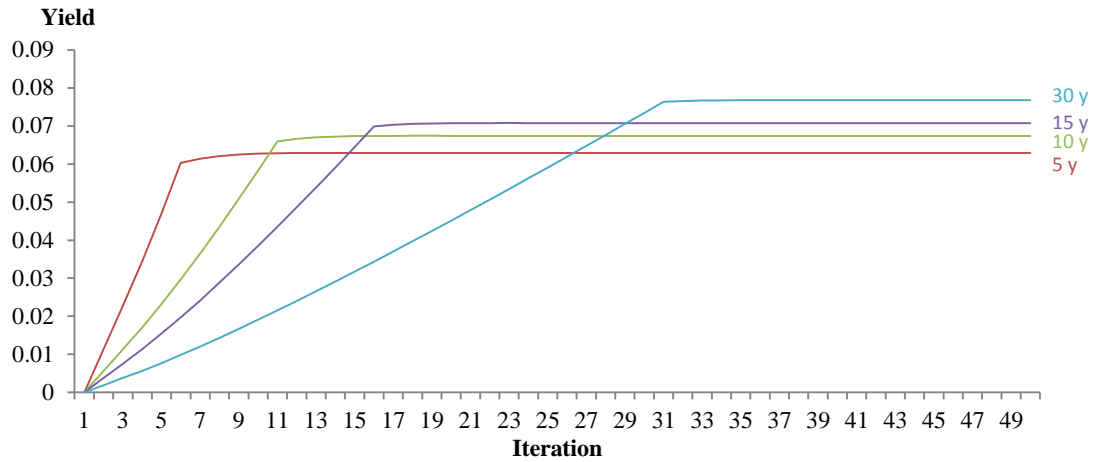


At the steady state / ZLB



Notes: The figures shows the effect of moving from an average duration of 2.7 years (blue) to 2.0 years (red) in both the baseline one-factor model (solid lines) and the model including the convenience-yield term (dashed lines). The top panel shows the case in which the short rate is at its sample mean of 5.8%. Bottom panels shows the case in which the short rate is at 0% and is anticipated to remain there with certainty for one year and then to follow its usual truncated-AR(1) process. Parameters are calibrated to the sample moments of yields, as described in the text.

Figure B1. Solution convergence in the one-factor model



Notes: The graphs show how the solution algorithm converges in the number of iterations (top), number of nodes in the grid (middle), and truncation point of the state space (bottom). All calculations are illustrated at the average value of the short rate.

Table 1. Parameter estimates for the three-factor model

	ϕ_0	ϕ_1	ϕ_2	$\log \sigma^f$	γ_0	γ_1	$\log \rho_0$	ρ_1	a
<i>Prior</i>	-0.005 (0.009)	0.8 (0.1)	0.4 (0.2)	-4.2 (0.4)	0.003 (0.005)	0.9 (0.2)	-20 (21)	0.003 (0.002)	12 (5)
<i>Posterior</i>	0.0005 (0.0010)	0.90 (0.01)	0.19 (0.03)	-4.0 (0.05)	0.0051 (0.0015)	0.71 (0.05)	-12.3 (2.1)	0.0023 (0.0007)	12.8 (1.0)

Notes: The table reports parameter estimate—means with standard deviations in parentheses—for the model in equations (33) through (35). The prior distribution is jointly normal and is based on maximum-likelihood estimation of the system using annual data on the one-year Treasury rate and core PCE inflation. The reported posterior values are based on 50,000 Metropolis draws, using monthly data on 5-, 10-, and 15-year yields from 1971 – 2008.

Table 2. Comparison of the three-factor model to the data.

A. Yields (%)

		1yr	5yr	10yr	15yr	30yr (beg. 12/85)
Data	Mean	6.4	7.1	7.5	7.7	6.6
	Std. Dev.	3.0	2.6	2.4	2.3	1.5
	1-year autocor.	0.81	0.87	0.88	0.88	0.91
	Corr w/short rate	1	0.96	0.92	0.89	0.71
Model	Mean	6.4	7.0	7.4	7.6	7.0
	Std. Dev.	3.0	2.7	2.3	2.1	1.3
	1-year autocor.	0.81	0.82	0.83	0.84	0.87
	Corr w/short rate	1	1.00	0.99	0.97	0.88
	R^2	--	0.92	0.87	0.84	0.73

B. Forward rates (%)

		1yr	5yr	10yr	15yr	30yr (beg. 12/85)
Data	Mean	6.4	7.5	8.0	8.1	6.0
	Std. Dev.	3.0	2.4	2.1	2.2	1.9
	1-year autocor.	0.81	0.88	0.87	0.88	0.83
	Corr w/short rate	1	0.88	0.82	0.81	0.61
Model	Mean	6.4	7.5	8.0	8.0	6.9
	Std. Dev.	3.0	2.4	1.8	1.4	1.1
	1-year autocor.	0.81	0.83	0.85	0.86	0.88
	Corr w/short rate	1	0.98	0.94	0.87	0.69
	R^2	--	0.80	0.77	0.68	0.54

See next page for notes.

Table 2 continued.

C. Annual excess returns (%)

		1yr	5yr	10yr	15yr	30yr (beg. 12/85)
Data	Mean	0	2.1	3.4	4.4	15.3
	Std. Dev.	0	6.5	12.4	17.9	29.8
	1-year autocor.	--	0.04	-0.09	-0.12	-0.58
	Corr w/short rate (end of period)	--	-0.48	-0.45	-0.40	-0.20
Model	Mean	0	1.7	2.9	3.6	2.7
	Std. Dev.	0	6.3	12.2	16.4	5.6
	1-year autocor.	--	0.20	0.19	0.18	0.01
	Corr w/short rate (end of period)	--	-0.44	-0.43	-0.43	-0.52
	R^2	--	0.63	0.40	0.33	0.14

Notes: Data are monthly averages of daily values, Dec. 1971 – Dec. 2008, except for 30-year yield. Model results are generated using the mean values of the parameters reported in Table 1, which were estimated over the 5-, 10-, and 15-year spot yields. Reported forward rates are one-year rates ending 1, 5, 10, 15, and 30 years ahead. Excess returns are computed on non-overlapping samples, December to December of each year, and are calculated relative to the initial one-year yield. Excess returns in the data are calculated from the Gurkaynak et al. (2007) zero-coupon yields. Shading indicates model output that exactly matches the data by construction.

Table 3. Contributions to the ten-year yield in the three-factor model

	Inflation factor	Supply factor	Both factors
Yield	0.07	0.06	0.06
Term premium	0.39	0.70	0.94

Notes: Table reports the percentage by which the total variance of the ten-year yield or term premium falls when the indicated factors are held at their sample-mean values, and investors price these levels in with certainty. Variances are computed over monthly first differences, 1971 – 2012.

Table 4. Summary of Federal Reserve asset-purchase programs

	Dates	Net quantity of Treasuries purchased (\$bil)	Net quantity of MBS purchased (\$bil)	Net quantity of reserves created (\$bil)	Assumed change in average Treasury duration outstanding	Empirical effect on ten-year term premium
LSAP I	Dec. 2008 – May 2010	\$300	\$941	\$1,318	0	40-100 bp
MBS reinvestment	Aug. 2010 – Oct. 2011	\$285	\$0	\$285	0	10-25 bp
LSAP II	Nov. 2010 – July 2011	\$600	\$0	\$600	0	15-55 bp
MEP	Oct. 2011 – Dec. 2012	\$0	\$0	\$0	-0.5	30-65 bp
Total		\$1,185	\$941	\$2,203	-0.5	95-245 bp

Notes: All quantities are net of redemptions and principal payments through December 2012. The assumed duration change is only that in Treasury securities (i.e., excluding agency debt, MBS, and the monetary base). The empirical effect of each program on the ten-year nominal Treasury yield is taken from the literature discussed in Section 6 of the text.

Table 5. Effects of alternative parameter assumptions on the LSAP impact

		Effect of LSAPs on 10y term premium (bp)	Ave. 10y slope, Dec. 2008 – Dec. 2012 (bp)
Data		95 - 245	268
<i>Parameters</i>	<i>Forward guidance</i>		
Estimated Model	2 years	18	214
$a = 12.7$	1 year	24	273
$a = 25.4$	2 years	46	328
	1 year	60	420

Notes: The estimated model is the three-factor model described in Section 6 of the text, using data from 1971 – 2008, with parameters evaluated at the posterior mean. The LSAP shock is assumed to be a decrease in average duration outstanding from 2.7 to 2.0 years. “Forward guidance” means that investors know that the short-term interest rate will be equal to its current value for the indicated time after the shock.