

News Shocks and the Slope of the Term Structure of Interest Rates

ONLINE APPENDIX*

André Kurmann
Federal Reserve Board

Christopher Otrok
University of Missouri

Federal Reserve Bank of St. Louis

April 5, 2012

This appendix reports details on

- VAR identification methodology
- Baseline VAR results reported in the paper
- Contemporaneous TFP shock identification
- Robustness checks for baseline VAR
 - with respect to OLS estimation
 - with respect to alternative variable definitions
- Larger VAR results reported in the paper
- The fit of interest rate rule with VAR results

*This appendix accompanies the paper with the same title and is not intended for publication. The results in this appendix do not necessarily represent the views of the Federal Reserve System or the Federal Open Market Committee.

A Identifying Structural Shocks: Two VAR Approaches

Here we present details on the two approaches to VAR identification in the paper. The first approach, proposed by Uhlig (2003), is purely statistical and extracts the largest 1 or 2 (or 3 or 4) shocks that explain the maximal amount of the forecast error variance (FEV) in a target variable, which in our case is the slope of the term structure. The second identification approach is motivated by a key result from the first identification: news about future TFP play an important role in explaining movements in the slope. To assess this interpretation formally, we follow Barsky and Sims (2010) and extend the FEV maximization approach of Uhlig (2003) by using TFP as the target variable and imposing the extra restriction that the identified shock is orthogonal to contemporaneous TFP.

A.1 Review of VAR basics

Consider a reduced-form VAR of the form

$$Y_t = B_1 Y_{t-1} + B_2 Y_{t-2} + \dots + B_p Y_{t-p} + u_t, \quad (1)$$

where Y_t is a $m \times 1$ vector of variables observed at time t ; and u_t is a $m \times 1$ vector of one-step-ahead prediction errors with variance-covariance matrix $E[u_t u_t'] = \Sigma$. Constant terms are dropped to save on notation. The vector moving average representation of this reduced-form VAR is

$$Y_t = [B(L)]^{-1} u_t = C(L) u_t, \quad (2)$$

where $B(L) \equiv I - B_1 L - \dots - B_p L^p$, and $C(L) \equiv I + C_1 L + C_2 L^2 + \dots$

Identification of the structural shocks amounts to finding a mapping A between the prediction errors u_t and a vector of mutually orthogonal shocks ε_t ; i.e. $u_t = A\varepsilon_t$. The key restriction on A is that it needs to satisfy $\Sigma = E[A\varepsilon_t \varepsilon_t' A'] = AA'$. This restriction is, however, not sufficient to identify A because for any matrix A , there exists some alternative matrix \tilde{A} such that $\tilde{A}Q = A$, where Q is an orthonormal matrix, that also satisfies $\Sigma = \tilde{A}\tilde{A}'$.¹ This alternative matrix maps u_t into another vector of mutually orthogonal shocks $\tilde{\varepsilon}_t$; i.e. $u_t = \tilde{A}\tilde{\varepsilon}_t$. For some arbitrary matrix \tilde{A} satisfying

¹To see this, consider an orthonormal matrix Q (i.e. $QQ' = I$) and define $\tilde{A}Q = A$ such that $Q\varepsilon_t = \tilde{\varepsilon}_t$. Then, $\Sigma = E[A\varepsilon_t \varepsilon_t' A'] = E[\tilde{A}Q\varepsilon_t \varepsilon_t' Q' \tilde{A}'] = E[\tilde{A}\tilde{\varepsilon}_t \tilde{\varepsilon}_t' \tilde{A}'] = \tilde{A}\tilde{A}'$ because $E[\tilde{\varepsilon}_t \tilde{\varepsilon}_t'] = QE[\varepsilon_t \varepsilon_t']Q' = QQ' = I$.

$\Sigma = \tilde{A}\tilde{A}'$ (e.g. the Cholesky decomposition of Σ), identification therefore reduces to choosing an orthonormal matrix Q .

A.2 Extracting the most important shocks

Uhlig's (2003) statistical approach consists of finding the $n < m$ columns of Q defining the n mutually orthogonal shocks that explain most of the FEV of some variable in Y_t over forecast horizon \underline{k} to \bar{k} . Formally, denote the k -step ahead forecast error of the i -th variable $y_{i,t}$ in Y_t by

$$y_{i,t+k} - E_t y_{i,t+k} = e_i' \left[\sum_{l=0}^{k-1} C_l \tilde{A} Q \varepsilon_{t+k-l} \right], \quad (3)$$

where e_i is a column vector with 1 in the i -th position and zeros elsewhere. Then Uhlig's (2003) approach solves

$$Q_n^* = \arg \max_{Q_n} e_i' \left[\sum_{k=\underline{k}}^{\bar{k}} \sum_{l=0}^{k-1} C_l \tilde{A} Q_n Q_n' \tilde{A}' C_l' \right] e_i \quad (4)$$

subject to $Q_n' Q_n = I$, where Q_n contains the columns of Q defining the n most important shocks.

To implement this problem, consider first finding the shock – i.e. the column q_1 of Q – that explains most of the FEV of variable y_i ²

$$q_1^* = \arg \max_{q_1} e_i' \left[\sum_{k=\underline{k}}^{\bar{k}} \sum_{l=0}^{k-1} C_l \tilde{A} q_1 q_1' \tilde{A}' C_l' \right] e_i \quad (5)$$

subject to $q_1' q_1 = 1$. The objective to be maximized can be expressed as

$$\begin{aligned} e_i' \left[\sum_{k=\underline{k}}^{\bar{k}} \sum_{l=0}^{k-1} C_l \tilde{A} q_1 q_1' \tilde{A}' C_l' \right] e_i &= \sum_{k=\underline{k}}^{\bar{k}} \sum_{l=0}^k \text{trace} \left[(e_i e_i') (C_l \tilde{A} q_1) (q_1' \tilde{A}' C_l') \right] \\ &= \sum_{k=\underline{k}}^{\bar{k}} \sum_{l=0}^k \text{trace} \left[(q_1' \tilde{A}' C_l') (e_i e_i') (C_l \tilde{A} q_1) \right] \\ &= q_1' \left[\sum_{k=\underline{k}}^{\bar{k}} \sum_{l=0}^k \tilde{A}' C_l' (e_i e_i') C_l \tilde{A} \right] q_1 \\ &= q_1' S q_1, \end{aligned}$$

²Without loss of generality, we order this vector first in Q .

with

$$S \equiv \sum_{k=\bar{k}}^{\bar{k}} \sum_{l=0}^k \tilde{A}' C'_l (e_i e'_i) C_l \tilde{A},$$

where the formulation in the second line takes advantage of the fact that for any three square matrices D , E , F of same dimension $trace(DEF) = trace(FDE)$. The maximization problem in (5) can therefore be expressed as a Lagrangian

$$L = q'_1 S q_1 - \lambda (q'_1 q_1 - 1) \tag{6}$$

with first-order condition

$$S q_1 = \lambda q_1.$$

Inspection of this solution reveals that this is simply the definition of an eigenvalue decomposition, with q_1 being the eigenvector of S that corresponds to eigenvalue λ . Furthermore, since $q'_1 q_1 = 1$, we can rewrite the first-order condition as $\lambda = q'_1 S q_1$; i.e. eigenvalue λ is the objective to be maximized. The partition q_1 that maximizes the variance is therefore the eigenvector associated with the largest eigenvalue λ . Likewise, q_2 is the second principal component and so forth for all the n components of Q_n that we want to extract.³

A.3 Identifying TFP News Shocks

In the second part of our empirical analysis, we identify a news shock about future innovations to TFP. As in Barsky and Sims (2010), we assume that TFP is an exogenous process driven by two orthogonal innovations: a traditional technology shock that affects TFP contemporaneously; and a news shock that is revealed today but impacts TFP only in the future. More formally, let the logarithm of TFP be defined by a moving average process of the form

$$\log TFP_t = v(L) \varepsilon_t^{current} + d(L) \varepsilon_t^{news}, \tag{7}$$

³As explained by Uhlig (2003), there are two other pairs (or rotations) of eigenvectors that explain together an equal fraction of the total FEV of the slope as the two eigenvectors associated with the two largest eigenvalues of S . We compared all of our results with these two other rotations and found that for each rotation, there exists one shock that explains over 50% of slope movements. This shock has very similar properties than the first shock we consider in the text.

where $\varepsilon_t^{current}$ and ε_t^{news} are uncorrelated innovations; and $v(L)$ and $d(L)$ are lag polynomials with the only restriction that $d(0) = 0$. This restriction defines $\varepsilon_t^{current}$ as the contemporaneous TFP shock and ε_t^{news} as the TFP news shock; i.e. while $\varepsilon_t^{current}$ can affect TFP immediately in period t , ε_t^{news} can affect TFP only in period $t + 1$ and later.

In a VAR with (an exogenous measure of) TFP ordered first, specification (7) with $d(0) = 0$ implies that the contemporaneous TFP shock $\varepsilon_t^{current}$ is identified as the shock associated with the first column of the matrix \tilde{A} obtained from the Cholesky decomposition.⁴ The news shock ε_t^{news} then necessarily covers all movements in TFP that are not related to $\varepsilon_t^{current}$. This restriction cannot be imposed to hold at all horizons. Instead, Barsky and Sims (2010) propose to identify ε_t^{news} as the innovation that explains most of future movements in TFP but nothing of current TFP. This is a restricted version of Uhlig’s (2003) approach presented above but now applied to TFP rather than the slope. It amounts to solving the Lagrangian problem in (6) subject to the side-constraint that the identified shock has no contemporaneous effect; i.e. that the first element of the extracted eigenvector q is zero.

A.4 Comparison of Uhlig’s approach applied to spread and long bond rate

An important question is whether Uhlig’s identification yields similar results independent of the target variable on which the approach is applied. To show that this is not the case, we applied Uhlig’s approach on other variables in the baseline VAR. Generally, the extracted shocks from these alternative identification is very weakly correlated with our slope shock (and therefore with the TFP news shock). As an illustration, we show here results for the case where we applied Uhlig’s approach to the long bond rate. This is an especially interesting alternative target variable because it is part of the definition of the spread. As can be seen from the below figures, the results are very

⁴To see this, recall that \tilde{A} obtained from the Cholesky decomposition is a lower-triangular matrix. Hence, the only shock in $\tilde{\varepsilon}_t$ associated with \tilde{A} that can have an immediate effect on the first variable in Y_t (i.e. TFP) is the first element in $\tilde{\varepsilon}_t$.

different from the ones obtained for the slope.

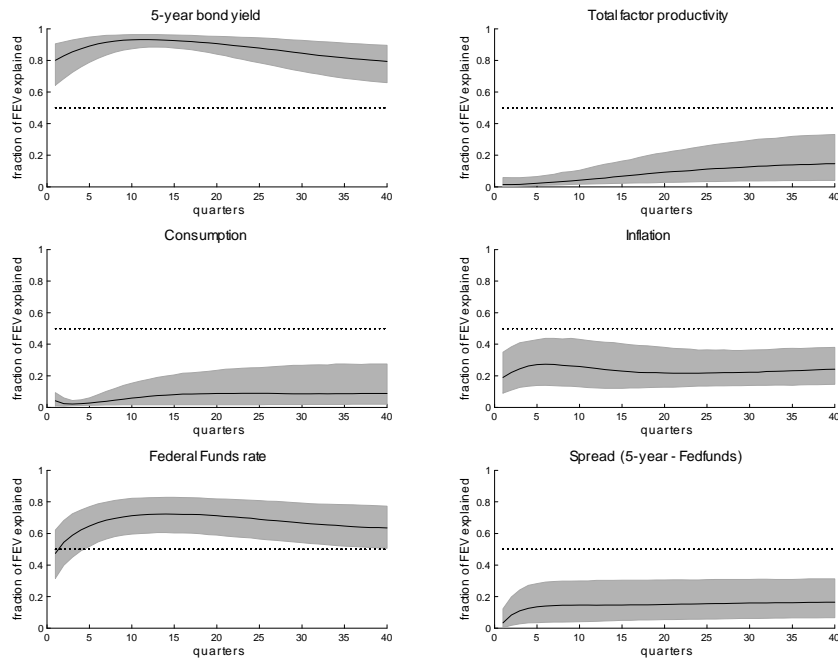


Fig. A.1: Fraction of FEV explained by level shock

The first figure shows that the extracted 'long-bond shock' accounts for 75% or more of the FEV of the long bond yield but only very little of TFP and the spread. Furthermore, the correlation between the extracted long-bond shock and the slope shock of the paper is 0.237 for the full sample; 0.49 for the pre-84 sample; and -0.44 for the post-84 sample (note the negative sign!). In other words, the shock that explains most of the FEV of the long-bond yield is very different from the slope shock (and therefore the TFP news shock). Likewise the IRFs to the level shock (shown below in solid

black lines) are very different from the IRFs to our slope shock (shown in dashed red lines).

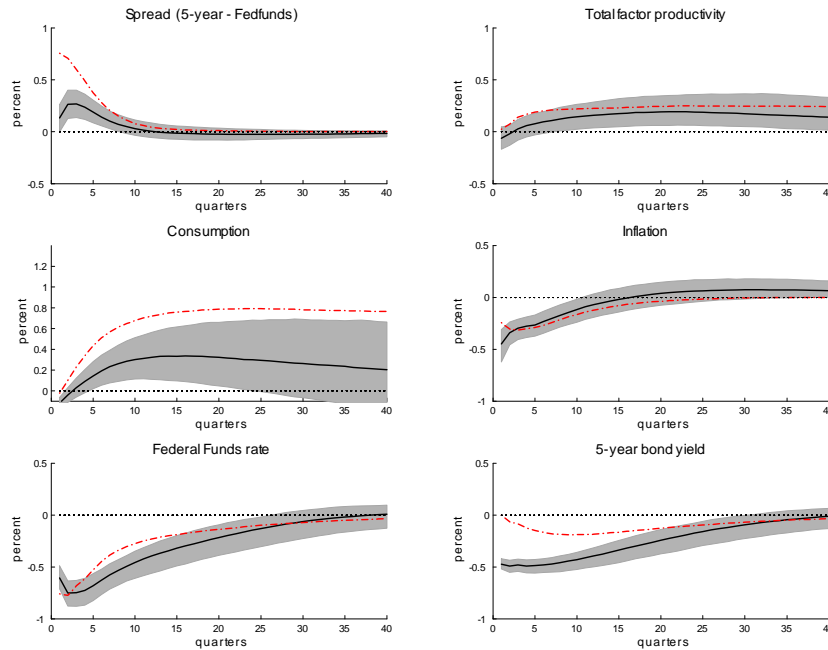


Fig. B.1.2: Impulse responses to 1% level shock (solid black lines and grey confidence intervals) and impulse responses to 1% slope shock (red dashed lines)

In contrast to the reaction to the slope shock, the spread hardly moves on impact of the long-bond shock. Conversely, the long-bond yield jumps down by almost 50 basis points whereas the long-bond yield barely reacts in response to the slope shock. The results suggest that the long-bond shock is akin to what the finance literature has identified as a 'level shock'; i.e. a shock that moves both the long and the short end of the term structure roughly equally. The long-bond shock also has a sizable impact on inflation but moves TFP and consumption relatively little. In particular, both TFP and consumption return back towards their initial level suggesting that the extracted long-bond shock does not relate to a permanent technology improvement. All of these results confirm that the extracted level shock looks like a very different shock than the slope shock (and therefore the TFP news shock) identified in the paper.

B Details on baseline VAR results reported in the paper

Here we report details for the baseline VAR of the paper. For both the slope identification and the TFP news identification, we show the full set of variance decompositions and impulse responses with corresponding confidence intervals. We also show a decomposition of the spread's impulse response into an expectations hypothesis part and a term premia part. This decomposition takes the spread between the nominal yield i_t^T on a T -period zero-coupon bond (in our case the 5-year treasury bond) and nominal yield i_t on a 1-period bill (in our case, the Federal Funds rate) and expresses it as the sum of two parts

$$i_t^T - i_t = \frac{1}{T} \sum_{i=0}^{T-1} \left[\left(1 - \frac{i}{T} \right) E[\Delta i_{t+i} | Y_t] \right] + tp_t,$$

where the $E[\Delta i_{t+i} | Y_t]$ denote expectations of future short rates as implied by the VAR based on information Y_t ; and tp_t denotes term premia. This type of decomposition is standard in the term structure literature. Notable examples are the seminal paper by Campbell and Shiller (1987) or more recently Diebold, Rudebusch and Aruoba (2006) and Evans and Marshall (2007).

B.1 Baseline VAR results for slope identification

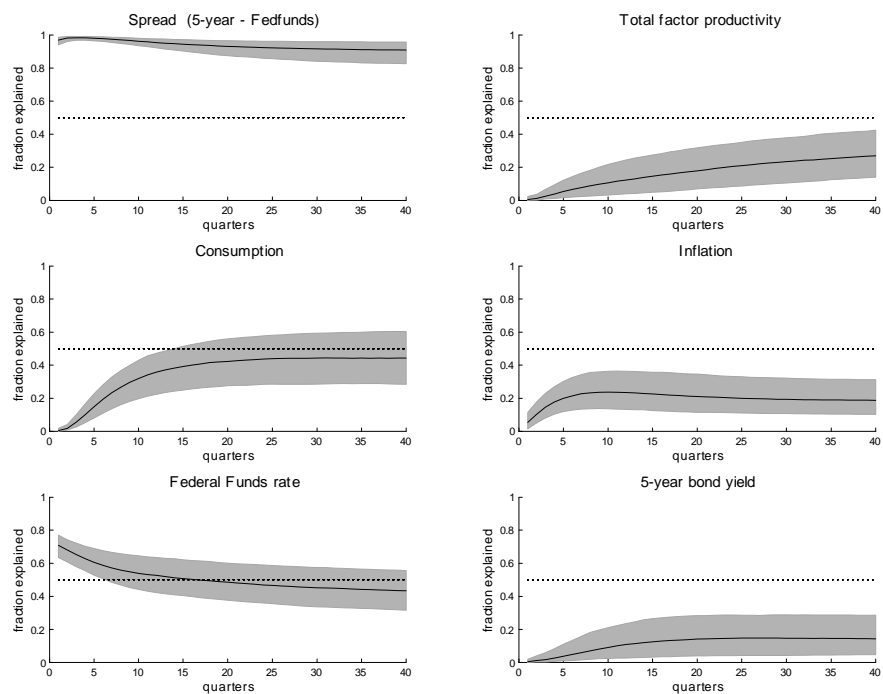


Fig. B.1.1: Fraction of FEV explained by slope shock

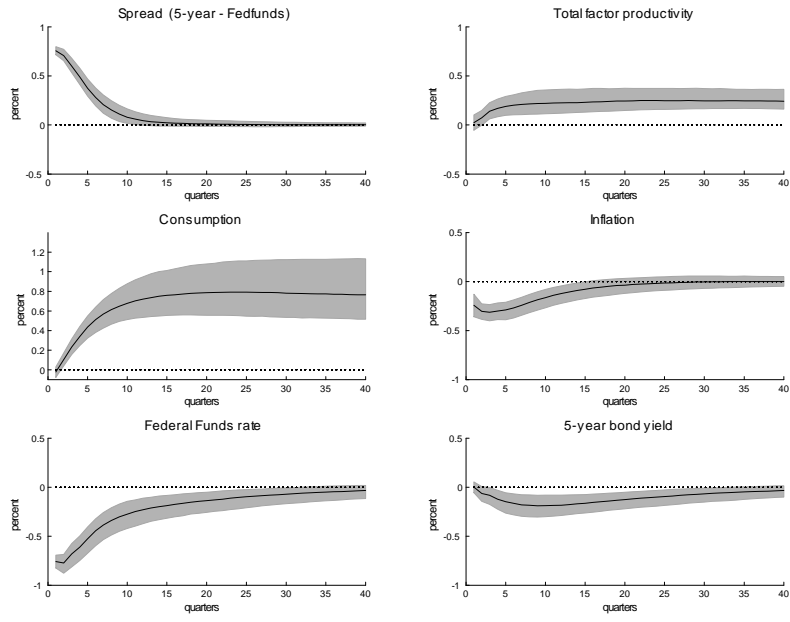


Fig. B.1.2: Impulse responses to 1% slope shock

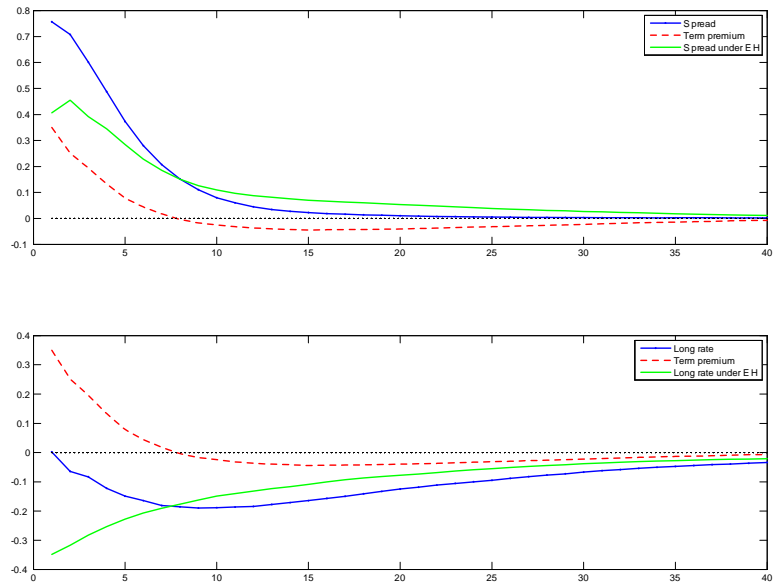


Fig. B.1.3: Decomposition of spread and long-rate response to slope shock into Expectations Hypothesis part and term premia part

B.2 Baseline VAR results for TFP news shock identification

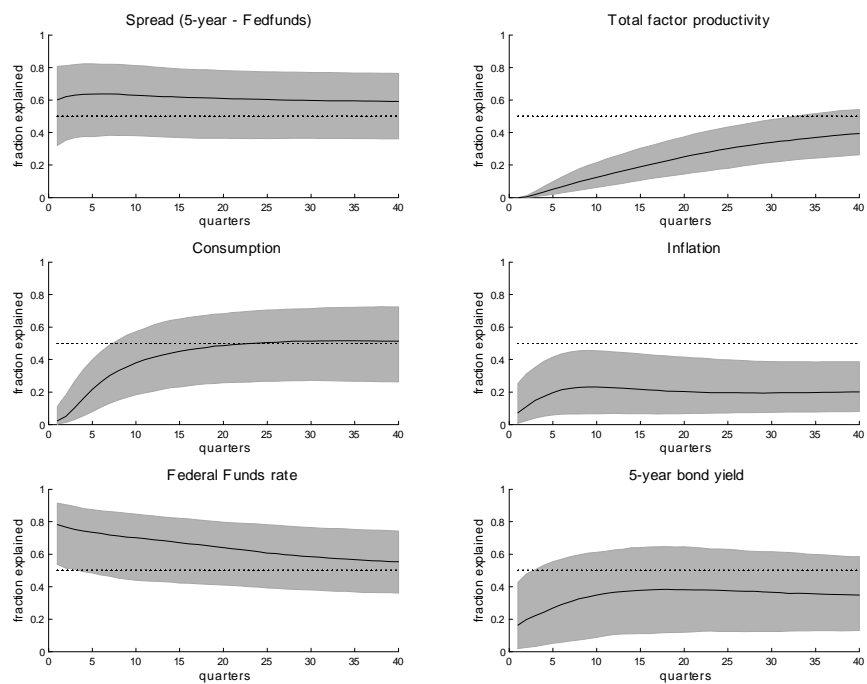


Fig. B.2.1: Fraction of FEV explained by TFP news shock

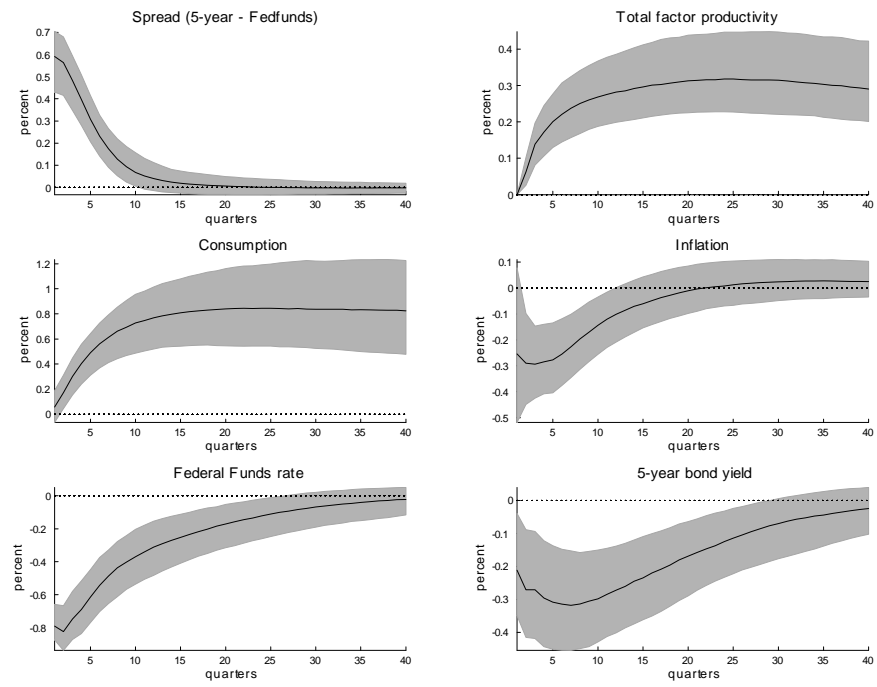


Fig. B.2.2: Impulse responses to 1% TFP news shock

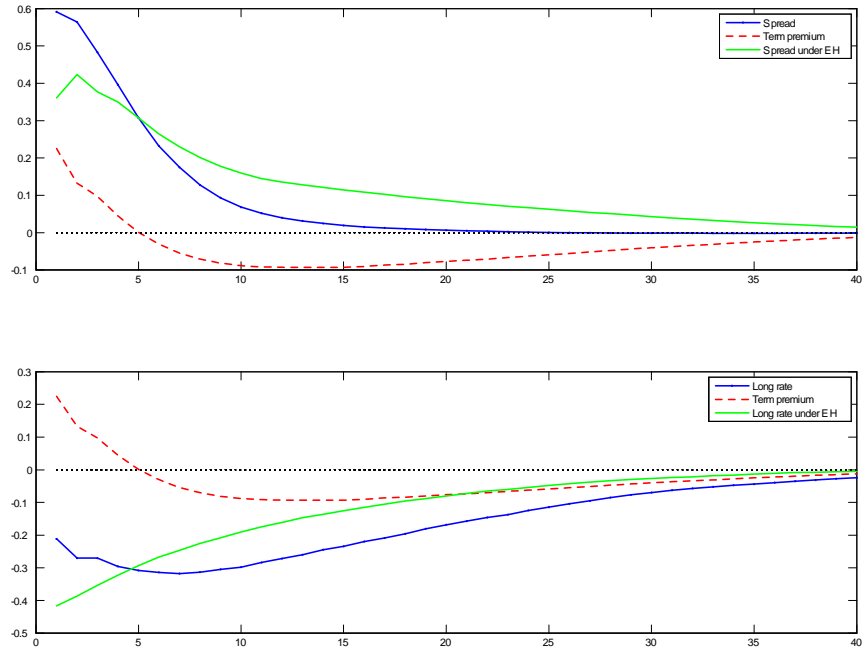


Fig. B.2.3: Decomposition of spread and long-rate response to TFP news shock into Expectations Hypothesis part and term premia part

B.3 Correspondence of extracted shocks from slope identification and TFP news identification

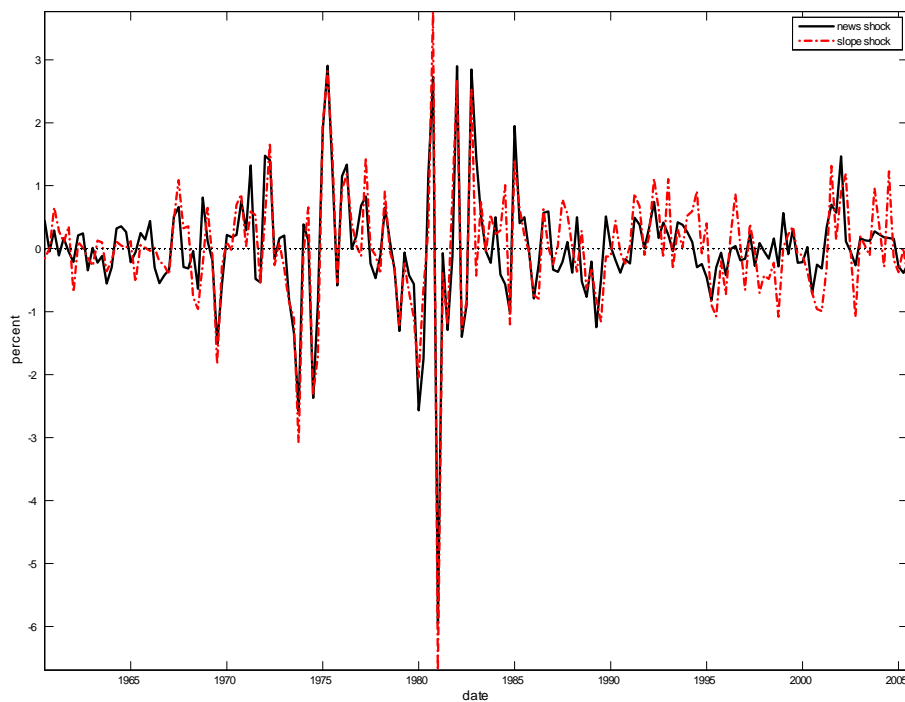


Fig. B.3.1: TFP news shock (solid black line) and Slope shock (dashed red line) extracted from baseline VAR

The correlation between the two shocks is 0.86 over the full sample; 0.91 over the pre-84 sample; and 0.58 over the post-84 sample.

C Results for contemporaneous TFP shock identification

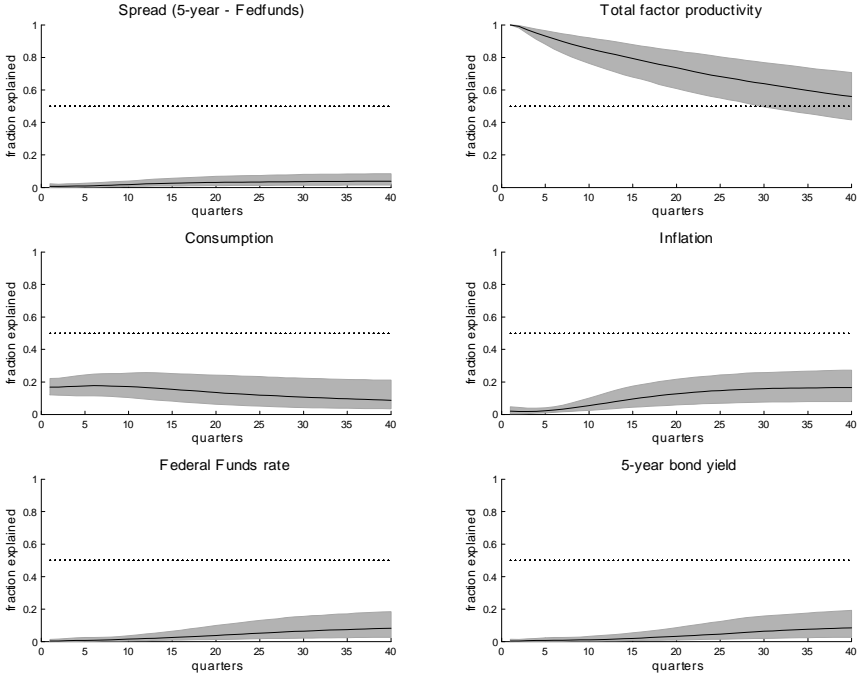


Fig. C.1: Fraction of FEV explained by contemporaneous TFP shock

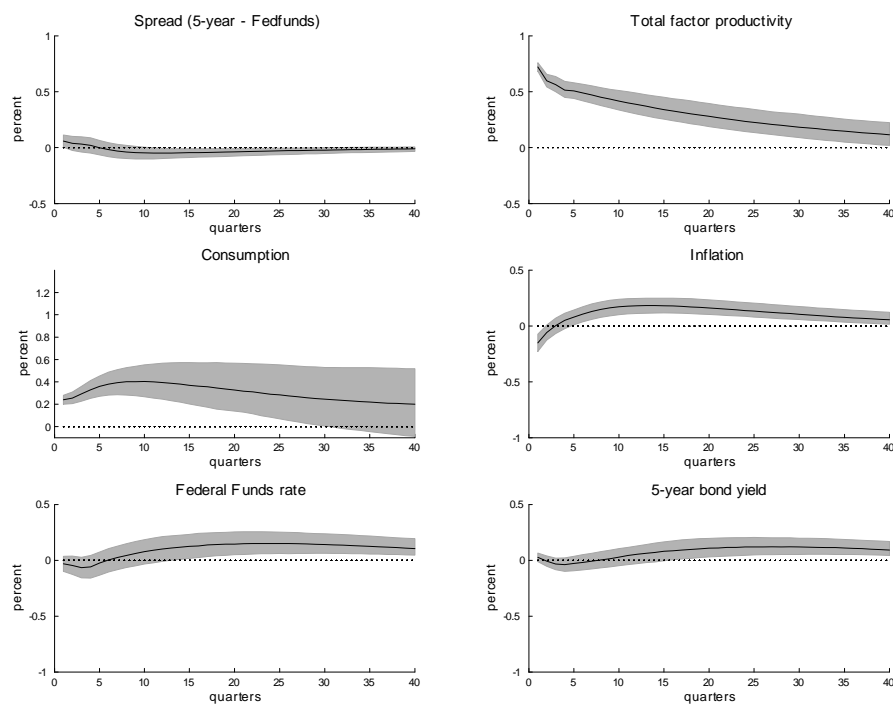


Fig. C.2: Impulse responses to 1% contemporaneous TFP shock

D Robustness checks

Here we report a variety of robustness checks of our baseline results. First, for the slope identification, we show result when the baseline VAR is estimated via OLS and confidence bands are obtained via bootstrapping. Second, we report robustness checks with respect to various alternative definitions of different variables in the VAR. In all cases, corresponding robustness checks for the news shock identification are very similar. We therefore do not report separate robustness checks for the news shock identification (results are, however, available from the authors upon request).

D.1 Baseline VAR results for OLS estimation

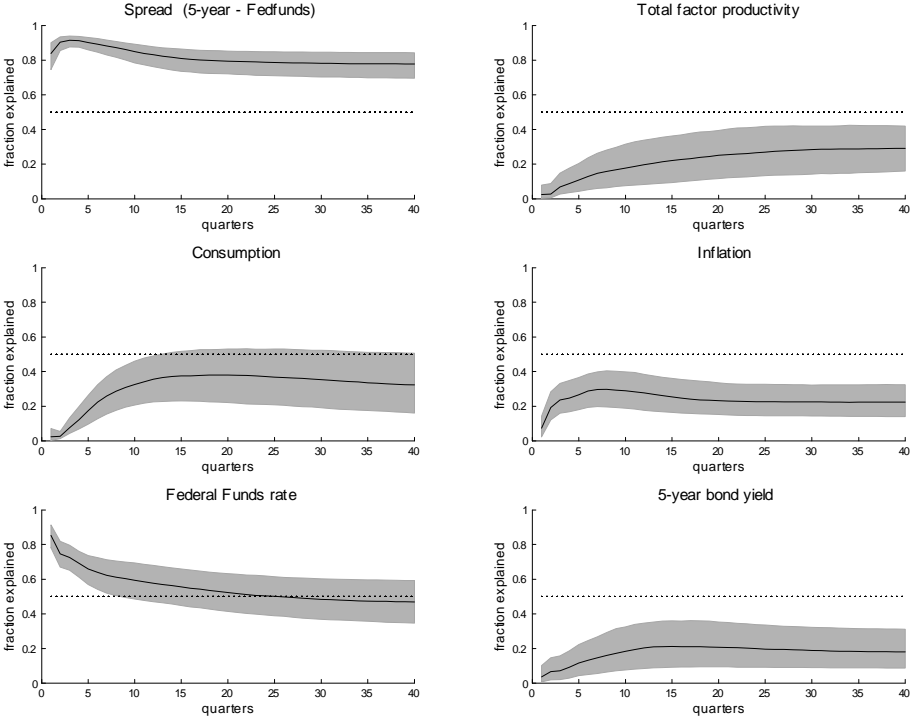


Fig. D.1.1: Fraction of FEV explained by slope shock (OLS estimates)

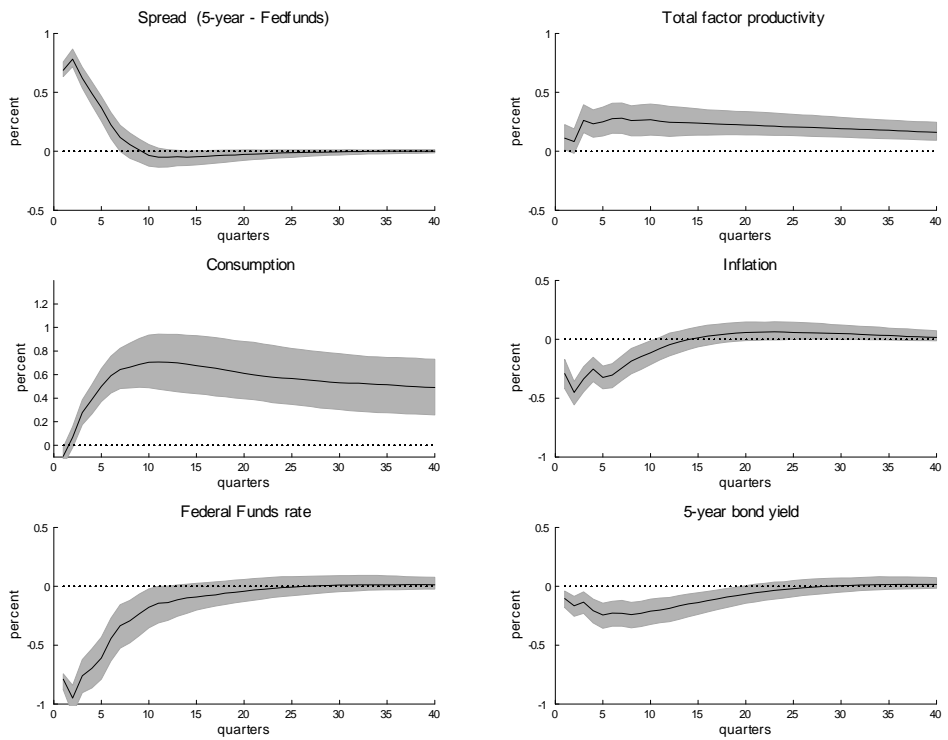


Fig. D.1.2: Impulse responses to 1% slope shock (OLS estimates)

D.2 Baseline VAR results for alternative variable definitions

The following robustness checks each replace one variable of the VAR with an alternative variable, leaving all other variables as in the baseline specification.

- Replacing total consumption by consumption of non-durables and services (nd&s)

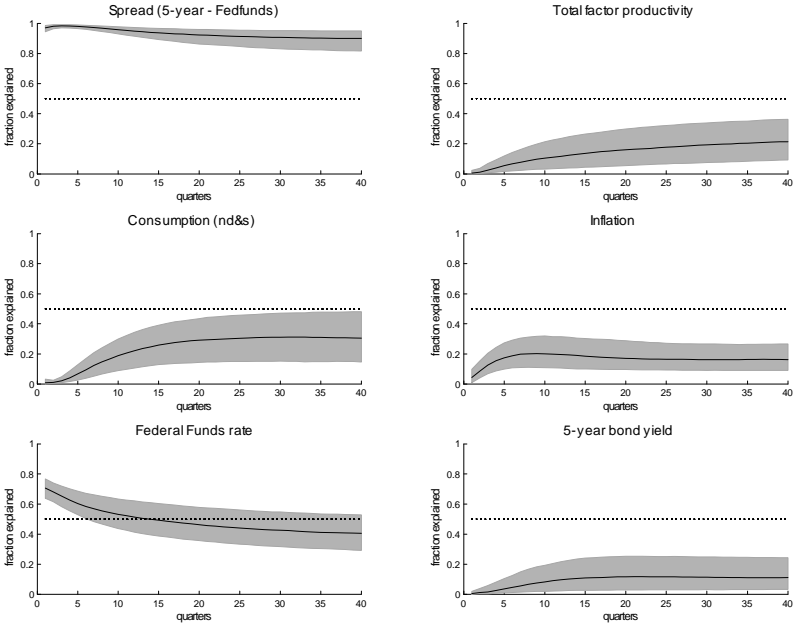


Fig. D.2.1: Fraction of FEV explained by slope shock

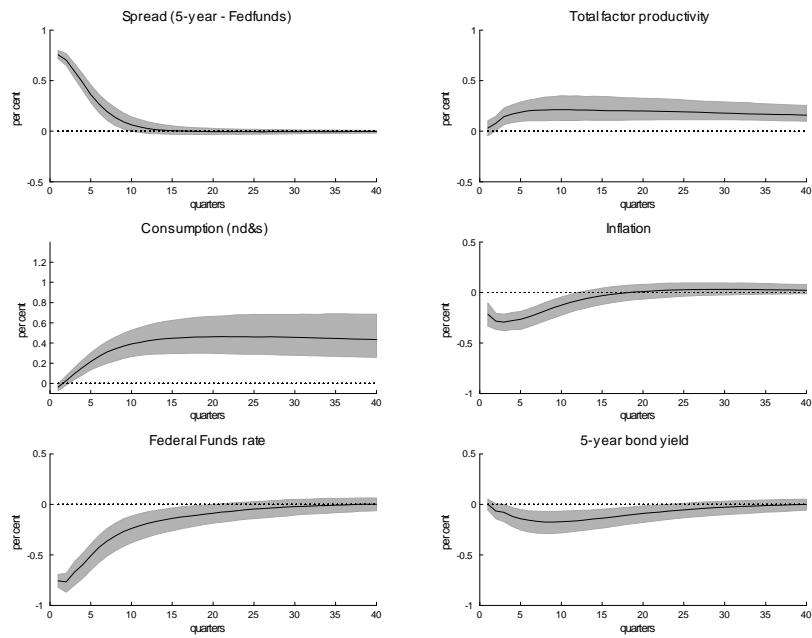


Fig. D.2.2: Impulse responses to 1% slope shock

- Replacing the GDP deflator as the measure for inflation with the PCE deflator

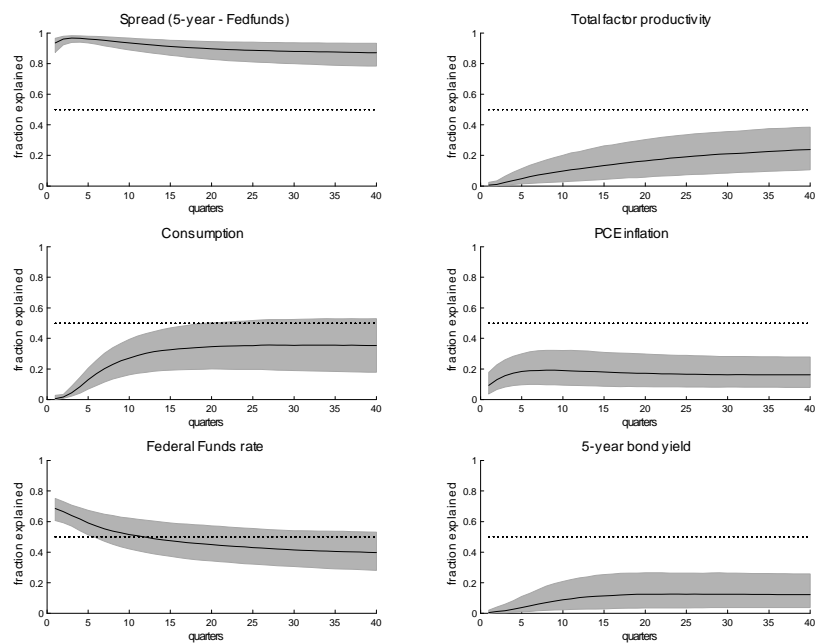


Fig. D.2.3: Fraction of FEV explained by slope shock

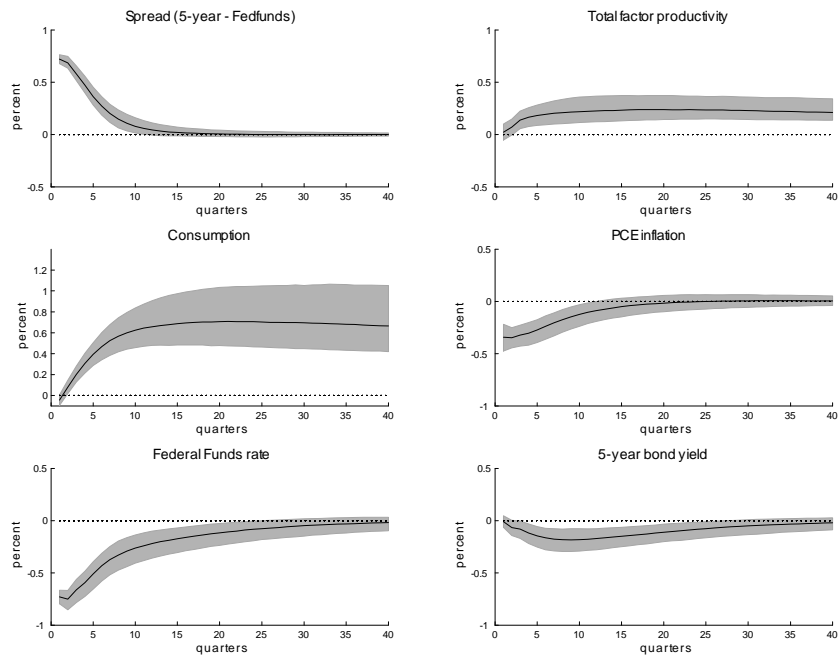


Fig. D.2.4: Impulse responses to 1% slope shock

- Replacing the GDP deflator as the measure for inflation with the CPI deflator⁵

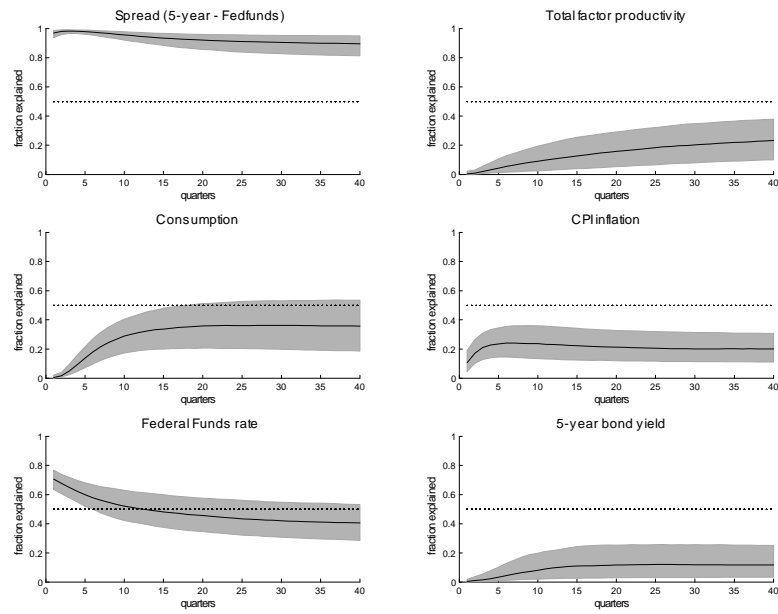


Fig. D.2.5: Fraction of FEV explained by slope shock

⁵The only exception is when we replace the GDP deflator with the CPI to compute inflation. As noted in the text, inflation in that case responds to a TFP news shock with a larger drop than the Fed Funds rate, implying an increase in the real short rate. We prefer either the GDP deflator or the PCE deflator over the CPI as a measure of inflation because the CPI suffers from substitution bias and is affected by large swings in food and energy prices that make it excessively volatile.

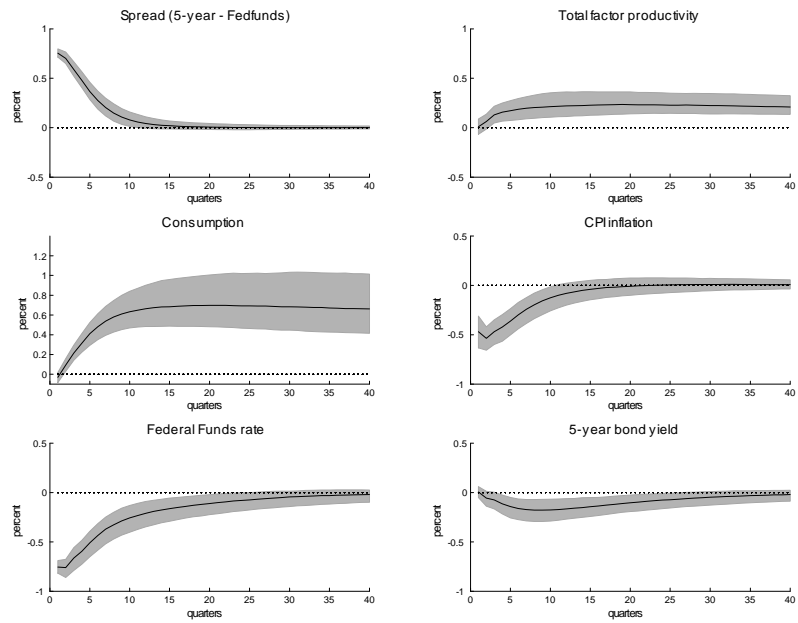


Fig. D.2.6: Impulse responses to 1% slope shock

- Replacing the Federal funds rate in the computation of the spread with the 3-month treasury bill rate

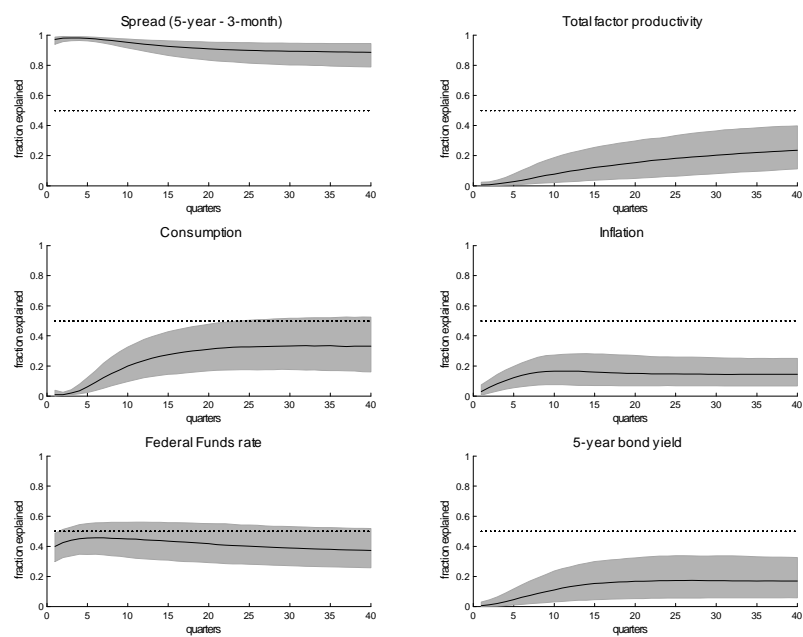


Fig. D.2.7: Fraction of FEV explained by slope shock

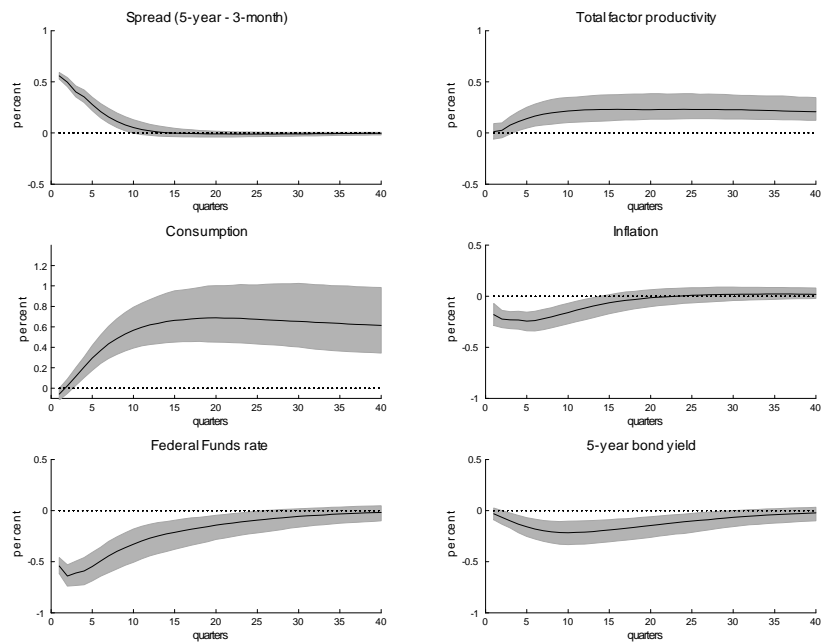


Fig. D.2.8: Impulse responses to 1% slope shock

- Replacing the 5-year treasury bond yield in the computation of the spread with the 10-year treasury bond yield⁶

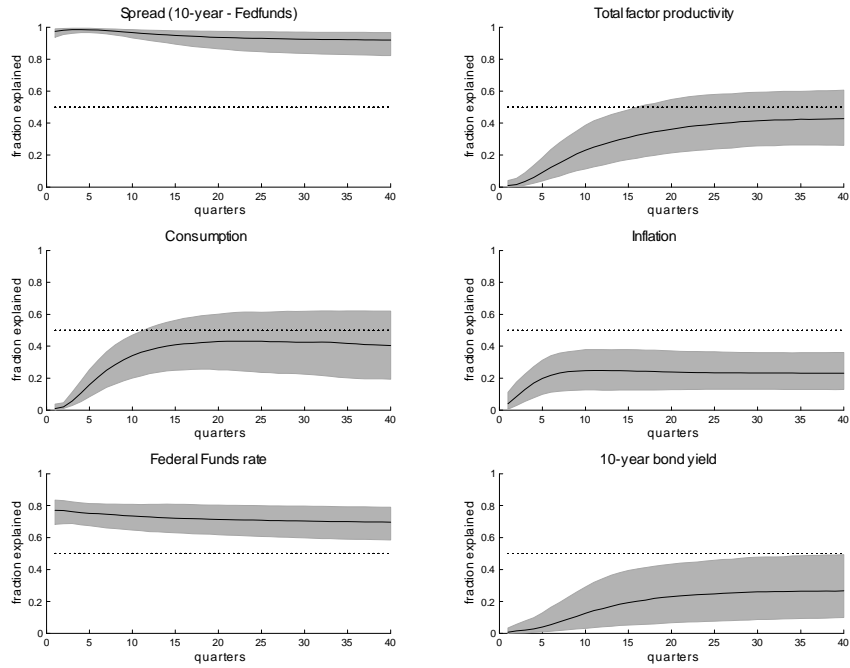


Fig. D.2.9: Fraction of FEV explained by slope shock

⁶The 10-year zero coupon bond yield series is taken from Gurkanyak, Sack and Wright (2007). Since this series is available only starting in 1971:4, the sample is shortened to 1971:4-2005:2.

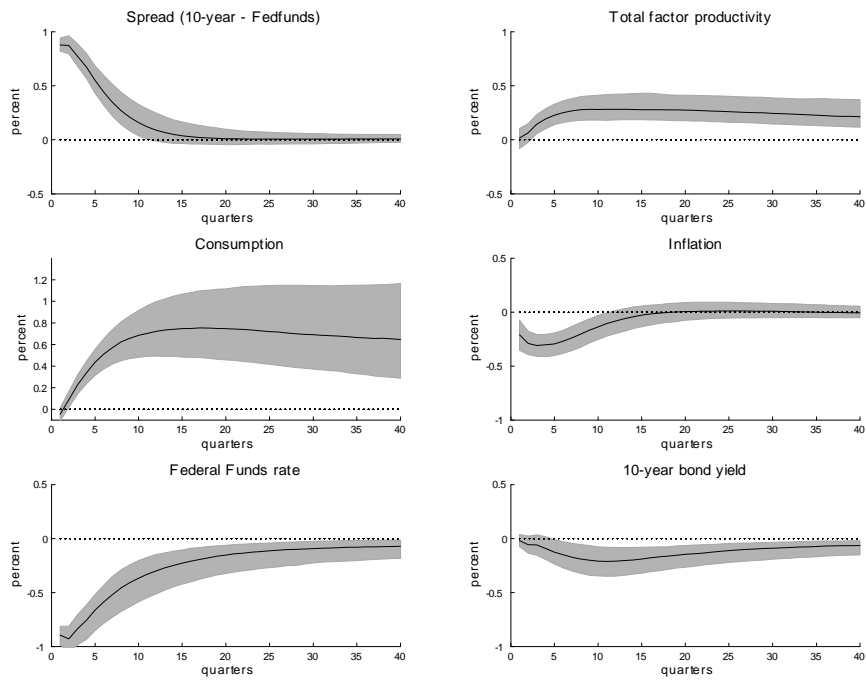


Fig. D.2.10: Impulse responses to 1% slope shock

- Replacing the 5-year bond – Federal funds rate spread with a term structure slope factor

computed as in Diebold and Li (2006)

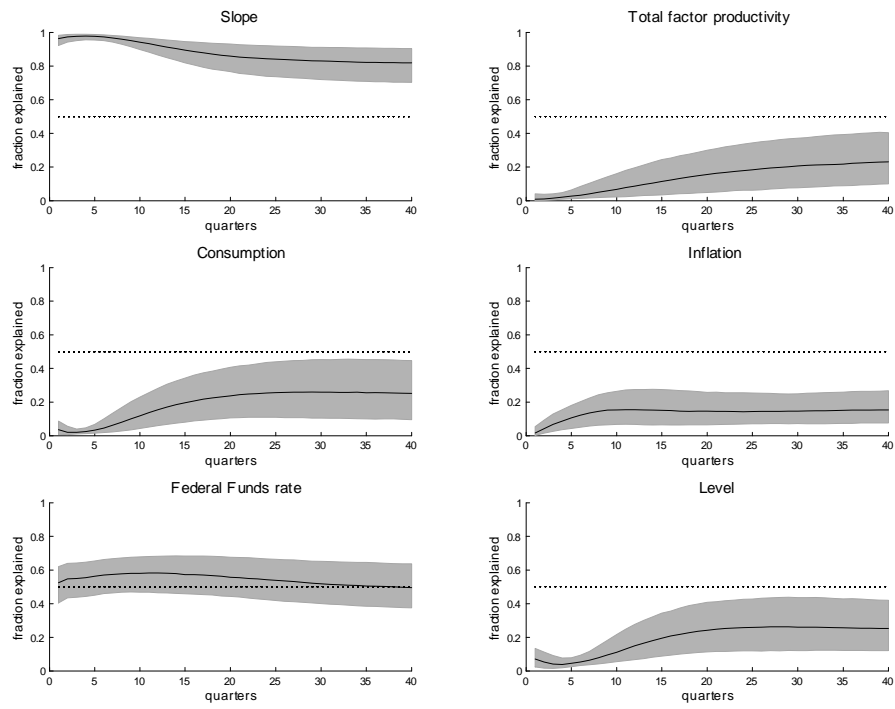


Fig. D.2.11: Fraction of FEV explained by slope shock

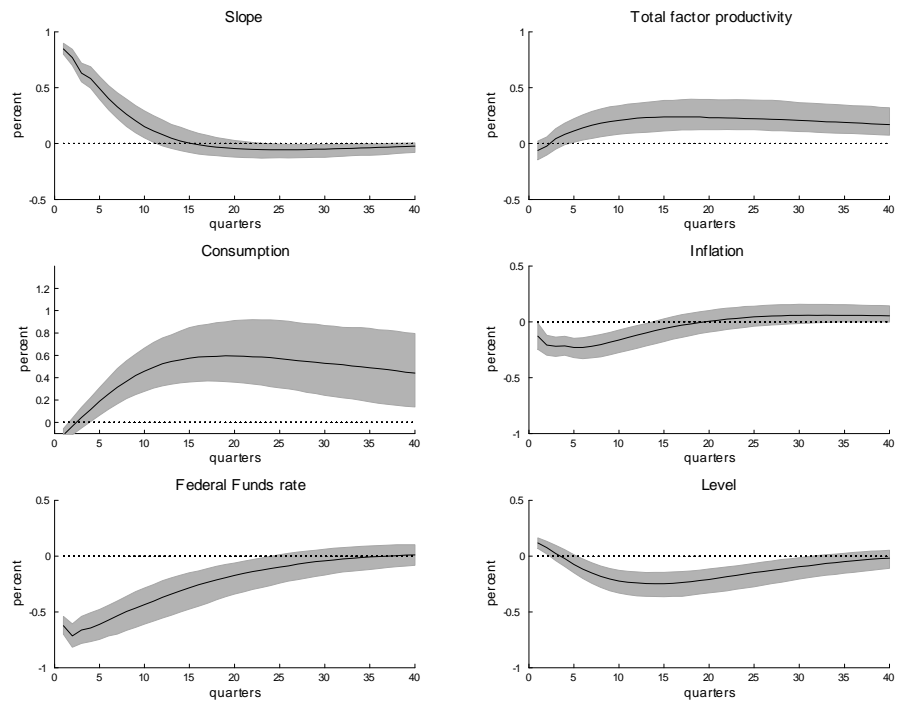


Fig. D.2.12: Impulse responses to 1% slope shock

E Robustness checks with larger VARs

Here we report details for the robustness checks with larger VARs in Section 5 of the paper. For both the slope identification and the TFP news identification, we show the full set of variance decompositions and impulse responses with corresponding confidence intervals.

E.1 Results for VAR with investment and equipment&software price deflators

- Slope identification

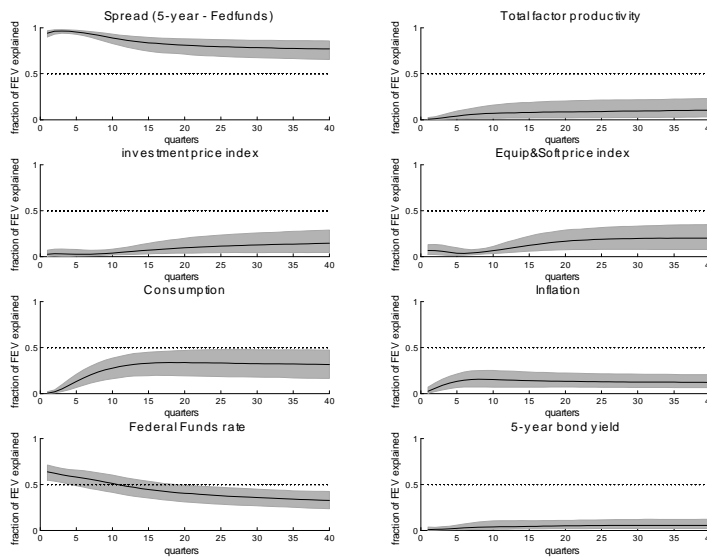


Fig. E.1.1: Fraction of FEV explained by slope shock

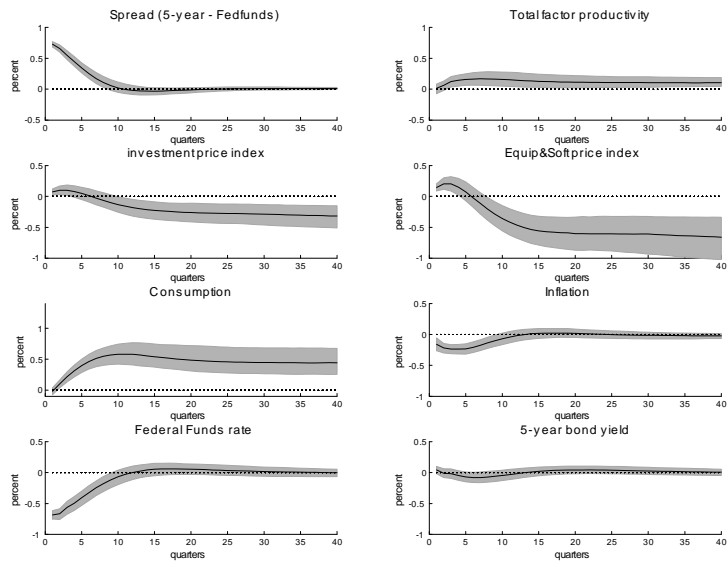


Fig. E.1.2: Impulse responses to 1% slope shock

- TFP news identification

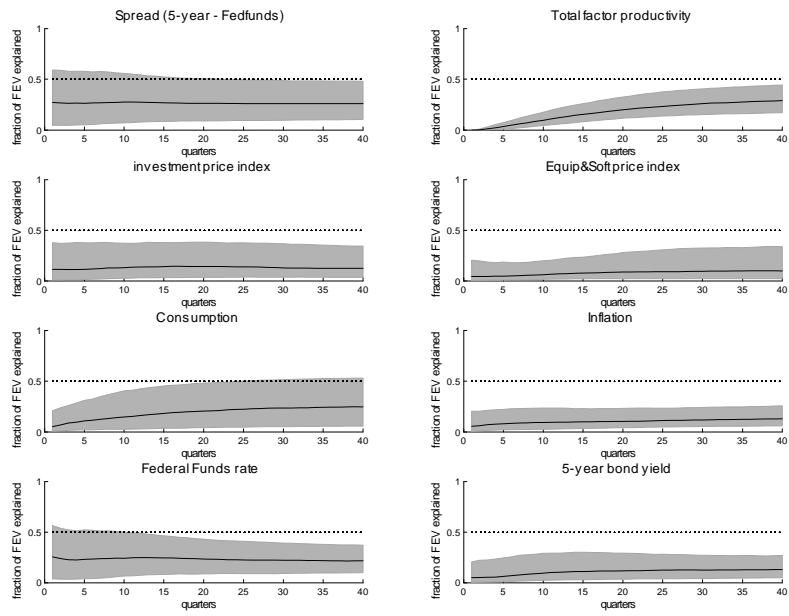


Fig. E.1.3: Fraction of FEV explained by TFP news shock

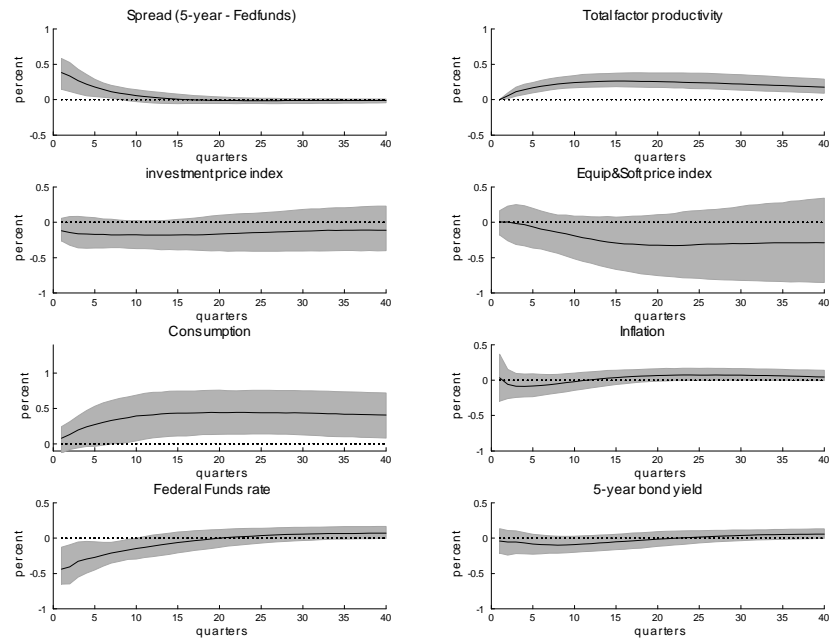


Fig. E.1.4: Impulse responses to 1% TFP news shock

E.2 Results for VAR with GDP, investment and real S&P500 price

- Slope identification

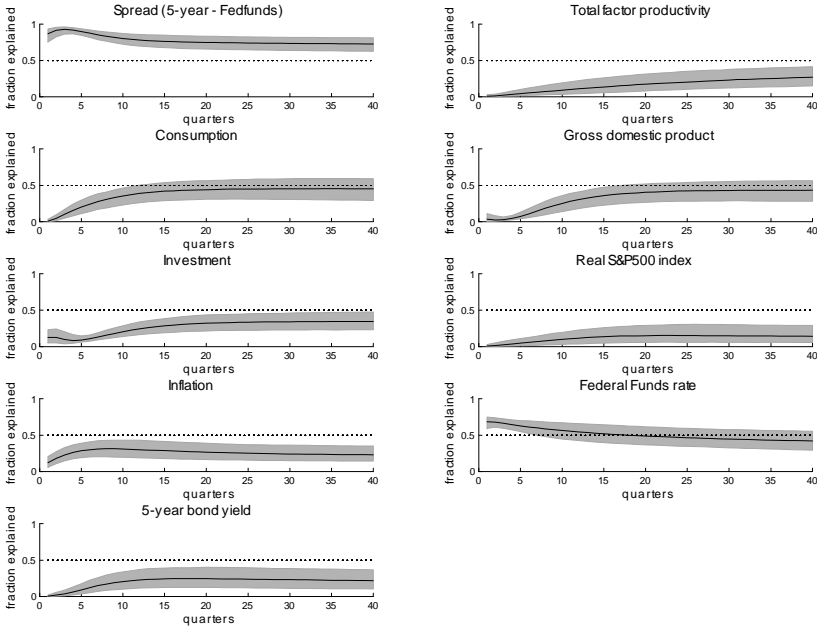


Fig. E.2.1: Fraction of FEV explained by slope shock

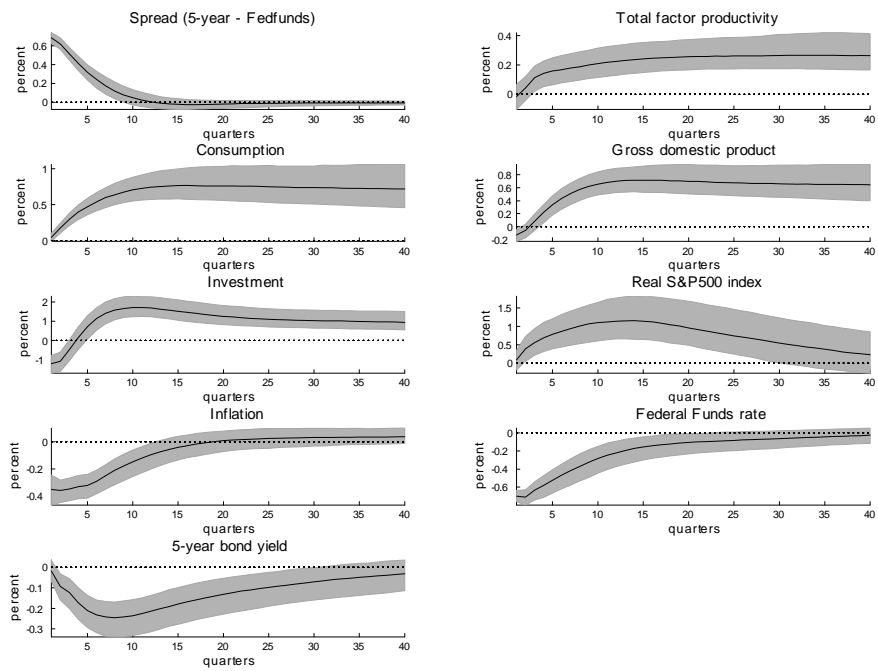


Fig. E.2.2: Impulse responses to 1% slope shock

- TFP news identification

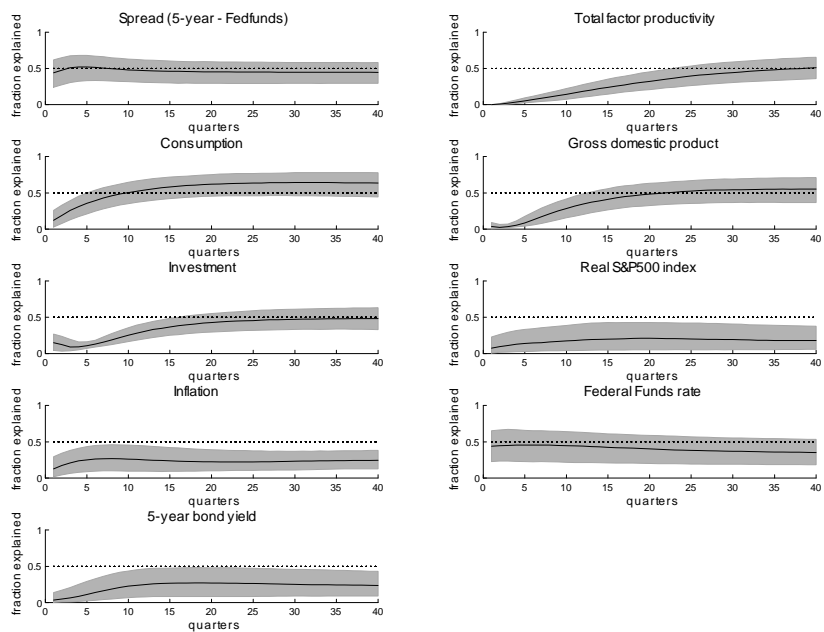


Fig. E.2.3: Fraction of FEV explained by TFP news shock

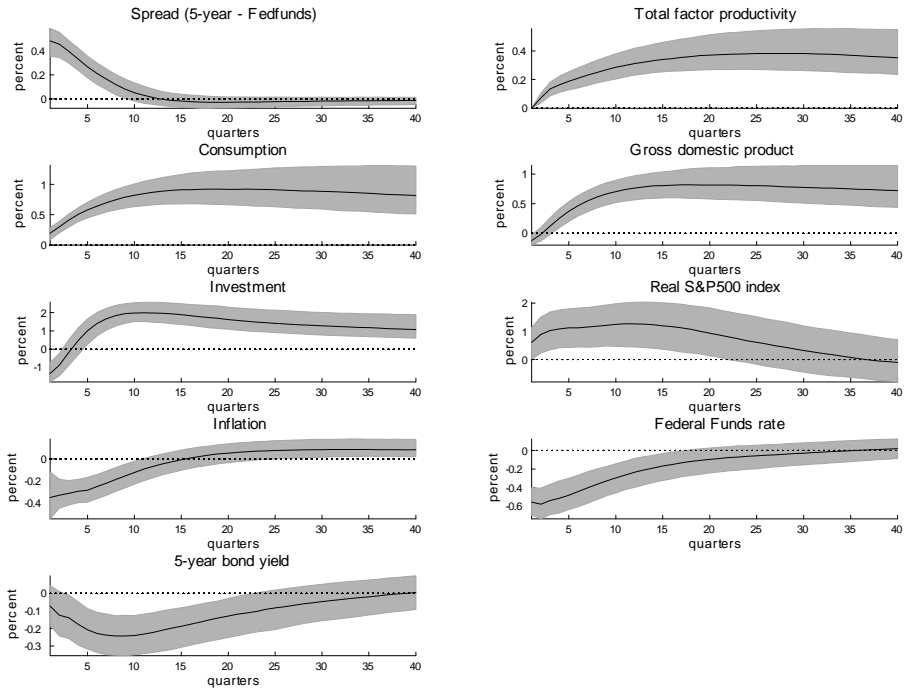


Fig. E.2.4: Impulse responses to 1% TFP news shock

- Correspondence between the two extracted shocks

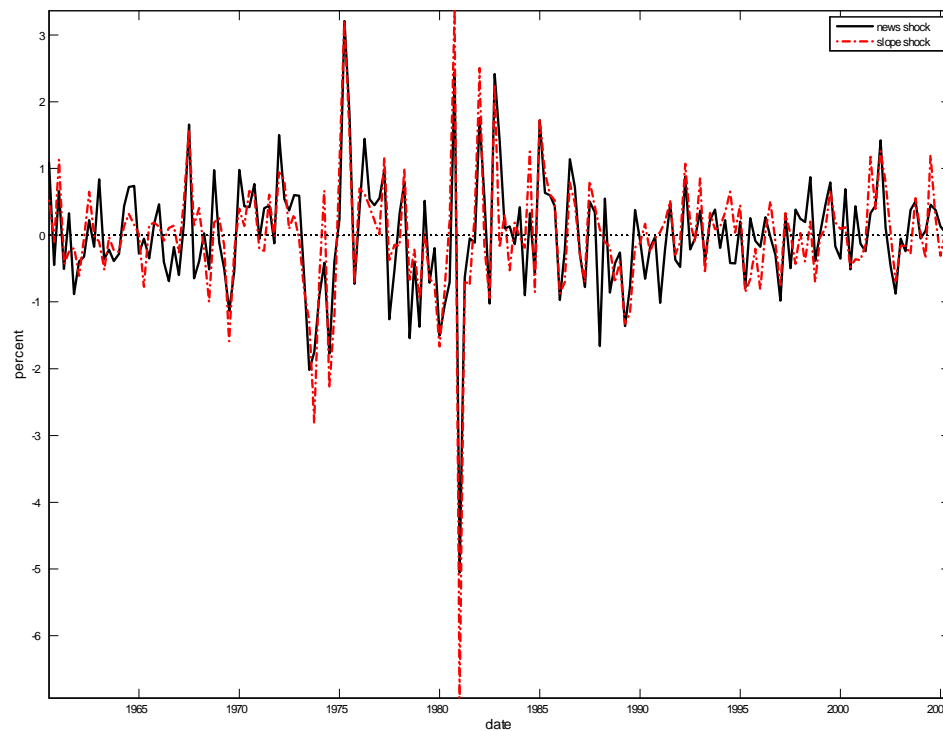


Fig. E.2.5: TFP news shock (solid black line) and Slope shock (dashed red line) extracted from larger VAR with output, investment and real S&P500 price

The correlation between the two shocks is 0.84 over the full sample; 0.88 over the pre-84 sample; and 0.68 over the post-84 sample.

F Fit of interest rate rule with VAR results

As discussed in the final section of the main text, an interesting question is whether the joint response of the Federal Funds rate, inflation and output to a TFP news shock as implied by our VAR estimates is consistent with an interest rate rule for monetary policy.

The interest rate rule we consider is

$$i_t = \gamma i_{t-1} + (1 - \gamma)[\theta_\pi \pi_t + \theta_y \Delta y_t],$$

where i_t denotes the Federal Funds rate; π_t denotes inflation; and Δy_t denotes output growth.⁷ The black solid line shows the impulse response of the Federal Funds rate to a TFP news shock as estimated from the large VAR above (Figure E.2.4), with the 68% confidence bands in grey. The blue dotted line shows the impulse response of the Federal Funds rate as implied by the above rule for $\gamma = 0.9$, $\theta_\pi = 1.5$, $\theta_y = 0.5$. The red solid line with dots shows the impulse response of the

⁷Results are very similar if we replace output growth by a measure of the output gap, specified as $y_t - \bar{y}_t$, where potential output \bar{y}_t is defined as

$$\bar{y}_t = \alpha \bar{y}_{t-1} + (1 - \alpha) trend_t$$

and $trend_t$ is proportional to the path of TFP in response to the news shock.

Federal Funds rate as implied by the same rule but for $\gamma = 0.5$, $\theta_\pi = 2$, $\theta_y = 0.5$.

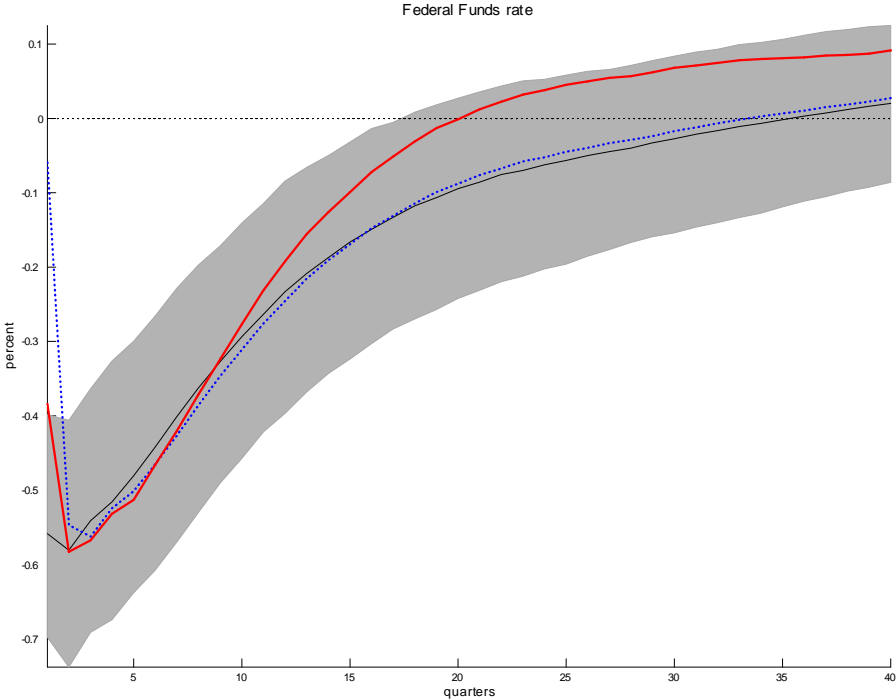


Fig. F.1. Impulse response of Federal Funds rate to TFP news shock as estimated from VAR (black solid line) and as implied by two calibrations of the interest rate rule (blue dotted line and red line with dots).

As can be seen from this figure, the first calibration (blue dotted line) fits the impulse response from the VAR very closely except for the impact response. This is because of the high degree of interest rate smoothing $\gamma = 0.9$. The second calibration (red line with dots) fits the impulse response from the VAR somewhat less closely after 10 quarters out but provides a better match of the impact response.