The Effects of Treasury Debt Supply on Macroeconomic and Term Structure Dynamics

Thomas Laubach and Min Wei *

Preliminary and Incomplete
September 2013

Abstract

JEL classification: E6, H6.

Keywords: Affine term structure models.

*Federal Reserve Board, thomas.laubach@frb.gov, min.wei@frb.gov. The views expressed in this paper are those of the authors and do not necessarily reflect views of the Federal Reserve Board or the Federal Reserve System.

1 Introduction

The relationship between the macroeconomy and asset pricing has been a long-standing area of research. One aspect that has received much attention is the interaction between the macroeconomy and the term structure of interest rates, and how this interaction is shaped by the conduct of monetary policy. Whereas the monetary policy rule determines the comovement of macroeconomic aggregates and expectations of future short-term interest rates, its implications, and those of other aspects of the macroeconomy, for the determination of term premia are less well understood. The use of large-scale asset purchases or "quantitative easing" (QE) by a number of central banks since the intensification of the financial crisis in 2008 has lent urgency to a better understanding of macroeconomic determinants of the yield curve, as those policy measures are thought to operate to a large extent by influencing term premia.

In this paper, we provide empirical evidence on the joint response of US Treasury yields, the embedded term premia, and macroeconomic variables to changes in the supply of Treasury securities. To do so, we follow the approach first developed in Ang and Piazzesi (2003) of combining an affine model of the term structure of Treasury yields with the assumption that the factors that matter for pricing these bonds are macroeconomic variables whose joint dynamics can be described by a linear VAR model. The focus of our analysis is on identifying exogenous variations in Treasury supply emanating from fiscal and monetary policy and to estimate the response of yields and term premia across the maturity spectrum. The objective is to provide empirical evidence on these responses that is based on minimal identifying assumptions, so as to uncover "stylized facts" for future study in more fully specified structural models.

An important motivation, as mentioned before, is the recent use of QE by a number of central banks. The macroeconomic effects of these policies are widely debated (some relevant studies will be discussed below). In many instances, the study of these effects has proceeded in two steps. First, there are a number of case studies that document the response of various asset prices to QE-related news in narrow event windows. Whereas with the proper choice of event window one can hope to capture the asset price responses of such news, the responses of slower-moving macroeconomic variables such as output and inflation can of course not be measured in this way. Therefore, in a second step, the asset price responses identified from the event studies are used in structural macroeconomic models to obtain results for the effects of QE on output, the unemployment rate, and inflation. This second step involves the choice of a particular structural model which implies numerous, and often contentious, assumptions about the transmission mechanism of the financial effects of QE.

In contrast to this two-step approach to the estimation of QE effects, in this study we model the dynamics of macro and financial variables jointly. In the spirit of the structural VAR literature, we aim to reduce the imposition of prior assumptions about the dynamic responses of macroeconomic variables to QE events to a minimum of identifying assumptions. We make use of data on both total supply of marketable securities by the Treasury and Treasury holdings by the Federal Reserve and foreign official institutions, and build on the literature on fiscal SVARs to disentangle exogenous innovations to Treasury supply emanating from fiscal policy from those emanating from monetary policy. Doing so is arguably important, as the effects of the former reflect both the effects of changes in the amount of Treasuries held in private portfolios and the effects of the tax and spending decisions that bring about the change in Treasury supply, whereas QE-style changes in Treasury supply are not associated with fiscal policy changes.

Aside from the study of the dynamic effects of QE, there is a long-standing literature examining the information content in the term structure, and especially its slope, for future economic activity. A premise of the macro-finance literature is that risk premia, including the term premia for bearing duration risk, are endogenous variables that can be affected by various macroeconomic disturbances. Hence, reduced-form regressions of measures of economic activity on the slope or some other measure of term premia are unlikely to uncover the partial effect of a change in term premia induced by a policy intervention such as QE. The impulse response functions of yields and term premia to various shocks produced by our model shed some light on the correlation patterns between term premia and real activity induced by these shocks.

Our [highly preliminary] findings are that (i) exogenous increases in Federal Reserve or foreign official holdings raise output significantly and by similar amounts as estimated in previous studies, but that inflation rises substantially, and that these effects are tempered by an increase in the short-term interest rate; (ii) that exogenous increases in Treasury supply of the same size induced by fiscal policy lead to significantly larger responses of output, inflation, and short-term interest rates; and (iii) that long-term Treasury yields rise in response to both types of shocks, but that the response to an increase in Federal Reserve or foreign official holdings is associated with a sizeable and significant decline in term premia, whereas fiscal policy-induced Treasury supply changes are not.

In the remainder of this introduction we discuss a few studies that are closely related to ours. In section 2, we discuss the specification of our term structure model, the identification strategy, and the data we use. Section 3 presents our empirical findings, and section 4 offers conclusions. The appendix spells out further details regarding model specification, data, and identification.

1.1 Related literature

The studies by Ang and Piazzesi (2003) and by Ang, Piazzesi, and Wei (2006) pioneered the use of macroeconomic variables as factors in affine term structure models, based on the intuition that, if the central bank varies short-term interest rates systematically in response to economic variables, these variables must be relevant for bond pricing. However, this use of macroeconomic variables as pricing factors is not uncontroversial. If macro variables were the only factors affecting bond pricing, regressions of macro variables on yields should show high R^2s, whereas in fact they do not (Kim, 2009; see Gürkaynak and Wright (2012) for further discussion). Instead there seems to be information in macro variables that is relevant for future short-term rates and future excess returns on longer-term bonds, but this information does not affect current bond pricing ("unspanned macro risks," see Duffee, 2011, and Joslin et al., 2013). We nonetheless model yields as being driven by fundamental macroeconomic disturbances to facilitate identification of economically interpretable shocks, recognizing that our model has only limited ability to explain yields.

Early case studies of the responses of long-term (Treasury and private) yields to QE-related announcements by the Federal Reserve include Gagnon et al. (2011) and Krishnamurthy and Vissing-Jorgensen (2011). Both studies find significant effects of announcements that raised expectations of Federal Reserve purchases of Treasuries or Agency MBS on yields of Treasury securities, MBS, and to a lesser extent corporates. An important question in this regard is the precise channel through which expectations of asset purchases reduce yields. They might do so by reducing the path of expected future short-term rates (the "signaling channel") or by reducing term premia that investors demand, for example because of the associated reduction in duration risk. Krishnamurthy and Vissing-Jorgensen

(2011) and D'Amico et al. (2012) provide evidence on the relative strength of various channels, where D'Amico et al. focus on evidence on Treasury yields at the CUSIP level. Kiley (2013a, b). (To be completed.)

Chung et al. (2012), Chen et al. (2012)

Li and Wei (2013), Vayanos and Vila (2009)

Blanchard and Perotti (2002), Dai and Philippon (2006).

2 The term structure model with supply factors

In this section we specify an affine model of the term structure of nominal US Treasury zero-coupon securities, in which the factors are macroeconomic variables. Our choice of factors reflects views on which fundamental sources of economic uncertainty are likely reflected in bond prices. These include long-run risks, such as perceptions of persistent changes in inflation or the equilibrium real interest rate, as well as variables relevant for determining the amount of duration risk held by private investors, such as total Treasury supply and foreign and domestic official holdings of Treasuries.

2.1 An arbitrage-free term structure model with supply factors

The model used to estimate the macroeconomic and term structure effects of Treasury supply shocks consists of a description of the relationships between major macroeconomic and Treasury supply variables, and a specification of the stochastic discount factor that ensures that pricing of bonds at various maturities is arbitrage-free. The model prices nominal zero-coupon bonds that are free of default risk.

For reasons discussed further below, we consider inflation π_t and the one-period real interest rate r_t as each being composed of a time-varying asymptote, denoted by a bar, and stationary deviations from this asymptote or trend:

$$\pi_t = \bar{\pi}_t + \tilde{\pi}_t \tag{1}$$

$$r_t = \bar{r}_t + \tilde{r}_t \tag{2}$$

The trend component of inflation may reflect time-varying perceptions of the central bank's inflation objective. Since we assume that $\tilde{\pi}_t$ has an unconditional mean of zero, the infinite-horizon expectation of inflation at time t is given by $\bar{\pi}_t$. Analogously, the trend component of the one-period real interest rate may reflect time-varying perceptions of trend productivity growth or changes in risk attitudes that affect the infinite-horizon expectation of the equilibrium real rate of return on short-term risk-free assets.[1]

With this notation, we define the vector of state variables or factors driving yields as

$$x_t = [\bar{\pi}_t,\ \bar{r}_t,\ \tilde{\pi}_t,\ q_t,\ t_t,\ s_t,\ \tilde{y}_{1t}]'$$

where $q_t,\ t_t,$ and s_t are measures of real activity, total marketable Treasury debt, and domestic and foreign official holdings of Treasuries, respectively, and the detrended one-period nominal yield $\tilde{y}_{1,t}$ is linked to the observed one-period yield $y_{1,t}$ by

$$y_{1,t} = \tilde{y}_{1,t} + \bar{\pi}_t + \bar{r}_t \tag{3}$$

The dynamics of the state vector follow a VAR

$$X_t = \mu + \Phi X_{t-1} + \Omega V_t,\ t = 1, \ldots, T \tag{4}$$

Under the assumption (discussed further below) that the last five elements of x_t follow a VAR(2), whereas trend inflation and the trend real short-term interest rate are assumed to follow univariate random walks, the vectors and matrices in (4) are of the form

$$X_t \equiv \begin{bmatrix} x_t \\ x_{t-1} \end{bmatrix},\ V_t \equiv \begin{bmatrix} v_t \\ 0 \end{bmatrix},\ \Phi = \begin{bmatrix} I_2 & 0 & 0 & 0 \\ 0 & \phi_1 & 0 & \phi_2 \\ I_2 & 0 & 0 & 0 \\ 0 & I_5 & 0 & 0 \end{bmatrix} \tag{5}$$

where ϕ_1 and ϕ_2 are coefficient matrices of size 5×5, and I_n symbolizes the identity matrix of size $n \times n$. Trend inflation $\bar{\pi}_t$ and the trend real interest rate $\bar{r}_t$ are latent factors, which, however, will be tightly constrained in the estimation, as discussed further below. The last five factors are observed without measurement error.

In specifying the stochastic discount factor that prices nominal bonds at different maturities, we follow the large literature on exponentially-affine term structure models (e.g. Duffee, 2002). Let

$$\xi_{t+1} = \exp\left(-\frac{1}{2}\lambda_t'\lambda_t - \lambda_t' V_{t+1}\right)$$

[1] Spencer (2008) presents a term structure model in which yields also depend on time-varying asymptotes of infation and the real short rate.

denote the Radon-Nikodym derivative that converts the data-generating to the risk-neutral probability measure, such that

$$\begin{aligned} P_{nt} &= E_t[M_{t+1}P_{n-1,t+1}] = E_t[e^{-y_{1,t}}\xi_{t+1}P_{n-1,t+1}] \\ &= E_t^{\mathbb{Q}}[e^{-y_{1,t}}P_{n-1,t+1}] \end{aligned}$$

where $E_t^{\mathbb{Q}}$ denotes expectation under the risk-neutral measure and λ_t are the prices associated with the macroeconomic risks. These risks are given by the (as yet to be identified) fundamental innovations ε_t, which are assumed to be i.i.d. standard normal.

A key assumption is the specification of the prices of risk λ_t as a general linear function of the states:

$$\lambda_t = \lambda_0 + \lambda_1 X_t \tag{6}$$

This specification plays an important role in enabling the model to explain the observed failures of the expectations hypothesis (Dai and Singleton, 2002) and to forecast yields (Duffee, 2002). We assume that the stochastic discount factor prices only current innovations v_t, and that the prices of risk λ depend only on the current state x_t. Other restrictions on λ_0 and λ_1 will be discussed below. Given the decomposition (3) of the short rate, it can be expressed as

$$y_{1t} = \delta_0 + \delta_1 X_t \tag{7}$$

with $\delta_0 = 0$ and δ_1 selecting the first two and the last element of x_t. The model then implies that the yield on a nominal zero-coupon bond with n periods to maturity is a linear function

$$y_{n,t} = a_n + b_n' X_t \tag{8}$$

where the coefficients $a_n,\ b_n$ are determined recursively.

2.2 Data and survey expectations used in estimation

For our empirical implementation we specify the model at quarterly frequency. This choice reflects that fact that we are interested in measuring the effects of changes in Treasury supply stemming from fiscal or monetary policy actions on macroeconomic variables and yields jointly, and recognizes that a lot variations in yields at higher data frequencies can likely not be attributed to these factors.

Because the decomposition of yields into contributions from expected future short-term interest rates and term premia depends on an accurate modeling of financial market participants' expectations, we follow Kim and Orphanides (2012) by making extensive use of survey expectations in both estimation and model evaluation. In particular, Kim and Orphanides provide evidence that, because of their high persistence, the physical dynamics of the yields are poorly estimated in samples of typical length, but at the same time are crucial for model implications. For example, if we were to estimate a VAR using actual inflation instead of decomposing it into trend inflation $\bar{\pi}_t$ and detrended inflation, the VAR would generate long-horizon inflation expectations, and thereby long-horizon expectations of short-term nominal interest rates, that are not nearly volatile enough over our sample (Kozicki and Tinsley, 2001). Using long-horizon survey expectations helps to disentangle transitory dynamics in inflation and real short-term rates from the secular movements as in (1)-(2), most notably the decline in long-horizon inflation expectations during the 1980s and 1990s. These two decompositions are motivated by the substantial fluctuations, shown in Figure 1, in expectations 7 to 11 years ahead of nominal 3-month Treasury bill yields, CPI inflation, and the expectations for real short-term interest rates implied by their difference.

In addition to the long-horizon expectatons for inflation and the 3-month yield, we use expectations at the 6- and 12-month horizons of the 3-month Treasury bill yield from the Blue Chip Financial Forecasts in the estimation, and impose that the VAR-implied expectations at the respective horizons are equal to these survey measures plus i.i.d. measurement errors. This assumption implies linear relationships of the form

$$E_t y_{1,t+k}^{svy} = \zeta_{y,k}' X_t + \epsilon_t^{y,k} \tag{9}$$

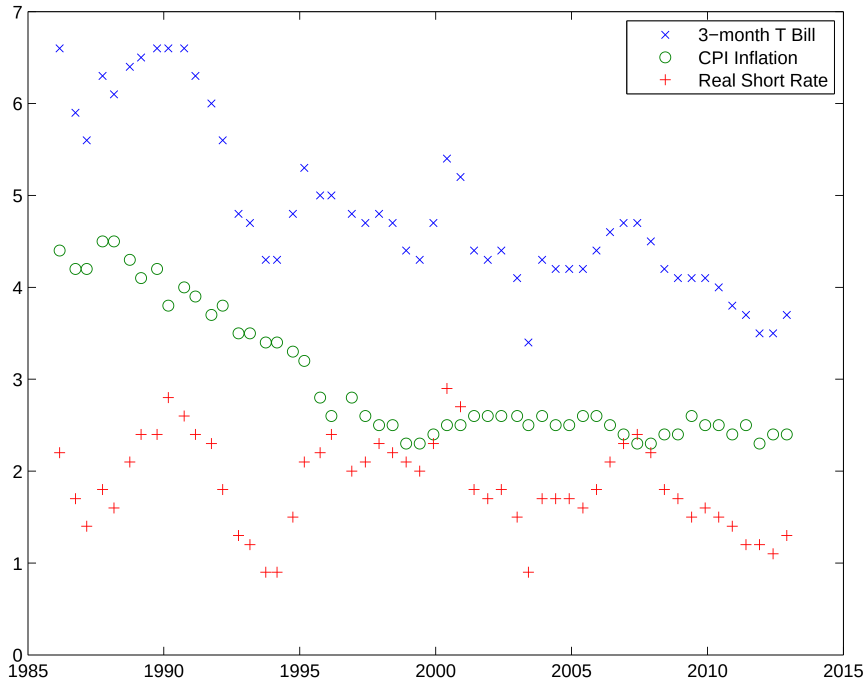
where $E_t y_{1,t+k}^{svy}$ denotes survey expectations of the 3-month T bill yield k periods ahead, and the coefficients $\zeta_{y,k}$ are functions of the VAR parameters.

A refinement of the decomposition (1) is motivated by the fact that CPI inflation, whether measured at monthly or quarterly frequency, is a very noisy process. Duffee (2011), Kim (2009), and Gürkaynak and Wright (2012) have argued that not all fluctuations in inflation are priced in bond yields. We therefore consider realized inflation π_{t+1} as composed of

$$\pi_{t+1} = \bar{\pi}_t + \tilde{\pi}_t + \epsilon_{t+1}^{\pi}$$

and that only $E_t \pi_{t+1} = \bar{\pi}_t + \tilde{\pi}_t$ affects bond pricing at date t, whereas the measurement

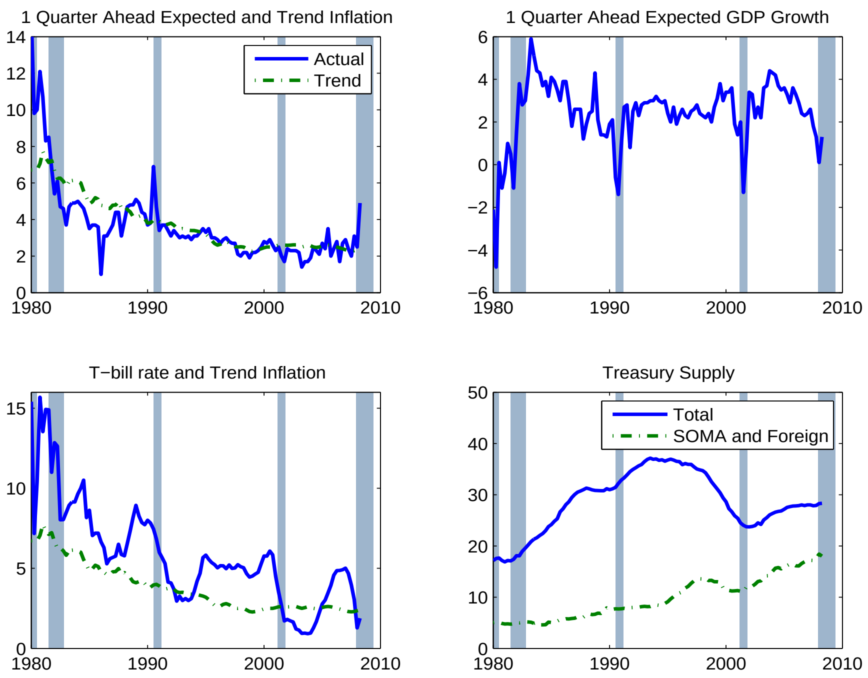
Figure 1: Long-Horizon Expectations of 3-month Yields and Inflation, Blue Chip



error ϵ_t^{π} does not. In particular, we use one-quarter-ahead survey forecasts of CPI inflation, denoted $E_t^{svy}\pi_{t+1}$ and assume that $E_t^{svy}\pi_{t+1} = \tilde{\pi}_t + \bar{\pi}_t$, thereby identifying $\tilde{\pi}_t$ conditional on an estimate of $\bar{\pi}_t$. Similarly, we use one-quarter ahead survey forecasts of real GDP growth as measure for real activity.

As will be discussed below, we seek to identify shocks that resemble the quantitative easing programs undertaken in recent years by several central banks. Usually, the main source of variation of the amount and duration of Treasury debt held by private investors is the Treasury itself. However, variations associated with fiscal policy actions affect macroeconomic variables importantly through changes in government spending and taxation. By contrast, variations in Treasury debt held by private investors engineered by central banks are not associated with fiscal measures and can therefore be expected to have different macroeconomic effects than variations due to fiscal policy. To allow us to disentangle these different sources of variations in Treasury debt held by private investors, we include two measures. The first, t_t, is the total amount of marketable debt outstanding, whereas the second, s_t, is the amount of Treasuries held by the Federal Reserve's System Open Mar-

Figure 2: Stationary variables in the state space



ket Account (SOMA) and foreign official institutions.[2] Both are expressed as percent of nominal GDP. The time series of the elements of x_t are shown in Figure 2. In light of the clear upward trend in s_t, we linearly detrend it before estimation of the VAR parameters. A fuller description of our state space model is provided in Appendix A, and further details on the data we use in Appendix B.

2.3 Identifying Treasury supply shocks

[Very rough and incomplete]

As discussed before, we would like to separately identify exogenous innovations to Treasury supply originating from fiscal policy on the one hand, and exogenous innovations to Federal Reserve and foreign holdings of Treasury securities on the other. Blanchard and Perotti (2002) proposed an identification strategy that takes account of endogenous responses of taxes and spending to output. Their approach to identification has been subject to a number of criticisms, one important of which is the fact that agents in the economy may be aware of the exogenous fiscal shocks before the econometrician observes them. However,

[2]The series of foreign official holdings is described in Beltran et al. (2012).

recent work by Leeper et al. (2012) and by Caldara and Kamps (2013) suggests that, in VARs that include more variables than Blanchard and Perotti's 3-equation model, adding measures of fiscal news doesn't qualitatively alter the results concerning dynamic responses to tax and spending shocks. Based on these results, we are adapting Blanchard and Perotti's identification strategy to our setting, in which we include not taxes and spending, but instead the total amount of marketable Treasury debt outstanding. In particular, we use the approximate relationship that

$$t_t = t_{t-1}/N_t + g_t - \tau_t$$

where g_t denotes the ratio of federal government expenditures to GDP, τ_t the ratio of federal tax revenues to GDP, and N_t the gross growth rate of nominal GDP between periods $t-1$ and t. This relationship allows us to convert idnetifying assumptions for tax and spending shocks separately into identifying assumptions for exogenous Treasury supply shocks induced by fiscal policy. Further details are presented in Appendix C.

The key challenge for identifying exogenous fiscal shocks is that there is clear evidence of contemporaneous causality running in both directions: Real revenues and spending are contemporaneously affected by changes in output and inflation because of the automatic stabilizers and lack of indexation of government wages, and output is contemporaneously affected by government spending and arguably by tax changes. By contrast, we assume that U.S. output and inflation are contemporaneously unaffected by exogenous changes in Federal Reserve and foreign official holdings of Treasury securities, for essentially the same reasons that most of the literature on identifying monetary policy shocks has assumed output and inflation to be contemporaneously unaffected by exogenous interest rate shocks. Hence, in addition to applying the Blanchard-Perotti identification strategy to fiscal shocks, we assume that exogenous innovations to the two monetary policy instruments, Federal Reserve and foreign official holdings and the short-term interest rate, do not contemporaneously affect any of the remaining variables in the VAR.

How to disentangle exogenous innovations to the two monetary policy instruments is a challenging question. For now, we assume a recursive ordering in which Federal Reserve and foreign official holdings of Treasury securities are chosen before the interest rate is determined, but we recognize that this is somewhat arbitrary. In future work we will want to explore alternative identifying assumptions in the spirit of Faust and Rogers (2003).

3 Estimation and Results

The model is estimated over the sample 1980q3 to 2008q2, using observations for 1980q1 and 1980q2 as initial lags. We start the sample only in 1980 because the systematic response of monetary policy to economic conditions is an important element of our factor VAR, and there is strong evidence for a break in this systematic component around 1980. We end the sample just before the intensification of the financial crisis in September 2008 because shortly thereafter the nominal short rate reached the zero lower bound, thereby introducing a nonlinearity in short-rate dynamics that our affine term structure model does not capture.[3] However, in discussing our results below, we will discuss the likely implications of our model for the effects of QE at the zero lower bound.

3.1 Estimation

We are estimating the model parameters jointly by maximum likelihood, treating the asymptotes of inflation $\bar{\pi}$ and the real one-period rate $\bar{r}$ as latent factors that are being filtered.[4] The total number of parameters to be estimated is rather large, and it is therefore important to use good starting values. We initialize the VAR parameters μ, ϕ_1, ϕ_2 and Ω with OLS estimates using proxies for $\bar{\pi}$ and $\bar{r}$. We impose in estimation the restriction that only current level risk is priced, i.e. only risk related to $\bar{\pi}_t$ and $\bar{r}_t$, but that the prices of these risks (the first two elements of λ_t) can be affected by all seven states. This restriction is consistent with the finding of Cochrane and Piazzesi (2008) that only level risk appears to be priced. Hence, the last five rows of λ_0 and λ_1 contain only zeros, reducing the number of elements of λ_0 and λ_1 to be estimated to 16.

To estimate the model, we use observations on zero-coupon Treasury yields with maturities 1, 2, 3, 7, 10, and 15 years from the data set described in Gürkaynak, Sack, and Wright (2006). Because the term structure model implies the exact linear relationships (8) between the states and the yields, with k yields and only two latent factors it is necessary to

[3] Li and Wei (2013) also end their sample in 2007, just before the onset of the financial crisis.

[4] An alternative approach would be to treat the long-horizon survey expectations as observations for the asymptotes, in which case we would estimate the parameters of the VAR for the factors by OLS, and only the risk price parameters λ_0 and λ_1 and the measurement error standard deviations for the yields by maximum likelihood. For taking this two-step approach, we would need to impute values for the long-horizon expectations of the T bill yield prior to 1986, the first year when the Blue Chip long-horizon expectations become available.

add measurement error to at least $k-2$ yields to avoid stochastic singularity. The n-period yield is therefore assumed to equal

$$y_{n,t} = a_n + b_n' X_t + \epsilon_t^n$$

.

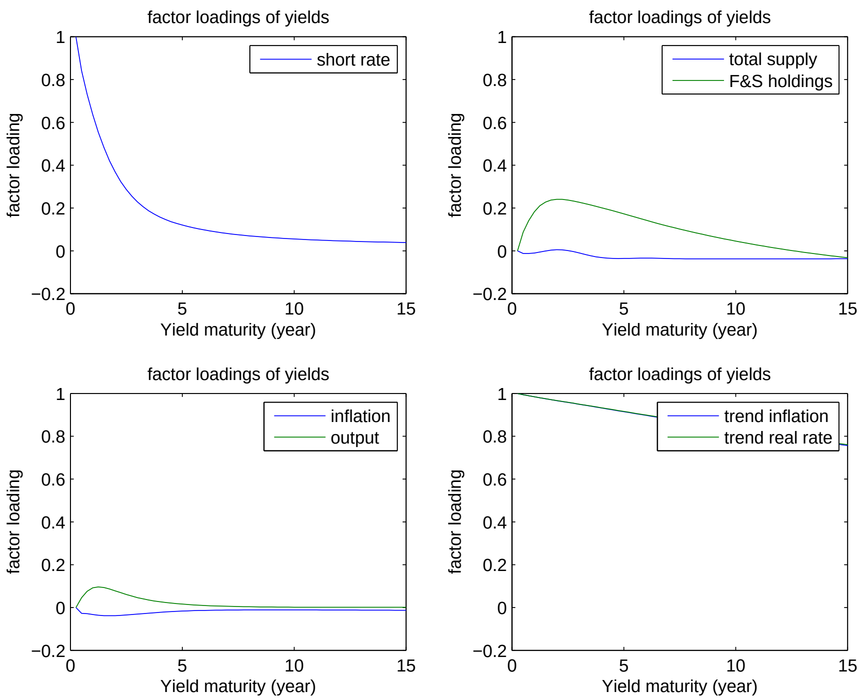
The model fit, as assessed by the standard deviations of the yield measurement errors, is quite impressive in comparison to other affine term structure models that use only observable factors. Note that, although the aymptotes of inflation and the real short rate are latent factors, we calibrate the measurement error standard deviations on the long-horizon survey expectations to 25 basis points, thus forcing the latent factors to closely follow the survey expectations. Of course, inclusion of the one-period yield among the factors implies that the shorter maturities are fit quite precisely, but the measurement error standard deviations out to maturit of 7 years are all close to 10 basis points. Only at the longest maturities does the standard deviation rise, to 20 basis points for the 10-year yield, and to 40 basis points for the 15-year yield.

Figure 3 presents the factor loadings of our estimated model, i.e. the contemporaneous responses of yields at various maturities to changes in the factors. These loadings lack a structural interpretation, as they do not distinguish between different sources for changes in the factors, and they are by construction static multipliers. Nonetheless they provide some insight into how the factors drive the yields. The upper left panel is consistent with the fact that the detrended short rate is by construction not highly persistent, and therefore its movements affect yields of longer maturities very little. The upper right shows that, whereas increases in total Treasury supply affects yields very little contemporaneously, increases in SOMA and foreign official holdings raise yields at intermediate maturities from 1 to 5 years upon impact. Of course, this does not imply that yields rise in response to exogenous innovations to foreign and SOMA holdings. Finally, the lower right panel shows that the asymptotes of inflation and the real short rate act as level factors.

3.2 Treasury supply shocks and the comovement of yields and macro variables

We now turn to some of the impulse responses to exogenous innovations to SOMA and foreign holdings, the short rate, and to total Treasury supply stemming from fiscal policy.

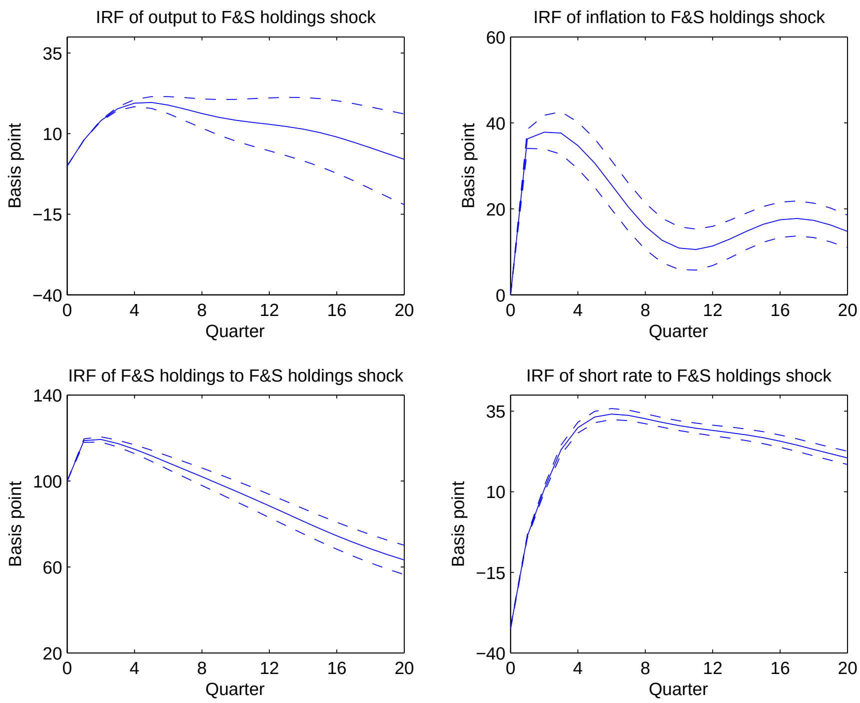
Figure 3: Factor loadings



We first focus on the responses of state variables to these shocks, and then on the responses of longer-term yields and term premia.

Figure 4 presents impulse responses to an exogenous increase in SOMA and foreign Treasury holdings in the amount of 1 percent of GDP (roughly \$150 billion at current levels). The upper left panel shows the response of the level of real GDP (the cumulative response of GDP growth). The level of real GDP rises gradually by close to 20 basis points over the four quarters following this innovation, and then slowly returns to its original level. The response of inflation (expressed at annual rate) is rapid in comparison to that in many VAR studies of the effects of monetary policy, but it should be kept in mind that this is the response of *detrended* inflation $\tilde{\pi}$, which shows less persistence than overall inflation (by assumption, trend inflation $\bar{\pi}$ does not respond to this shock, neither contemporaneously nor lagged). As shown in the lower left, the shock leads to a very persistent rise in SOMA and foreign Treasury holdings, even after detrending the series of s_t. In terms of the relationship between these holdings and the traditional short-rate tool of monetary policy, the lower right shows that the short rate declines by about 30 basis points upon impact, but then gradually

Figure 4: Impulse response functions to F&S shock

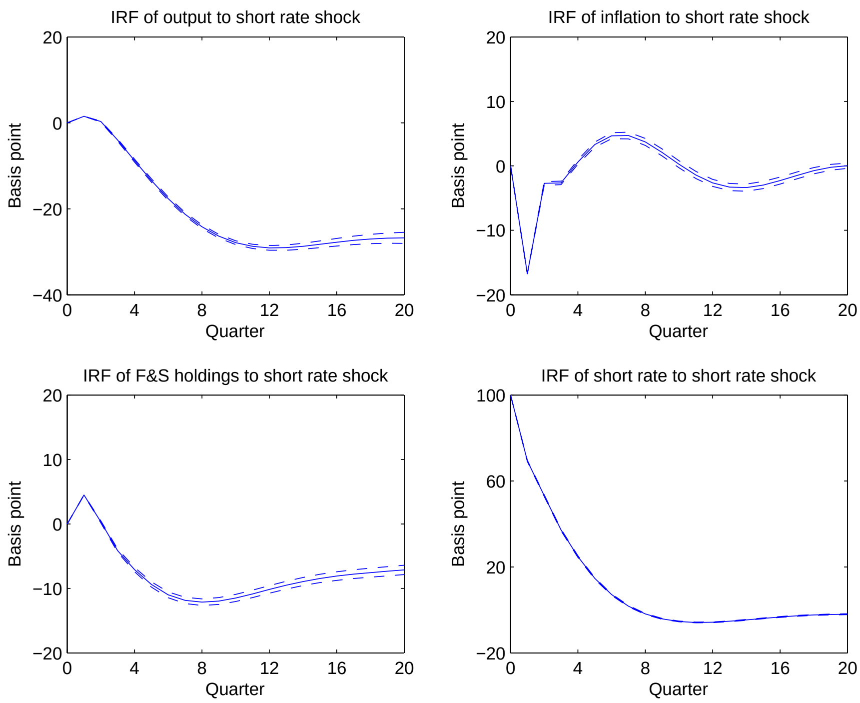


rises in response to the increase in output and inflation.

Figure 5 presents impulse responses to a monetary policy innovation of the traditional short-rate variety. In reponse to a 100 basis point increase (at annual rate) of the 3-month yield that dies out only gradually, the level of output declines by about 30 basis points over the 10 quarters following the shock before returning to its original level. The response of inflation is again quite rapid, reflecting likely the less persistent dynamics of detrended inflation. Finally, the lower left panel indicates that SOMA and foreign holdings act as complements to short-term interest rates by declining by about 10 basis points ($15 billion at current levels) during the first eight quarters.

The responses to a fiscal shock are presented in Figure 6. The increase in Treasury supply by 1 percent of GDP could reflect either an increase in spending or a decrease in taxes; since we only include total Treasury supply, we cannot distinguish between these two sources. Output rises upon impact by about 60 basis points and reaches a peak of 80 basis points within two quarters. Inflation jumps upon impact, but the response is short-lived. The increase in output and inflation leads to an initial rise in the 3-month T bill yield of

Figure 5: Impulse response functions to short-rate shock



about 100 basis points that dies out within 10 quarters.

Finally, in Figure 7 we show impulse responses of the 10-year Treasury yield and the associated term premium to the two shocks associated with monetary policy. As the bottom two panels show, the 10-year yield rises in response to an exogenous shock to the 3-month yield by about 6 basis points upon impact and then gradually returns to zero. This increase almost entirely reflects the "expectations hypothesis" effect stemming from persistently higher 3-month yields (as shown in Figure 5), whereas the term premium remains essentially unchanged. By contrast, an increase in SOMA and foreign holdings leads to a *rise* in the 10-year yield that peaks at about 6 basis after 4 quarters, but in this case the expectations hypothesis effect, which is substantially larger than in the case of the short-term rate shock, is largely offset by a persistent decline in the term premium by about 15 basis points. This estimate is qualitatively similar, but somewhat larger than the estimates reported in Li and Wei (2013).

Figure 6: Impulse response functions to fiscal shock

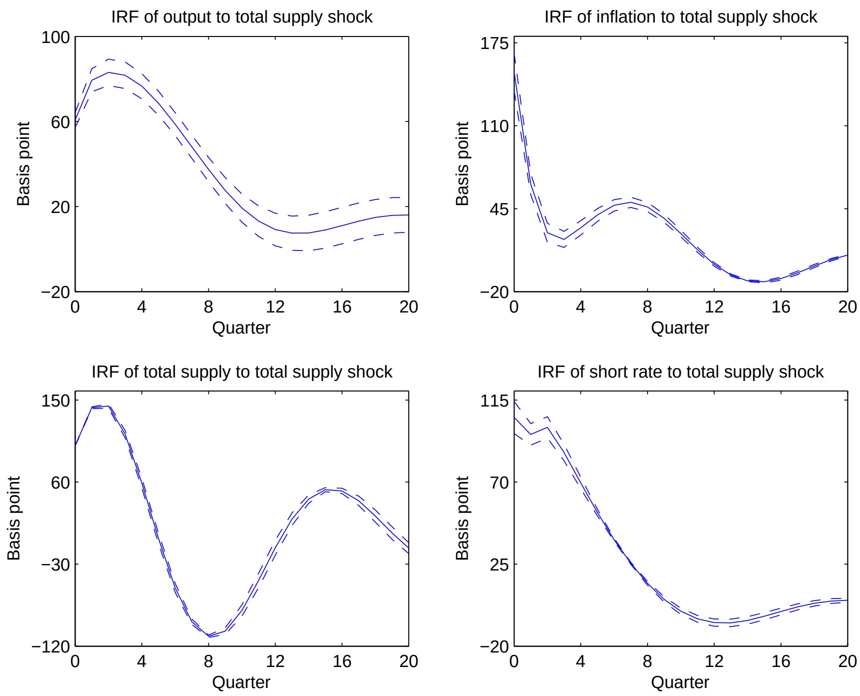
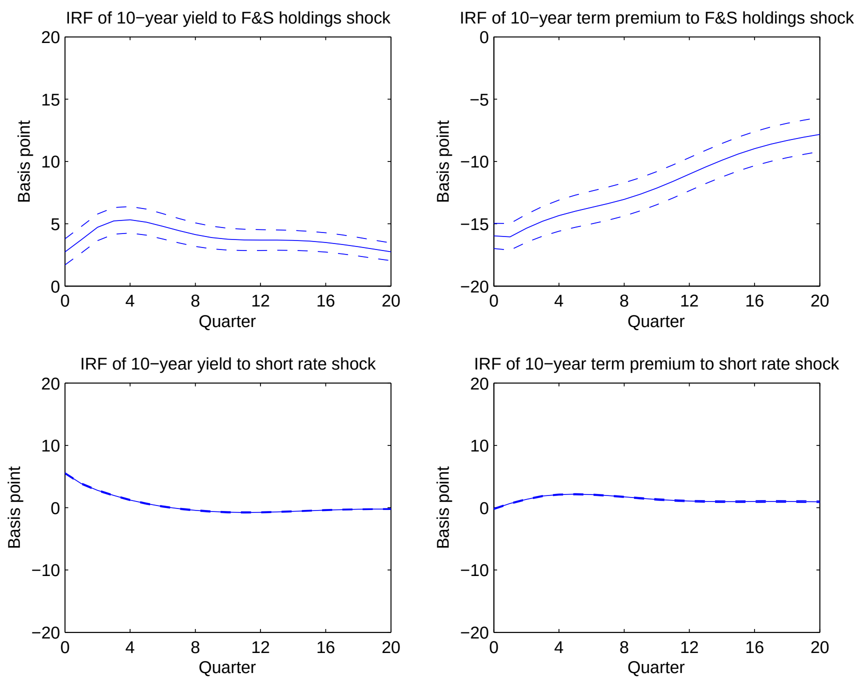


Figure 7: Impulse responses of 10-year yield and term premium



4 Conclusions

Still to be done:

- Include yield spread in VAR as in Ang-Piazzesi-Wei.
- Are identified F/S, fiscal shocks in accordance with narrative record?
- Identify exogenous changes to Treasury maturity composition from historical records. Combine SVAR and narrative approaches to identification in the manner of Stock and Watson (Brookings 2012), Mertens and Ravn (JME 2013).
- Revisit assumption that SOMA and foreign holdings don't respond contemporaneously to exogenous short-rate shocks.
- Simulate yields over period since 08Q2, decompose into contributions from SOMA purchases, forward guidance, fiscal.
- Examine robustness of results to restrictions on risk price parameters.
- One- vs. two-step estimation, calculation of standard errors.

References

[1] Ang, Andrew, and Monika Piazzesi, 2003. "A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables." *Journal of Monetary Economics* 50 (4), 745-787.

[2] Ang, Andrew, Monika Piazzesi, and Min Wei, 2006. "What does the yield curve tell us about GDP growth?" *Journal of Econometrics* 131 (1-2), 359-403.

[3] Beltran, Daniel, Maxwell Kretchmer, Jaime Marquez, and Charles Thomas, 2012. "Foreign holdings of U.S. Treasuries and U.S. Treasury yields." Federal Reserve Board, International Finance Discussion Paper 1041.

[4] Bernanke, Ben, Vincent Reinhart, and Brian Sack, 2004. "Monetary policy alternatives at the zero lower bound: An empirical assessment." *Brookings Papers on Economic Activity* 2, 1-100.

[5] Blanchard, Olivier, and Roberto Perotti, 2002. "An empirical investigation of the dynamic effects of shocks to government spending and taxes to output." *Quarterly Journal of Economics* 117, 1329-1368.

[6] Caldara, Dario, and Christophe Kamps, 2013. "The identification of fiscal multipliers and automatic stabilizers in SVARs." Working paper, Federal Reserve Board.

[7] Chen, Han, Vasco Cúrdia, and Andrea Ferrero, 2012. "The macroeconomic effects of large-scale asset purchase programs." *Economic Journal* 122, F289-F315.

[8] Chung, Hess, Jean-Philippe Laforte, David Reifschneider, and John Williams, 2012. "Have we underestimated the likelihood and severity of zero lower bound events?" *Journal of Money, Credit, and Banking* 44, 47-82.

[9] Dai, Qiang, and Thomas Philippon, 2006. "Fiscal policy and the term structure of interest rates." Working paper, New York University, November.

[10] Dai, Qiang and Kenneth Singleton, 2002. "Expectations puzzles, time-varying risk premia, and affine models of the term structure." *Journal of Financial Economics* 63, 415-441.

[11] D'Amico, Stefania, William English, David López-Salido, and Edward Nelson, 2012. "The Federal Reserve's large-scale asset purchase programs: Rationale and effects." Federal Reserve Board, Finance and Economics Discussion Series 2012-85.

[12] Duffee, Gregory, 2002. "Term premia and interest rate forecasts in affine models." *Journal of Finance* 57, 405-443.

[13] Duffee, Gregory, 2011. "Information in (and not in) the term structure." *Review of Financial Studies* 24, 2895-2934.

[14] Faust, Jon, and John Rogers, 2003. "Monetary policy's role in exchange rate behavior." *Journal of Monetary Economics* 50, 1403-1424.

[15] Favero, Carlo and Francesco Giavazzi, 2012. "Measuring tax multipliers: The narrative method in fiscal VARs." *American Economic Journal: Economic Policy* 4, 69-94.

[16] Gagnon, Joseph, Matthew Raskin, Julie Remache, and Brian Sack, 2011. " Large-scale asset purchases by the Federal Reserve: Did they work?" Federal Reserve Bank of New York *Economic Policy Review* 17, 41-59.

[17] Gürkaynak, Refet, Brian Sack, and Eric Swanson, 2005. "Do actions speak louder than words? The response of asset prices to monetary policy actions and statements." *International Journal of Central Banking* 1, 55-93.

[18] Gürkaynak, Refet, and Jonathan Wright, 2012. "Macroeconomics and the term structure." *Journal of Economic Literature*

[19] Joslin, Scott, Marcel Priebsch and Kenneth Singleton, 2013. "Risk premiums in dynamic term structure models with unspanned macro risks." *Journal of Finance*, forthcoming.

[20] Kiley, Michael, 2013a. "The response of equity prices to movements in long-term interest rates associated with monetary policy statements: Before and after the zero lower bound." Federal Reserve Board, Finance and Economics Discussion Series 2013-15.

[21] Kiley, Michael, 2013b. "Monetary policy statements, Treasury yields, and private yields: Before and after the zero lower bound." Federal Reserve Board, Finance and Economics Discussion Series 2013-16.

[22] Kim, Don, 2009. "Challenges in macro-finance modeling." Federal Reserve Bank of St. Louis *Review* 91 (5, Part 2), 519-544.

[23] Kim, Don and Athanasios Orphanides, 2012. "Term structure estimation with survey data on interest rate forecasts." *Journal of Financial and Quantitative Analysis* 47, 241-272.

[24] Kozicki, Sharon, and Peter Tinsley, 2001. "Shifting endpoints in the term structure of interest rates." *Journal of Monetary Economics* 47, 613-652.

[25] Krishnamurthy, Arvind, and Annette Vissing Jorgensen, 2011. "The effects of quantitative easing on interest rates: Channels and implications for policy." *Brookings Papers on Economic Activity* 42, 215-265.

[26] Li, Canlin, and Min Wei (2013). "Term Structure Modelling with Supply Factors and the Federal Reserve's Large Scale Asset Purchase Programs." *International Journal of Central Banking* 9, 3-39.

[27] Perotti, Roberto, 2004. "Estimating the Effects of Fiscal Policy in OECD Countries." Working paper, Universita' Bocconi, October.

[28] Rudebusch, Glenn, Brian Sack, and Eric Swanson, 2007. "Macroeconomic implications of changes in the term premium." Federal Reserve Bank of St. Louis *Review* July/August 2007, 241-269.

[29] Spencer, Peter, 2008. "Stochastic volatility in a macro-finance model of the U.S. term structure of interest rates 1961-2004." *Journal of Money, Credit and Banking* 40, 1177-1215.

[30] Vayanos, Dimitri, and Jean-Luc Vila, 2009. "A preferred-habitat model of the term structure of interest rates." NBER working paper No. 15487.

A Specification of the affine term structure model

The term structure model is given by the law of motion for the vector of factors, which is assumed to be a VAR(2) under the historical probability measure,

$$X_t = \mu + \Phi X_{t-1} + V_t,\ t = 1, \ldots, T$$

where the vectors X_t and V_t and the matrix Φ are defined as in (5) in the main text, by the affine market price of risk specification (6), the short-rate specification (7), and the stochastic discount factor. This is an "essentially affine" term structure model as defined in Duffee (2002). Given the normality of V_{t+1}, the SDF

$$M_{t+1} = \exp\left(-r_t - \frac{1}{2}\lambda_t'\Omega\lambda_t - \lambda_t' U_{t+1}\right)$$

is log-normal. Under these assumptions it is possible to express the price of the n-period bond as

$$P_{nt} = \exp(A_n + B_n' X_t) \tag{10}$$

where the scalars A_n and vectors B_n evolve recursively as

$$A_{n+1} = A_n + B_n'(\mu - \Sigma\lambda_0) + \frac{1}{2}B_n'\Sigma\Sigma' B_n - \delta_0 \tag{11}$$

$$B_{n+1}' = B_n'(\Phi - \Sigma\lambda_1) - \delta_1' \tag{12}$$

with $A_1 = -\delta_0,\ B_1 = -\delta_1$, and Σ defined as the square root of the covariance matrix Ω of V_t. Expression (10) then leads to (8) with $a_n = -A_n/n$ and $b_n = -B_n/n$.

A.1 The state space model with survey information

The model is specified at the quarterly frequency. In the estimation of the model, we use yields of maturities 1, 2, 3, 7, 10, and 15 years in addition to the 3-month T bill yield. The vector of observables Y_t therefore consists of

$$Y_t = [y_t^1,\ y_t^4,\ y_t^8,\ y_t^{12},\ y_t^{28},\ y_t^{40},\ y_t^{60},\ \bar{\pi}_t^{svy},\ \bar{r}_t^{svy},\ E_t^{svy}[\pi_{t+1}],\ E_t^{svy}[q_{t+1}],\ t_t,\ s_t,\ E_t^{svy}[y_{1,t+2}],\ E_t^{svy}[y_{1,t+4}]]'$$

We treat the survey expectations $\bar{\pi}_t^{svy}$ and $\bar{r}_t^{svy}$ as if they had a constant forecast horizon, i.e. as if they represented expectations of average inflation over the horizon 25 to 44 quarters ahead (details on the actual series are provided below). The model-implied expectation of average inflation over this horizon is then given by $\zeta_\pi' X_t$ with

$$\zeta_\pi = 0.05(\iota_1 + \iota_3)\Phi^{25}(\sum_{i=0}^{19}\Phi^i)$$

where ι_j selects the j-th element of X_t.

With these assumptions, the state space model consists of the transition equation given by (5) and a measurement equation

$$Y_t = A + BX_t + e_t$$

given by

$$
\begin{bmatrix} y_{1,t} \\ y_{m_1,t} \\ \vdots \\ y_{m_N,t} \\ E_t^{svy}[\pi_{\text{long}}] \\ E_t^{svy}[y_{1,\text{long}}] \\ E_t^{svy}[\pi_{t+1}] \\ E_t^{svy}[q_{t+1}] \\ t_t \\ s_t \\ E_t^{svy}[y_{1,t+2}] \\ E_t^{svy}[y_{1,t+4}] \end{bmatrix} = \begin{bmatrix} 0 \\ \\ A_y \\ \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ & & & & & & \\ & & & B_y & & & \\ & & & & & & \\ & & & \zeta_{\bar{\pi}}' & & & \\ & & & \zeta_{\bar{r}}' & & & \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ & & & \zeta_y^{2\prime} & & & \\ & & & \zeta_y^{4\prime} & & & \end{bmatrix} \begin{bmatrix} \bar{\pi}_t \\ \bar{r}_t \\ \tilde{\pi}_t \\ q_t \\ t_t \\ s_t \\ \tilde{y}_{1,t} \end{bmatrix} + \begin{bmatrix} 0 \\ \epsilon_t^{m_1} \\ \vdots \\ \epsilon_t^{m_N} \\ \epsilon_t^{\bar{\pi}} \\ \epsilon_t^{\bar{r}} \\ 0 \\ 0 \\ 0 \\ 0 \\ \epsilon_t^{y,2} \\ \epsilon_t^{y,4} \end{bmatrix} \tag{13}
$$

where $m_1, \ldots, m_N$ denotes the maturities of Treasury securities included in estimation, A_y and B_y the vector and matrix of the stacked corresponding a_n and b_n from (8), $E_t^{svy}[\pi_{\text{long}}]$ and $E_t^{svy}[y_{1,\text{long}}]$ denote the Blue Chip forecasts of CPI inflation and the 3-month T bill yield at the longest horizon (the projected average over the horizon roughly 7 to 11 years ahead), and $E_t^{svy}[y_{1,t+2}]$ and $E_t^{svy}[y_{1,t+4}]$ are the Blue Chip forecasts of the 3-month T bill yield 2 and 4 quarters ahead. The corresponding measurement errors are denoted by ϵ.

B The data

[To be completed]

The series of inflation expectations consists of three different pieces. Until 1981:Q1, the series is an estimated step function based on the changepoint model developed in Kozicki and Tinsley (2001). From 1981:Q2 until 1988:Q4, the series is based on the Hoey survey of bond market participants, which was conducted on a quarterly basis by Richard Hoey, an economist at Drexel Burnham Lambert. Participants in this survey were polled for their expectation of CPI inflation over the second five years of a 10-year horizon. From 1989:Q4 onwards the series is based on the expectations for the average CPI inflation rate roughly 7 to 11 years ahead from the Blue Chip Financial Forecasts. Because the Financial Forecasts poll respondents only twice a year for their long-horizon forecasts, we interpolate the data to quarterly frequency. For the years from 1986 on, we use the long-horizon forecasts of the 3-month T bill yield from the Financial Forecasts, and likewise interpolate to quarterly frequency. Before 1986 no such long-horizon forecasts are available, and we treat them as missing observations in the estimation.

C Identification and estimation of the VAR

Following the notation used in (4), let v_t denote the vector of reduced-form residuals, and ε_t the vector of structural innovations of which we seek to identify several elements. As stated in the main text, we assume that the asymptote of inflation $\bar{\pi}_t$ and of the one-period real rate $\bar{r}_t$ follow univariate random walks with innovations $\varepsilon_t^{\bar{\pi}}$ and $\varepsilon_t^{\bar{r}}$.

We are interested in identifying the structural shocks to total Treasury supply, including privately-held Treasuries, and those to Treasury holdings by the SOMA and foreign official institutions. Following Blanchard and Perotti (2002), we assume that fiscal policy cannot respond contemporaneously to macroeconomic developments except by the automatic stabilizers embedded in the tax and spending policies in place. Hence, the reduced-form innovations to total Treasury supply are composed of a response to current shocks to economic activity q and inflation π as implied by the automatic stabilizers, and any exogenous fiscal policy shocks that are unrelated to current macroeconomic conditions. Note in particular that total Treasury supply is assumed to be contemporaneously unaffected by monetary policy, be that SOMA asset holdings or the one-period interest rate (where we follow the convention of assuming no contemporaneous response of real activity and inflation to innovations to $y_{1,t}$). With these assumptions, the contemporaneous relationship between the reduced-form innovations v_t^t, v_t^q, and v_t^π to Treasury supply, real activity, and inflation respectively, and the structural fiscal (Treasury supply) shock ε_t^t is

$$v_t^t = \eta^{t,\pi} v_t^\pi + \eta^{t,q} v_t^q + \varepsilon_t^t$$

where the coefficients $\eta^{t,x}$ can be constructed as $\eta^{t,x} = \eta^{\tau,x} - \eta^{g,x}$ from the underlying calibrated parameters in the equations for log real taxes τ and log real spending g

$$\begin{aligned} v_t^\tau &= \eta^{\tau,\pi} v_t^\pi + \eta^{\tau,q} v_t^q + \varepsilon_t^\tau \\ v_t^g &= \eta^{g,\pi} v_t^\pi + \eta^{g,q} v_t^q + \varepsilon_t^g \end{aligned}$$

Constructing the coefficients $\eta^{t,\pi}$ and $\eta^{t,q}$ in the manner of Blanchard and Perotti (2002) is critical because inflation and real activity are assumed to be contemporaneously affected by fiscal policy and hence by ε_t^t.

By contrast, SOMA (and foreign official) Treasury holdings f_t are also assumed to respond contemporaneously to real activity and inflation, whereas they are not assumed to affect real activity and inflation, in analogy to the conventional assumption in the literature that monetary policy shocks (in the form of structural shocks to the short-term interest rate) do not affect these variables contemporaneously. The relationships between the reduced-form residuals v_t and the structural innovations ε_t can thus be written as $v_t = \boldsymbol{\eta}\varepsilon_t$ with

$$\begin{bmatrix} v_t^{\bar{\pi}} \\ v_t^{\bar{r}} \\ v_t^{\pi} \\ v_t^{q} \\ v_t^{t} \\ v_t^{s} \\ v_t^{y} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ ? & ? & 1 & 0 & ? & 0 & 0 \\ ? & ? & ? & 1 & ? & 0 & 0 \\ ? & ? & \eta^{t,\pi} & \eta^{t,q} & 1 & 0 & 0 \\ ? & ? & ? & ? & 0 & 1 & 0 \\ ? & ? & ? & ? & ? & ? & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_t^{\bar{\pi}} \\ \varepsilon_t^{\bar{r}} \\ \varepsilon_t^{\pi} \\ \varepsilon_t^{q} \\ \varepsilon_t^{t} \\ \varepsilon_t^{s} \\ \varepsilon_t^{y} \end{bmatrix} \tag{14}$$

The parameters $\eta^{t,\pi}$ and $\eta^{t,q}$ are calibrated based on the values for these parameters reported in Perotti (2004). The parameters denoted with "?" are estimated by instrumental variables. Specifically, the structural residuals $\varepsilon_t^{\bar{\pi}}$ and $\varepsilon_t^{\bar{r}}$ are simply the first differences of the series $\bar{\pi}_t$ and $\bar{r}_t$; the unknown parameters in the third row of the matrix are estimated by regressing v_t^{π} on $\varepsilon_t^{\bar{\pi}}$, $\varepsilon_t^{\bar{r}}$, and v_t^t,, using ε_t^t as instrument etc.