Université de Neuchâtel
Institut de recherches économiques

IRENE Working Paper 24-03

Multi-dimensional monetary policy shocks based on heteroscedasticity

Marc Burri and Daniel Kaufmann

Multi-dimensional monetary policy shocks based on heteroscedasticity*

Marc Burri[†] Daniel Kaufmann[‡ §]

19 August 2024

Abstract: We propose a two-step approach to estimate multi-dimensional monetary policy shocks and their causal effects requiring only daily financial market data and policy events. First, we combine a heteroscedasticity-based identification scheme with recursive zero restrictions along the term structure of interest rates to disentangle multi-dimensional monetary policy shocks and derive an instrumental variables estimator to estimate dynamic causal effects. Second, we propose to use the Kalman filter to compute the linear minimum mean-square-error prediction of the unobserved monetary policy shocks. We apply the approach to examine the causal effects of US monetary policy on the exchange rate. The heteroscedasticity-based monetary policy shocks display a relevant correlation with existing high-frequency surprises. In addition, their dynamic causal effects on the exchange rate are similar. This suggests the approach is a valid alternative if high-frequency identification schemes are not applicable.

JEL classification: C3, E3, E4, E5, F3

Keywords: Monetary policy shocks, forward guidance, large-scale asset purchases, identification through heteroscedasticity, instrumental variables, term structure of interest rates, exchange rate

*We thank Jean-Marie Grether, Leif Anders Thorsrud, Eric Swanson, Mark Watson, Linyan Zhu, as well as participants at the SSES Annual Congress and the Workshop on Applied Macroeconomics and Monetary Policy in St. Gallen for helpful discussions and comments.

[†]University of Neuchâtel, Institute of Economic Research, Rue A.-L. Breguet 2, CH-2000 Neuchâtel, marc.burri@unine.ch

[‡]University of Neuchâtel, Institute of Economic Research, Rue A.-L. Breguet 2, CH-2000 Neuchâtel, daniel.kaufmann@unine.ch

[§]KOF Swiss Economic Institute, ETH Zurich

1 Introduction

Since the global financial crisis, most major central banks regularly used non-conventional tools, such as forward guidance and large-scale asset purchases (LSAP). Understanding how these tools affect financial markets and the economy more broadly is, therefore, key. Recent research has identified the causal effects of multiple dimensions of central bank decisions based on intraday financial market data. High-frequency identification schemes exploit changes in financial market variables in a narrow window around central bank announcements and impose additional restrictions to disentangle multiple dimensions (see, e.g., Gürkaynak et al., 2005, Altavilla et al., 2019, Swanson, 2021). However, if high-frequency data is lacking, or if the exact time of a policy event is unknown, such an approach is infeasible.[1] This paper proposes a two-step approach to estimate multi-dimensional monetary policy shocks and their causal effects, requiring only daily financial market data and policy events.

First, following Rigobon (2003) and Rigobon and Sack (2004), we exploit the change in the variance-covariance matrix of daily financial market variables on monetary policy announcement days to identify the causal effects of monetary policy shocks. We additionally impose recursive zero restrictions on the impact matrix along the term structure of interest rates to disentangle three orthogonal dimensions. Specifically, we identify a shock to the short-term interest rate (target shock), medium-term interest rate (path shock or forward guidance), and term spread (term premium shock or LSAP). The zero restrictions impose that a path shock has no immediate impact on the short-term interest rate, and the term premium shock has no immediate impact on short- and medium-term interest rates. In contrast, a target shock can affect all financial market variables on the same day. Importantly, these restrictions are unnecessary in identifying the overall effect of a linear combination of all monetary policy shocks. However, they allow us to disentangle multiple dimensions if they exist. We also show that we can recursively estimate each column of the impact matrix up to a scale using a modified version of the instrumental variable (IV) estimator proposed by Rigobon and Sack (2004).[2]

Second, using the Kalman filter, we derive the unobserved shocks' linear minimum mean-square-error (MSE) prediction. Recovering these shocks is helpful, for example, to estimate the causal effects on low-frequency macroeconomic variables with IV-structural

[1] Although databases on high-frequency surprises exist for various countries, due to the high data requirements, they usually only start after the mid-1990s (see, e.g., Altavilla et al., 2019, Braun et al., 2024).

[2] In addition, we can estimate dynamic causal effects using local projections (Jordà, 2005) and apply weak instruments tests in the presence of multiple endogenous regressors and instruments suggested by Lewis and Mertens (2022).

vector autoregressions (IV-SVAR) or IV-local projections (IV-LP). The heteroscedasticity-based identification scheme delivers all necessary inputs: an estimate of the impact matrix (up to a scale), the variance-covariance matrix of the data on policy event days, and the data. There are two main advantages compared to existing approaches.[3] To extract a heteroscedasticity-based monetary policy shock series Bu et al. (2021) regress interest rate changes across the term structure on the impact responses in the spirit of Fama and MacBeth (1973). This approach requires assumptions that may be violated in empirical applications. Specifically, it requires an orthogonality condition between the impact response to the monetary policy shocks and the impact response to other shocks. If this condition is violated, the monetary policy shock series will be contaminated by other shocks occurring on policy event days. In addition, the cross-sectional regressions do not yield the optimal prediction of the unobserved shocks in the minimum MSE sense.

We then apply the two-step approach to estimate the causal impact of US monetary policy on the US dollar exchange rate for two reasons. First, we can compare the results to existing multi-dimensional high-frequency surprises by Swanson (2021). Second, the identification scheme does not directly constrain the US dollar exchange rate. Our findings suggest that the heteroscedasticity-based identification scheme is a valid alternative to estimate multi-dimensional monetary policy shocks. The heteroscedasticity-based target, path, and term premium shocks correlate with Swanson's high-frequency Federal Funds Rate (FFR), path, and LSAP surprises, respectively (0.66, 0.56, and 0.50). In addition, the daily impulse responses to the heteroscedasticity-based shocks are qualitatively identical to those of their high-frequency counterparts. A surprise monetary policy tightening, be it an increase in the short-term interest rate target, the medium-term interest rate path, or an increase in the term premium, all appreciate the US dollar exchange rate within a few working days. Therefore the daily data do not point to a relevant delayed overshooting puzzle often found in low-frequency VARs (see, e.g., Eichenbaum and Evans, 1995, Kim et al., 2017). However, we also find that term premium shocks, and to a lesser extent path shocks, suffer from a weak instruments problem. Meanwhile, the target shock passes the weak instruments tests. Again, we find qualitatively identical results using high-frequency surprises.

We then conduct various robustness tests, qualitatively confirming the main results. Perhaps the most relevant differences emerge when estimating monthly impulse responses of macroeconomic variables to the heteroscedasticity-based monetary policy shocks in an IV-LP

[3]The approach based on the Kalman filter can be used for other identification schemes where the shocks are not directly observed, as long as these three inputs are available.

framework. The exchange rate appreciation to the target tightening is delayed by about 30 months, while we observe an appreciation after 10 months in response to a path tightening. Therefore, there is a stark difference between the qualitative response at daily and monthly frequency, as well as between different monetary policy dimensions. However, the weak instruments problem for the path and term premium shocks carry over to the monthly data for both heteroscedasticity-based and high-frequency identification schemes.

Our paper contributes to three strands of the literature. Following the seminal work by Kuttner (2001), most researchers use high-frequency identification schemes to identify multiple dimensions of monetary policy. Gürkaynak et al. (2005) show that path surprises, associated with longer maturity interest rates, have a stronger effect on long-term interest rates than target surprises. In the wake of the financial crisis, researchers aimed to identify additional dimensions. Swanson (2021, 2023b) and Altavilla et al. (2019) estimate multiple factors from a cross-section of high-frequency financial market data and rotate them so that they can be interpreted as a target, path, or LSAP surprises. While high-frequency surprises have become a widespread identification scheme, it is not without problems. Brennan et al. (2024) have shown that different measures of high-frequency surprises yield varying results. In addition, the surprises may still be contaminated by background noise and other shocks. Nakamura and Steinsson (2018) show that these high-frequency surprises comprise an 'information effect', that is, news that the central bank communicates about the state of the economy. These information effects may bias the results because they may lead to positive co-movement of interest rates and stock prices in response to supposedly exogenous monetary policy decisions. The assumption that variation in high-frequency surprises is only caused by monetary policy shocks has been questioned. Studies have shown that the surprises are predictable by the information available to the public before the FOMC announcements (see, e.g. Bauer and Swanson, 2022, 2023, Miranda-Agrippino and Ricco, 2021, Zhu, 2023).[4] Jarociński (2024) and Georgiadis and Jarociński (2023) aim to circumvent this problem by exploiting fat tails in high-frequency surprises to identify multiple dimensions of US monetary policy. We also exploit a change in the distribution of financial market variables on policy announcement days. However, we propose to use the change in the variance-covariance matrix and additional zero restrictions to identify multi-dimensional monetary policy shocks without using high-frequency data.

Therefore, the paper is related to studies identifying causal effects of monetary policy and other events via heteroscedasticity (see, e.g., Rigobon, 2003, Rigobon and Sack, 2004, Wright,

[4] Recently, Schlaak et al. (2023) show how to combine high-frequency identification schemes with heteroscedasticity. As they exploit more exogenous variation, they can test the relevance and exogeneity of the instruments. They find evidence against the validity of high-frequency monetary policy surprises.

2012). In addition, we follow Canetg and Kaufmann (2022) who use additional zero restrictions in a heteroscedasticity-based identification scheme to disentangle multi-dimensional shocks of the Swiss National Bank's debt security auctions.[5] They assume the overnight rate shock immediately affects all financial market variables through the short-term interest rate. However, the signaling shock on impact affects only forward-looking variables, such as stock prices, but not the short-term interest rate. We expand this literature by showing that dynamic causal effects can be estimated with an IV approach and that the unobserved shocks can be predicted using the Kalman filter.

Finally, we also contribute to a literature estimating the causal impact of monetary policy on the exchange rate.[6] There is still a controversy about whether the exchange rate immediately overshoots after a monetary policy shock, as in the theory by Dornbusch (1976) and no consensus about the size and persistence of the effects.[7] Theoretically, the response depends on rigidity in goods markets and the monetary policy regime (Benigno, 2004), or portfolio adjustment costs (Bacchetta and van Wincoop, 2021). Empirically, the response depends on the sample period (Kim et al., 2017) and the identification scheme (Faust et al., 2003, Kearns and Manners, 2006, Scholl and Uhlig, 2008, Bjørnland, 2009). Some authors suggest that the exchange rate does not overshoot at all (Schmitt-Grohé and Uribe, 2022) or displays a delayed overshooting, for example, due to information rigidity (Müller et al., 2023). Other authors find that the exchange rate often depreciates on a day with a policy tightening (Gürkaynak et al., 2021).[8] To the best of our knowledge, no paper examined whether the response differs depending on the specific dimension of monetary policy.

The remainder of the paper is structured as follows. Section 2 presents the identification and estimation strategy, as well as the Kalman-filter approach to extract multi-dimensional monetary policy shocks. Section 3 discusses the results of the empirical application before the last section concludes.

[5] Lewis (2019) estimates multiple dimensions of unconventional monetary policy announcements using intraday heteroscedasticity. His approach allows for the varying importance of shocks across announcements. He also finds that forward guidance has relevant effects. However, he also finds evidence of relevant information and large-scale asset purchase shocks.

[6] There are various studies on global spillovers of US monetary policy using high-frequency surprises (see, e.g., Georgiadis and Jarociński, 2023, Ricco et al., 2020).

[7] To answer this question, researchers resorted to theoretical and empirical approaches (see, e.g., Dornbusch, 1976, Eichenbaum and Evans, 1995, Rogoff, 2002, Bjørnland, 2009, Schmitt-Grohé and Uribe, 2022, Bacchetta and van Wincoop, 2021).

[8] Wright (2012) identifies monetary policy shocks at the effective lower bound through heteroscedasticity and shows that these shocks immediately affect bilateral exchange rates. However, he does not report dynamic causal effects.

2 Heteroscedasticity-based identification with zero restrictions

We propose identifying and estimating the causal effects of multiple dimensions of (monetary policy) shocks based on heteroscedasticity and zero restrictions. Imposing recursive zero restrictions on top of assuming that some shocks occur only in specific periods allows us to disentangle the causal effects of multi-dimensional shocks. Therefore, we can use an IV approach to estimate the causal effects of these multiple dimensions. In addition, we can extract the underlying unobserved structural shocks using the Kalman filter.

2.1 Identification and estimation of causal effects

Suppose the data-generating process reads:

$$\begin{aligned} y_t &= \Psi\varepsilon_t + \Gamma v_t \quad \text{for } t \in P \\ y_t &= \Gamma v_t \quad \text{for } t \in C \end{aligned} \tag{1}$$

where y_t is a vector of N dependent variables, ε_t is a vector of E i.i.d. structural shocks on policy event days (P), and v_t is a vector of R i.i.d. other shocks on policy event as well as control days (P and C).[9] Furthermore, Γ and Ψ denote impact matrices of dimensions $N \times R$ and $N \times E$, respectively. Finally, we assume that Ψ is lower triangular.

We will justify the identifying assumptions in the empirical application below. In what follows, we show that these assumptions allow us to sequentially identify and estimate the causal impact of ε_t using the change in the variance-covariance of y_t on policy event days. The first equation of the model reads:

$$\begin{aligned} y_{1t} &= \Psi_{11}\varepsilon_{1t} + \Gamma_1 v_t \quad \text{for } t \in P \\ y_{1t} &= \Gamma_1 v_t \quad \text{for } t \in C \end{aligned}$$

where Ψ_{ij} denotes the ith row and jth column of Ψ and Γ_i denotes the ith row of Γ.

Because Ψ is lower triangular, only the first structural shock (ε_{1t}) affects y_{1t}. Therefore, the same insights as in Rigobon (2003) apply. The variance of y_{1t} differs between policy event and

[9] We drop constant terms, lagged dependent, and other exogenous variables for ease of exposition. In addition, we focus on the identification and estimation of the impact matrix. We can extend this framework to include additional control variables and estimate dynamic causal effects using local projections following Jordà (2005). Details are given in Appendix A.

control days:

$$\begin{aligned} \mathbb{V}[y_{1t}] &= \Psi_{11}^2\sigma_{1\varepsilon}^2 + \textstyle\sum_{r=1}^R \Gamma_{1r}^2\sigma_{rv}^2 && \text{for } t \in P \\ \mathbb{V}[y_{1t}] &= \textstyle\sum_{r=1}^R \Gamma_{1r}^2\sigma_{rv}^2 && \text{for } t \in C \end{aligned}$$

where $\sigma_{e\varepsilon}^2$ and σ_{rv}^2 denote the variances of structural shock e and other shock r, respectively. Thus, we can identify Ψ_{11} up to a scale from the difference in the variance on policy event and control days ($\Psi_{11}^2\sigma_{1\varepsilon}^2$).

We can identify the impact on y_{2t} from the change in the covariance between the first two variables:

$$\begin{aligned} \mathbb{COV}[y_{1t}, y_{2t}] &= \Psi_{11}\Psi_{21}\sigma_{1\varepsilon}^2 + \textstyle\sum_{r=1}^R \Gamma_{1r}\Gamma_{2r}\sigma_{rv}^2 && \text{for } t \in P \\ \mathbb{COV}[y_{1t}, y_{2t}] &= \textstyle\sum_{r=1}^R \Gamma_{1r}\Gamma_{2r}\sigma_{rv}^2 && \text{for } t \in C \end{aligned}$$

where $\Psi_{11}\Psi_{21}\sigma_{1\varepsilon}^2$ corresponds to the difference in the covariance between policy event and control days. Having identified Ψ_{11}, this difference allows us to identify Ψ_{21} up to a scale. Similarly, the impacts on other variables ($\Psi_{31}, \ldots, \Psi_{N1}$) can be identified from the corresponding covariances.

The impact of the second structural shock can be identified from the second equation of the model:

$$\begin{aligned} y_{2t} &= \Psi_{21}\varepsilon_{1t} + \Psi_{22}\varepsilon_{2t} + \Gamma_2 v_t && \text{for } t \in P \\ y_{2t} &= \Gamma_2 v_t && \text{for } t \in C \end{aligned}$$

where the variance of y_{2t} corresponds to:

$$\begin{aligned} \mathbb{V}[y_{2t}] &= \Psi_{21}^2\sigma_{1\varepsilon}^2 + \Psi_{22}^2\sigma_{2\varepsilon}^2 + \textstyle\sum_{r=1}^R \Gamma_{2r}^2\sigma_{rv}^2 && \text{for } t \in P \\ \mathbb{V}[y_{2t}] &= \textstyle\sum_{r=1}^R \Gamma_{2r}^2\sigma_{rv}^2 && \text{for } t \in C \end{aligned}$$

The difference between the variance of y_{2t} on event and control days ($\Psi_{21}^2\sigma_{1\varepsilon}^2 + \Psi_{22}^2\sigma_{2\varepsilon}^2$) comprises the effects of shocks 1 and 2. Therefore, applying the approach by Rigobon (2003) using y_{2t} identifies the effect of a linear combination of the two shocks. However, having identified Ψ_{21} from the first equation, Ψ_{22} is subsequently identified up to a scale. Similarly, we can identify the effect of shock 2 on the remaining variables using the covariances, for example, between y_{2t} and y_{3t}:

$$\begin{aligned} \mathbb{COV}[y_{2t}, y_{3t}] = &\ \Psi_{21}\Psi_{31}\sigma^2_{1\varepsilon} + \Psi_{22}\Psi_{32}\sigma^2_{2\varepsilon} + \textstyle\sum_{r=1}^{R}\Gamma_{2r}\Gamma_{3r}\sigma^2_{rv} && \text{for } t \in P \\ \mathbb{COV}[y_{2t}, y_{3t}] = &\ \textstyle\sum_{r=1}^{R}\Gamma_{2r}\Gamma_{3r}\sigma^2_{rv} && \text{for } t \in C \end{aligned}$$

Conditional on having identified Ψ_{21} and Ψ_{31} from the first equation, and Ψ_{22} from the second equation, we can identify Ψ_{32} up to a scale from the difference in the two covariances $(\Psi_{21}\Psi_{31}\sigma^2_{1\varepsilon} + \Psi_{22}\Psi_{32}\sigma^2_{2\varepsilon})$. Similar arguments apply to identifying the effect of shock 3 on other variables and the effects of additional dimensions.

The additional recursive zero restrictions also allow us to estimate each column of Ψ up to a scale using an IV approach. Only the first shock affects the first variable. As argued before, we can use a standard heteroscedasticity-based identification scheme and, therefore, the standard IV estimator (see Rigobon and Sack, 2004, Lewis, 2022). The instrument, the first, as well as the second stage read:

$$\begin{aligned} Z_{1t} &= \left[\mathbf{1}(t \in P)\frac{T}{T_P} - \mathbf{1}(t \in C)\frac{T}{T_C}\right] y_{1t} \\ y_{1t} &= \alpha_1 + \beta_1 Z_{1t} + u_{1t} \\ y_{it} &= \alpha_i + \tilde{\Psi}_{i1}\hat{y}_{1t} + e_{it} \end{aligned}$$

where $\mathbf{1}(t \in X)$ denotes an indicator function that equals one if the condition in parentheses is true and zero otherwise, and T, T_P, and T_C are the number of total, policy event and control days, respectively. In addition, Z_{1t} and $\hat{y}_{1t} = \hat{\beta}_1 Z_{1t}$ denote the instrument, as well as the first-stage prediction. Finally, α_i, β_i, and $\tilde{\Psi}_{ij}$ are regression parameters and u_{it} and e_{it} are regression residuals.

Three comments are in order. First, the instrument is uncorrelated with v_t, and therefore e_{it}, because ε_{1t} occurs only during policy event days (see e.g. Lewis, 2022, for a detailed discussion). Second, the instrument is also uncorrelated with structural shocks $e = 2, \ldots, E$ because the variance of y_{1t} changes on event days only due to ε_{1t} (recursive zero restriction). Finally, we estimate the impulse responses only up to a scale because the impact on y_{1t} is normalized to unity, that is, $\tilde{\Psi}_{i1} = \Psi_{i1}/\Psi_{11}$.

For structural shock 2, the standard heteroscedasticity-based IV estimator will fail. The variance of y_{2t} changes on policy event days due to the first and second structural shocks. If we fail to control for shock 1, the instrument will be correlated with the error term.

We propose to estimate the impact of structural shock 2 by controlling for variation in y_{1t} caused by shock 1. We can construct an additional instrument using the second variable and then use both to estimate Ψ_{i2}. The instrument, the first, as well as the second stage, then read:

$$\begin{aligned} Z_{2t} &= \left[\mathbf{1}(t \in P)\frac{T}{T_P} - \mathbf{1}(t \in C)\frac{T}{T_C}\right] y_{2t} \\ y_{1t} &= \alpha_1 + \beta_{11} Z_{1t} + \beta_{12} Z_{2t} + u_{1t} \\ y_{2t} &= \alpha_2 + \beta_{21} Z_{1t} + \beta_{22} Z_{2t} + u_{2t} \\ y_{it} &= \alpha_i + \tilde{\Psi}_{i2} \hat{y}_{2t} + \tilde{\Psi}_{i1} \hat{y}_{1t} + e_{it} \end{aligned}$$

On event days, y_{2t} and Z_{2t} are correlated with shocks ε_{1t} and ε_{2t}. However, by including $\hat{y}_{1t}$ as a control in the second stage, the instrument will not be correlated with the error term.[10]

More generally, we can recursively estimate the impact matrix of E-dimensional structural shocks using the following instruments, first and second stages:

$$\begin{aligned} Z_{et} &= \left[\mathbf{1}(t \in P)\frac{T}{T_P} - \mathbf{1}(t \in C)\frac{T}{T_C}\right] y_{et}, \quad e = 1, \dots, E \\ y_{et} &= \alpha_e + \sum_{j=1}^{e} \beta_{ej} Z_{jt} + u_{et}, \quad e = 1, \dots, E \\ y_{it} &= \alpha_i + \sum_{j=1}^{e} \tilde{\Psi}_{ij} \hat{y}_{jt} + e_{it}, \quad e = 1, \dots, E\ , \ i = 1, \dots, N \end{aligned}$$

In this framework, we have multiple endogenous regressors and instruments for $E > 1$. Therefore, we can use the heteroscedasticity- and autocorrelation-robust (HAR) test for weak instruments proposed by Lewis and Mertens (2022). This test is also helpful in examining whether there are multiple dimensions on policy event days because the corresponding IV estimates will suffer from a weak instruments problem.[11]

2.2 Extraction of unobserved shocks

It is helpful to extract the unobserved shock series to compare the results to existing multi-dimensional high-frequency surprises or use the shocks in low-frequency IV-LP or IV-SVAR models. We propose to obtain the linear minimum MSE prediction of the unobserved monetary policy shocks using the Kalman filter.[12]

[10] Again, we identify Ψ_{i2} only up to a scale because the initial response on y_{2t} is normalized to unity, that is, $\tilde{\Psi}_{i2} = \Psi_{i2}/\Psi_{22}$.

[11] Appendix B illustrates this point using simulated data.

[12] We thank Mark Watson for guiding us in this direction.

Recall that we can obtain the least squares forecast of a state vector in a state-space model using the Kalman filter. Following Hamilton (1994), p. 372, consider the state-space system:

$$
\begin{aligned}
\xi_t &= F\xi_{t-1} + \nu_t \\
y_t &= A'x_t + H'\xi_t + w_t \\
Q &= \mathbb{E}[\nu_t \nu_t'] \\
R &= \mathbb{E}[w_t w_t']
\end{aligned}
$$

where ν_t and w_t are serially uncorrelated shock vectors. In addition, they are uncorrelated with each other at all leads and lags. y_t is a vector of observed variables, ξ_t a vector of unobserved states, and x_t is a vector of exogenous or predetermined variables. F, A', H' are conformable matrices of parameters.

The unobserved states represent the unobserved structural shocks. We can use formulae based on the Kalman filter to obtain predictions of these states.[13] We are ultimately interested in $\xi_{t|t} \equiv \mathbb{E}[\xi_t|\Omega_t]$, the linear projection of the unobserved states on the information set $\Omega_t \in \{y_t, y_{t-1}, \ldots, x_t, x_{t-1}, \ldots\}$ on day t. The Kalman filter recursively provides the linear projection in the state-space model (Hamilton, 1994, p. 379):

$$\xi_{t|t} = \xi_{t|t-1} + P_{t|t-1}H(H'P_{t|t-1}H + R)^{-1}(y_t - A'x_t - H'\xi_{t|t-1})$$

where $\xi_{t|t-1}$ is the one-step-ahead prediction of the unobserved states and $P_{t|t-1}$ is the one-step-ahead prediction MSE matrix.

The unobserved shocks in Equation (1) are serially uncorrelated. Therefore, $F = \mathbf{0}$, where $\mathbf{0}$ is a conformable matrix of zeros. It follows that $\xi_{t|t-1} = \mathbf{0}$, the unobserved states are unpredictable based on past data, and $P_{t|t-1} = Q$, the MSE matrix corresponds to the unconditional variance of the shocks in the state equation (see Hamilton, 1994, p. 380). In addition, $H'QH + R = \mathbb{E}[(y_t - A'x_t)(y_t - A'x_t)'] \equiv \Sigma$ is the variance-covariance matrix of the residuals between the data and the predetermined variables. Therefore, we can predict the unobserved states based on Q, H and Σ, and data available at time t:

$$\xi_{t|t} = QH\Sigma^{-1}(y_t - A'x_t)$$

[13] Using the Kalman smoother is unnecessary because the unobserved states are serially uncorrelated. Therefore, future observations do not contain useful information about the current unobserved state.

We can cast the model in Equation (1) in state-space form. For $t \in P$ we obtain:

$$
\begin{aligned}
\xi_t &= \varepsilon_t \\
y_t &= \Psi \xi_t + \Gamma v_t \\
Q &= \mathbb{E}[\varepsilon_t \varepsilon_t'] = \Sigma_\varepsilon \\
R &= \mathbb{E}[\Gamma v_t v_t' \Gamma']
\end{aligned}
$$

Therefore, the prediction formula for the unobserved shocks reads:

$$
\varepsilon_{t|t} = \Sigma_\varepsilon \Psi' \Sigma^{-1} y_t \quad \text{for } t \in P
$$

This implies that we can compute the minimum MSE prediction of the shocks if we know the variance-covariance matrix of the structural shocks (Σ_ε), the impact matrix (Ψ), the variance-covariance matrix $\Sigma \equiv \Psi \Sigma_\varepsilon \Psi' + R$, and the actual data.[14]

Five comments are in order. First, we can estimate each column in Ψ up to a scale using the IV estimator.[15] Second, we can estimate $\Sigma = \Psi \Sigma_\varepsilon \Psi' + R$ as the variance-covariance matrix of y_t on policy event days ($t \in P$). Third, although we do not directly observe the diagonal matrix Σ_ε, it only affects the scale of the shocks. Therefore, we can arbitrarily normalize it to an identity matrix to extract the shocks up to a scale. Fourth, if we include lags of the dependent variables or other exogenous variables in the IV regressions to estimate Ψ, we can replace y_t with the residuals ($y_t - A'x_t$) and also compute the variance-covariance matrix Σ based on these residuals.[16] Finally, the number of variables in y_t, N, should be relatively large compared to the number of unobserved shocks, E, to obtain an accurate prediction.[17]

Our approach has several advantages. In particular, it provides the linear minimum MSE prediction of the shocks. In addition, at least with a sufficiently high number of variables in y_t, the multiple dimensions are orthogonal to each other and other shocks affecting y_t on policy event days. This is not guaranteed using existing approaches. Bu et al. (2021) propose a two-step approach to estimate a one-dimensional unobserved monetary policy shock. First, they estimate the causal impact of a monetary policy shock on interest rate changes across the

[14]The MSE of this updated projection reads $P_{t|t} = \Sigma_\varepsilon - \Sigma_\varepsilon \Psi' \Sigma^{-1} \Psi \Sigma_\varepsilon$ (see Hamilton, 1994, p. 380).

[15]The fact that we can estimate Ψ only up to a scale only affects the scale of the extracted shocks. Suppose $\tilde{\Psi} = \Psi S$, where S is a diagonal matrix scaling each column by a different scalar. $\tilde{\varepsilon}_t = \Sigma_\varepsilon (\Psi S)' \Sigma^{-1} y_t = \Sigma_\varepsilon S' \Psi' \Sigma^{-1} y_t = \tilde{\Sigma}_\varepsilon \Psi' \Sigma^{-1} y_t$, where $\tilde{\Sigma}_\varepsilon = \Sigma_\varepsilon S'$ is again a diagonal matrix. Therefore, using $\tilde{\Psi}$ instead of Ψ to extract the shocks corresponds only to a different scaling of the variance of the structural shocks.

[16]Even if no lags are included, the regressions will likely include a constant. In this case, y_t is replaced by the demeaned data.

[17]We illustrate this point using simulated data in Appendix B.

term structure via the heteroscedasticity-IV approach by Rigobon and Sack (2004). Second, in the spirit of Fama and MacBeth (1973), they perform a cross-sectional regression of interest rate changes along the term structure on the impact vector for every day with an FOMC decision. The OLS coefficients of these regressions are then proportional to the underlying unobserved monetary policy shocks.

However, the regression-based approach requires stronger assumptions than the Kalman-filter approach. Suppose we know each column of Ψ up to a scale, $\tilde{\Psi} = \Psi S$, where S is a conformable diagonal matrix with scaling factors. The OLS estimator of y_t on $\tilde{\Psi}$ on day t reads:

$$\hat{\varepsilon}_t = (\tilde{\Psi}'\tilde{\Psi})^{-1}\tilde{\Psi}' y_t = S^{-1}\varepsilon_t + (\tilde{\Psi}'\tilde{\Psi})^{-1}\tilde{\Psi}'\Gamma v_t \tag{2}$$

The OLS estimate depends proportionally on the true shocks and an error term. To verify whether OLS is an unbiased estimator of the shocks up to a scale, we can compute the expectation conditional on the information set on day t:

$$\mathbb{E}[\hat{\varepsilon}_t|\Omega_t] = S^{-1}\varepsilon_t + \mathbb{E}[(\tilde{\Psi}'\tilde{\Psi})^{-1}\tilde{\Psi}'\Gamma v_t|\Omega_t]$$

OLS is an unbiased estimator only if the error term $(\tilde{\Psi}'\tilde{\Psi})^{-1}\tilde{\Psi}'\Gamma v_t$ is zero in expectation. Using the law of total expectations, we can show that this is the case if $\mathbb{E}[\Gamma|\tilde{\Psi}, \Omega_t] = \mathbf{0}$, that is, we need an orthogonality condition between $\tilde{\Psi}$ and Γ. This implies that the cross-sectional variation in the responses of the dependent variables to the structural shocks ($\tilde{\Psi}$) has to be orthogonal to the variation in the responses to other shocks (Γ).

If this assumption is violated, the shock estimates will suffer from an unobserved variables bias because we fail to control for variation in Γ in the cross-sectional regression.[18] If the orthogonality assumption does not hold, the extracted shocks are contaminated by other shocks occurring on policy event days (v_t). In addition, it may introduce a correlation between the multiple dimensions of structural shocks.[19] Whether the assumption holds depends on the application and is, therefore, an empirical question.

[18] In monetary policy applications, for example, the assumption is violated if long-term interest rates respond less strongly than short-term interest rates to both monetary policy and productivity shocks. Whether this assumption is violated in practice is an empirical question, and we will compare the resulting shock series with both approaches to existing multi-dimensional high-frequency surprises.

[19] In Appendix B, we show, using simulated data, that even if the orthogonality assumption is satisfied, the Kalman filter yields a more accurate prediction of the underlying shocks.

3 Empirical application

We use the two-step approach to analyze how various dimensions of US monetary policy affect the US dollar exchange rate. The US application allows us to compare the results to the multi-dimensional high-frequency surprises by Swanson (2021). In addition, using the exchange rate as an outcome is helpful because the identification scheme does not directly constrain its impulse response.

3.1 Data

In the baseline specification, we set:[20]

$$\begin{aligned} y_t \;=\; & [i_t^{3m},\, i_t^{2y},\, i_t^{10y-2y},\, \text{neer}_t, \\ & i_t^{6m},\, i_t^{1y},\, i_t^{3y},\, i_t^{5y},\, i_t^{7y},\, i_t^{30y}, \\ & \text{AAA}_t,\, \text{BAA}_t,\, \text{SP500}_t,\, \text{NASDAQ}_t,\, \text{OIL}_t,\, \text{COM}_t]' \end{aligned}$$

The variables in the first row (3M and 2Y interest rates, the 10Y-2Y term spread, and the nominal effective exchange rate) are the main variables of interest. We will impose recursive zero restrictions on the first three variables to disentangle multiple dimensions of monetary policy shocks and use the exchange rate as the main outcome of interest.

Then, we follow Bu et al. (2021) and use interest rate data along the term structure to extract the monetary policy shocks (6M, 1Y, 3Y, 5Y, 7Y, and 30Y rates).[21] In addition, we include two corporate bond spreads (AAA and BAA), as well as financial market prices that may respond to monetary policy announcements (S&P 500 and NASDAQ stock price indices, WTI oil prices, and a commodity futures index).[22]

The exchange rate and financial market prices are included as log-changes multiplied by 100, so the cumulative impulse responses are measured in percent. All interest rates and spreads are included in first differences, so the cumulative impulse responses are measured in percentage points.[23]

As monetary policy events, we obtained the 323 FOMC announcement dates for 1988–2019

[20] A detailed description of the data sources is given in Appendix C.

[21] Bu et al. (2021) use data from Gürkaynak et al. (2007) featuring more maturities. We use the same data as they do in a robustness test. The baseline uses Treasury bill yields from the Board of Governors because we know the exact timestamp of the daily data.

[22] The exchange rate data is recorded at noon EST. Most other series are recorded at market close. There are very few missing observations that we linearly interpolate.

[23] As these (log-)changes are likely relatively independent over time, the baseline specification does not include lags of the dependent variables. All specifications include constant terms.

from Bauer and Swanson (2022) and extended them for 2020–2022. We end up with 344 FOMC announcement dates.[24] In the baseline, we use scheduled and unscheduled announcements as policy event days and all other working days as control days.[25]

3.2 Identifying assumptions

The ordering of variables in y_t implies that we impose recursive zero restrictions on the response of the 3M interest rate, 2Y interest rate, and 10Y-2Y term spread to identify $E = 3$ dimensions of monetary policy shocks. Therefore, the first shock potentially affects all variables and increases the variance of the 3M interest rate.[26] We, therefore, interpret this dimension as a shock to the Fed's short-term interest rate target. The second shock does not immediately impact the 3M interest rate but has to increase the variance of the 2Y interest rate. As this shock affects medium-term interest rates, we interpret it as a shock to expectations about the Fed's interest rate path.[27] Finally, the third shock does not immediately impact 3M and 2Y interest rates. However, the shock has to increase the variance of the term spread.[28] We interpret this shock as a non-conventional monetary policy dimension because we assume that it affects financial markets via the term premium.

Existing approaches extracting multi-dimensional surprises from high-frequency data use similar restrictions. The target surprise by Altavilla et al. (2019) is a factor estimated on changes in high-frequency data around the ECB's press release and primarily loads on short-term interest rates (although this is not imposed in the model). In addition, they impose that the path and LSAP surprises are orthogonal to the short-term interest rate. Finally, the Quantitative Easing (QE) shock is assumed to have a small variance before the financial crisis. Similarly, Swanson (2021) imposes that path, and large-scale asset purchase (LSAP) surprises do not affect the current FFR. In addition, he imposes that the LSAP shock is as small as possible before the effective lower bound on short-term interest rates becomes binding.

[24] 284 are regularly scheduled FOMC meetings. Unscheduled meetings include events before 1994, as the FOMC only started then to announce its decisions for the FFR target after each FOMC meeting. In addition, we have 60 FOMC (intermeeting) conference calls that we treat as unscheduled events. Note that these numbers differ from Swanson and Jayawickrema (2023) because we use daily data and define events daily. In contrast, they sometimes have two events on one day. We refer to their paper for more details about the selection of event dates.

[25] We control for various contaminating events, such as speeches or data releases, in a series of robustness tests.

[26] If the first shock does not increase the variance of the 3M rate, the heteroscedasticity-based instrument suffers from a weak instruments problem. We use the 3M rate instead of the FFR to identify the target shock because FOMC meetings do not occur daily. A target shock should be expected to persistently change the overnight rate and, therefore, the 3M rate. However, we also report results using the Federal Funds Rate in the robustness section.

[27] To impose the zero restriction in the IV regressions, we therefore include the first and second instruments and have two endogenous variables.

[28] To impose the zero restriction in the IV regressions, we therefore include all three instruments and have three endogenous variables.

While high-frequency approaches assume that no other shocks occur in a narrow window around policy announcements, our main identifying assumption is that policy events occur only on pre-determined days. For example, suppose the Federal Reserve is more likely to take decisions during economic crises. In that case, the variance of other shocks is different between policy event days and control days. Focusing on scheduled FOMC decisions fulfills this requirement. As our baseline includes scheduled and unscheduled announcements, we will distinguish between scheduled and unscheduled FOMC decisions in a robustness test.

Furthermore, we assume that other shocks occur randomly across policy and control days. This assumption is violated, for example, if the FOMC schedules its meetings after economic data is released. If this is the case, the variance on control days is affected by data release surprises that do not occur on policy event days. Therefore, other shocks are not randomly distributed between policy events and control days. In addition, other central banks may schedule their meetings briefly after the FOMC to consider US policy surprises in their decisions. We account for this issue by excluding other events from the control days in the robustness checks.

3.3 Comparison of monetary policy shocks

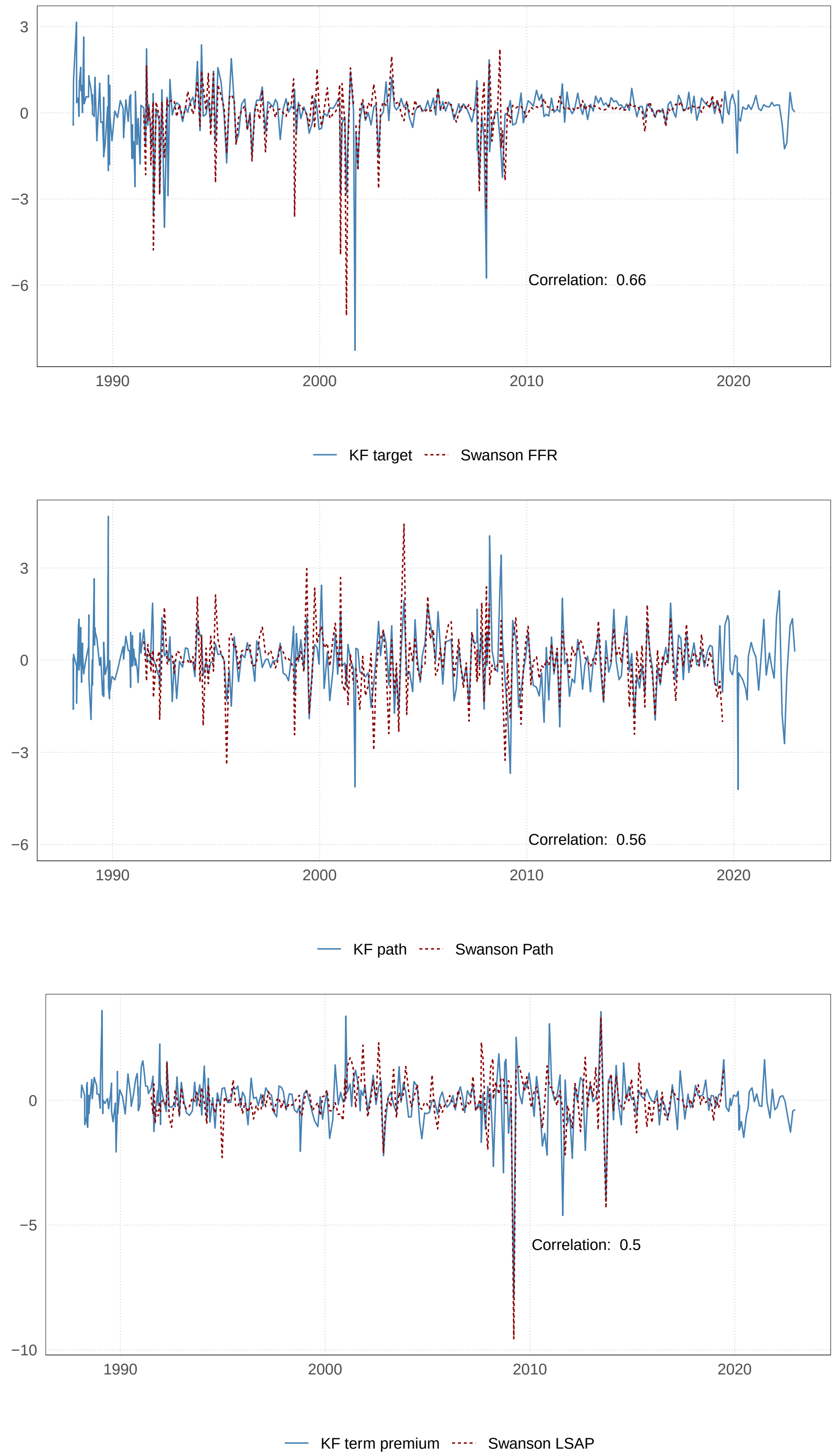
How does our two-step procedure perform in practice? Figure 1 compares the shocks predicted by the Kalman filter (KF) with the high-frequency surprises by Swanson (2021). The three dimensions show a relevant correlation with the high-frequency counterparts (between 0.50 and 0.66).[29] Although the correlations are lower than unity, they are sizeable, given that our approach relies on a different data set and identification strategy. Compared to the high-frequency surprises, the KF target shock exhibits a slightly larger variability starting in late 2008, when short-term rates were constrained by the effective lower bound (ELB). This may be related to the fact that we use a 3M rate rather than the FFR. Nevertheless, the variance of target shocks declined substantially over time, especially at the ELB. The exceptions are severe economic crises. In addition, the variance of the term premium shock increased substantially during the period at the ELB, even though we did not impose this restriction. Meanwhile, path shocks were quite volatile throughout the sample.

Recall that the Kalman filter yields the linear minimum MSE prediction of the unobserved shocks. Alternatively, we can use Fama and MacBeth (1973) (FM) regressions to estimate the shocks under stronger assumptions. Table 1 shows that the correlations of the FM shocks with the corresponding high-frequency surprises are relatively low.[30] In addition, the various

[29] Figure 9 in Appendix E shows the same graphs aggregated to monthly frequency.

[30] This is in line with Brennan et al. (2024), showing that the correlation between various approaches, including the FM approach, to measure monetary policy shocks is relatively low.

Fig. 1: Heteroscedasticity-based shocks and high-frequency surprises



Notes: Daily shocks extracted using the Kalman filter (KF) and high-frequency surprises by Swanson (2021). All series normalized to a mean of zero and a standard deviation of one.

Tab. 1: Correlation matrix of heteroscedasticity-based shocks and high-frequency surprises

	KF target	KF path	KF term premium
KF target	1.00		
KF path	0.06	1.00	
KF term premium	-0.03	0.20	1.00
Swanson FFR	0.66	0.05	-0.11
Swanson path	0.03	0.56	0.17
Swanson LSAP	-0.11	0.16	0.50

	FM target	FM path	FM term premium
FM target	1.00		
FM path	-0.54	1.00	
FM term premium	0.43	-0.76	1.00
Swanson FFR	0.30	-0.16	0.28
Swanson path	0.21	-0.05	0.14
Swanson LSAP	-0.09	-0.02	0.01

Notes: Correlation of shocks extracted using the Kalman filter (KF) and Fama-MacBeth regressions (FM) with high-frequency surprises by Swanson (2021).

monetary policy dimensions are correlated when using FM regressions. This may indicate that, in this specific application, the FM shocks are contaminated by other shocks affecting financial market variables. By contrast, the various dimensions of KF shocks show a lower cross-correlation in absolute value.

Tab. 2: Relationship between heteroscedasticity-based shocks and high-frequency surprises

	SW FFR	SW path	SW LSAP	SW FFR	SW path	SW LSAP
	(1)	(2)	(3)	(4)	(5)	(6)
KF target	0.74***	0.07	−0.09			
	(0.12)	(0.10)	(0.07)			
KF path	0.11	0.60***	0.05			
	(0.07)	(0.10)	(0.10)			
KF term premium	−0.09	0.04	0.45***			
	(0.07)	(0.06)	(0.17)			
FM target				0.27***	0.22***	−0.12*
				(0.09)	(0.07)	(0.07)
FM path				0.19	0.21	−0.05
				(0.26)	(0.22)	(0.11)
FM term premium				0.32	0.19	0.02
				(0.28)	(0.21)	(0.11)
Constant	0.01	−0.01	0.01	0.04	0.02	−0.002
	(0.04)	(0.05)	(0.05)	(0.05)	(0.06)	(0.06)
N	241	241	241	241	241	241
R^2	0.45	0.32	0.26	0.13	0.06	0.01
Adjusted R^2	0.44	0.32	0.25	0.12	0.05	−0.001

Notes: OLS regressions of the high-frequency surprises by Swanson (2021) (SW) on shocks extracted using the Kalman filter (KF) and Fama-MacBeth regressions (FM). All series have been normalized to have zero mean and variance 1. Significance levels are given by *p<0.1; **p<0.05; ***p<0.01. HAC robust standard errors are in parentheses.

Because the FM shocks are correlated, we also perform multivariate regressions of Swanson's high-frequency surprises on the KF and FM shocks, respectively. Table 2 shows a significant relationship between the KF target, path, and term premium shocks with their high-frequency counterparts. However, the KF path and term premium shocks, for example, do not display a significant relationship with the high-frequency FFR shock. We expect this if our approach yields i.i.d. monetary policy shocks that reasonably approximate the corresponding high-frequency surprises. Meanwhile, the FM target shock is positively related to high-frequency FFR and path surprises. The other dimensions do not exhibit a significant relationship with any of the high-frequency shocks.

3.4 Comparison of impulse response functions

Next, we turn to daily impulse response functions, focusing on the US dollar exchange rate. While the correlation with the high-frequency surprises is imperfect, this may stem from classical measurement errors that do not bias the impulse responses in an IV-LP approach. This section therefore investigates whether we obtain qualitatively similar economic effects using heteroscedasticity-based and high-frequency approaches.[31]

Figure 2 shows the impulse responses to a target, path, and a term premium shock, respectively. The blue solid lines show the local projection responses using the heteroscedasticity-based instruments with 90% and 95% confidence intervals. The red dotted lines show responses using the corresponding high-frequency surprises as instruments.[32]

Focusing on the heteroscedasticity-based identification scheme, a target shock increases the 3M and 2Y interest rates and lowers the term spread. This is as expected if a target shock persistently, but still temporarily, increases the short-term interest rate. All exchange rates are defined as one USD in terms of foreign currency. A decrease in the exchange rate is, therefore, an appreciation of the USD. We observe an immediate appreciation of the US dollar, which vanishes after ten working days. The response is slightly delayed. However, the delay is not statistically significant.[33]

A path shock tightening does not affect the 3M interest rate on impact due to the recursive zero restriction. In addition, the 3M rate and the term spread hardly respond. However, we observe a rapid and statistically significant exchange rate appreciation. The exchange rate remains persistently stronger for slightly less than 20 working days. Finally, the term premium shock raises the term spread for up to ten working days while not significantly affecting the 3M or 2Y interest rates, even after the imposed zero impact effect. The exchange rate also appreciates. Although the response is slightly delayed, it has reached the trough after about five working days.

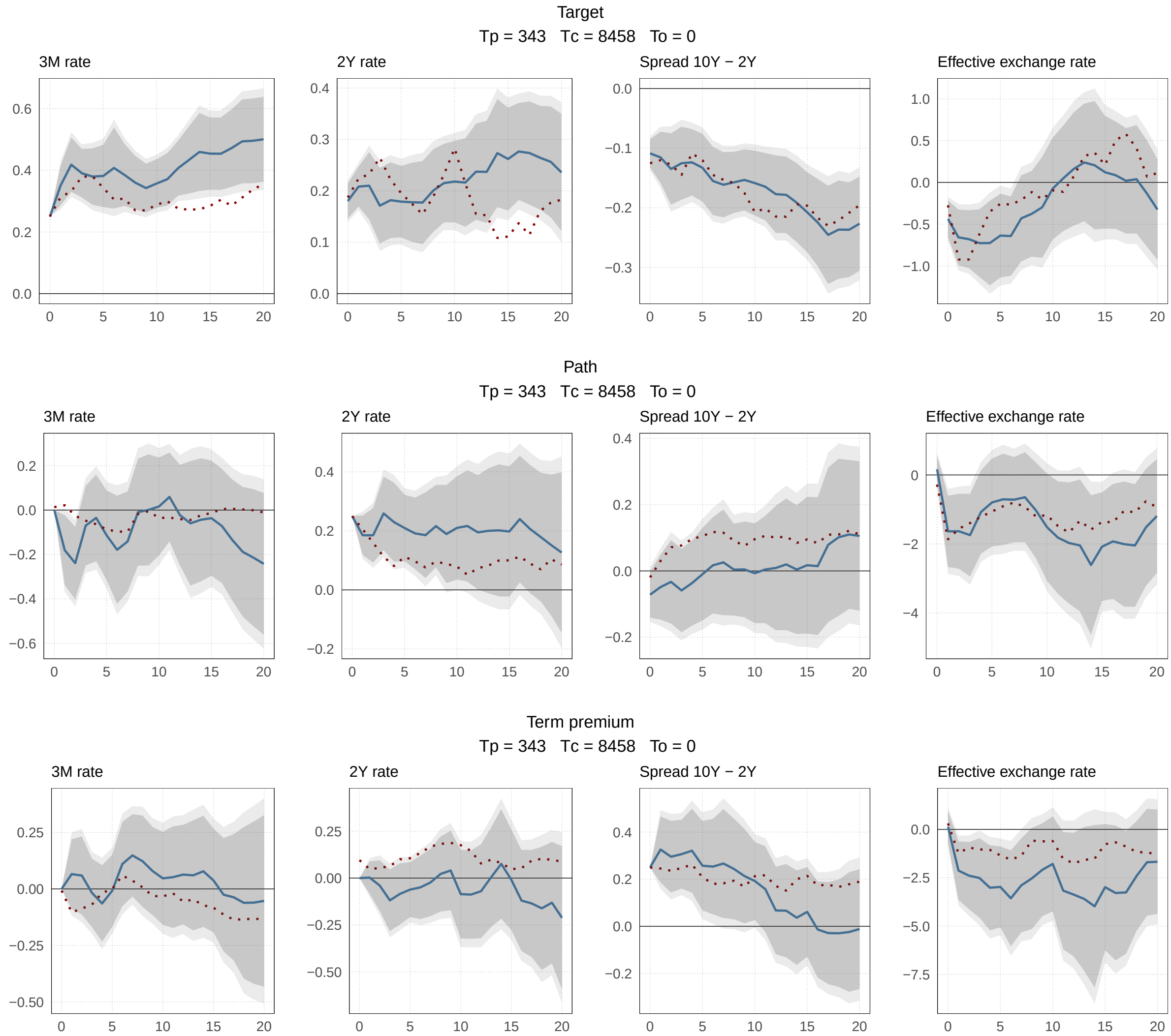
The responses are very similar for the high-frequency surprises. Although we do not impose the zero restriction, the 3M rate (3M and 2Y rate) does not respond much to the path (LSAP) shock. The target shock also leads to a decline in the term spread, while the LSAP shock leads

[31]We focus on IV-LP models throughout the paper, as IV-SVARs may yield too narrow confidence intervals if the lag length of the model is misspecified (Montiel Olea et al., 2024).

[32]These responses use the FFR, path, and LSAP high-frequency surprises by Swanson (2021) to instrument the 3M rate, 2Y rate, and 10Y-2Y term spread, respectively. Figure 10 in Appendix E shows these responses with confidence intervals. In addition, it shows the responses when additionally imposing the same zero restrictions as in the heteroscedasticity-based identification scheme. Qualitatively, the results are identical.

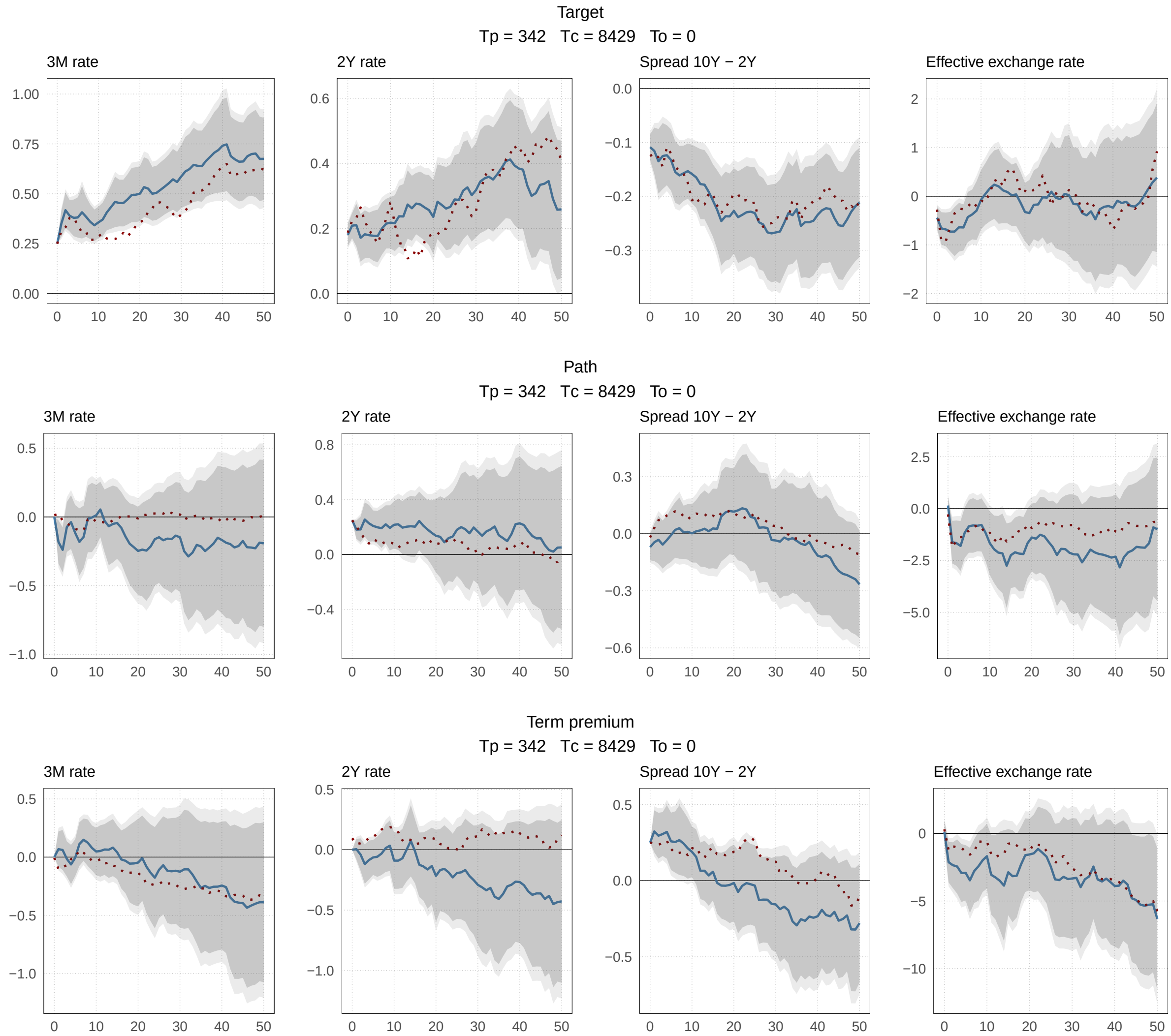
[33]This may be partly related to the fact that exchange rates are recorded at noon, so it takes one working day until the full effect of the monetary policy shock is recorded in the data.

Fig. 2: Impulse responses to orthogonal monetary policy shocks



Notes: Impulse responses to monetary policy shocks (target, path, and term premium). The responses are normalized to a 25 bp increase in the 3M rate, 2Y rate, and 10Y - 2Y spread, respectively. The blue solid lines show the heteroscedasticity-based responses. The red dotted lines are the responses to the high-frequency surprises by Swanson (2021). The horizontal axis measures working days (excluding weekends and holidays). The models are estimated in first (log-)differences, but the impulse responses are cumulated. Therefore, all interest rate responses are measured in percentage points and the exchange rate responses are measured in percent. 90% and 95% confidence intervals are based on HAC-robust standard errors. T_p, T_c, T_o denote the number of policy event days, control days, and other days, respectively.

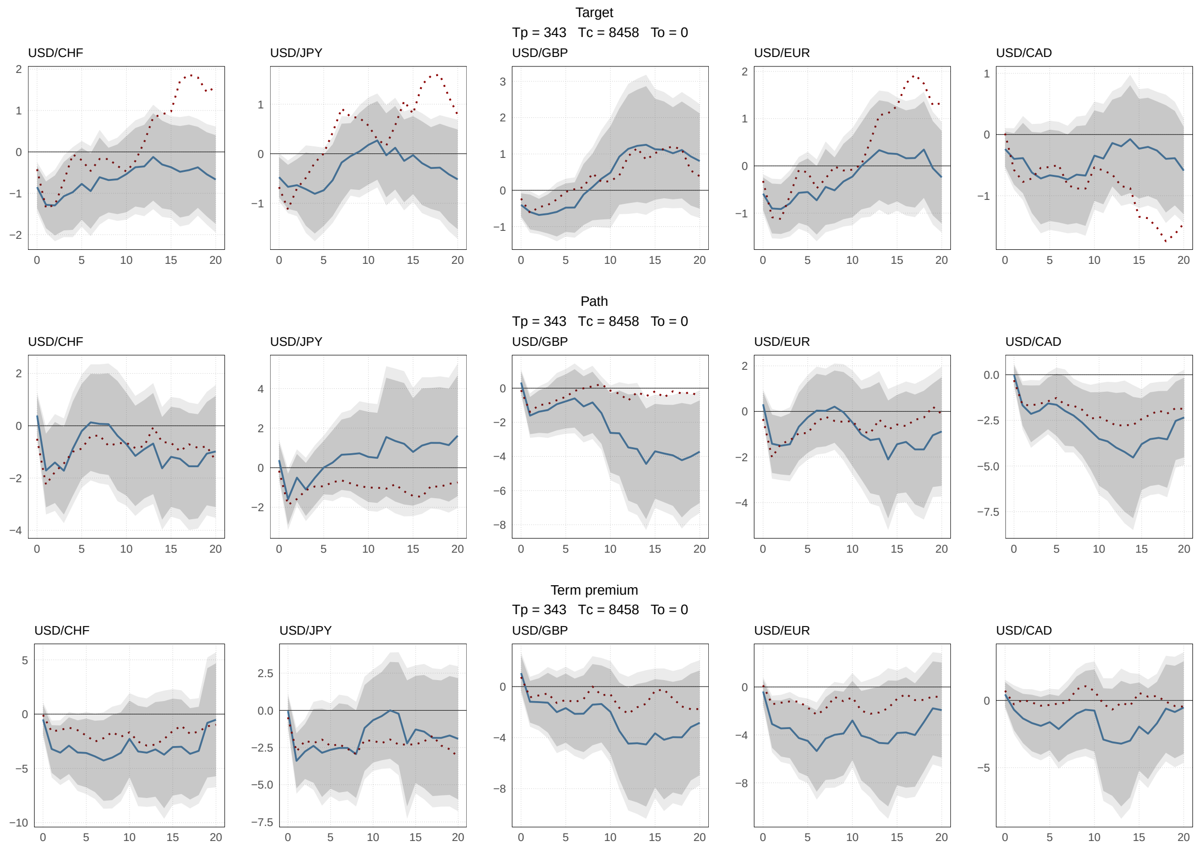
Fig. 3: Impulse responses in the longer run



Notes: Impulse responses to monetary policy shocks (target, path, and term premium) up to a horizon of 50 working days. The responses are normalized to a 25 bp increase in the 3M rate, 2Y rate, and 10Y - 2Y spread, respectively. The blue solid lines show the heteroscedasticity-based responses. The red dotted lines are the responses to the high-frequency surprises by Swanson (2021). The horizontal axis measures working days (excluding weekends and holidays). The models are estimated in first (log-)differences, but the impulse responses are cumulated. Therefore, all interest rate responses are measured in percentage points and the exchange rate responses are measured in percent. 90% and 95% confidence intervals are based on HAC-robust standard errors. T_p, T_c, T_o denote the number of policy event days, control days, and other days, respectively.

to an increase. Finally, for all three shocks, the exchange rate almost immediately appreciates. Only for the LSAP shock is the exchange rate response smaller and less significant than the heteroscedasticity-based response. Overall, this suggests that, at least in the daily data, there is no delayed overshooting for any of the monetary policy dimensions and identification schemes.

Fig. 4: Impulse responses of bilateral exchange rates



Notes: Impulse responses of bilateral exchange rates to monetary policy shocks (target, path, and term premium). A separate model is estimated for every bilateral exchange rate. The responses are normalized to a 25 bp increase in the 3M rate, 2Y rate, and 10Y - 2Y spread, respectively. The blue solid lines show the heteroscedasticity-based responses. The red dotted lines are the responses to the high-frequency surprises by Swanson (2021). The horizontal axis measures working days (excluding weekends and holidays). The models are estimated in first (log-)differences, but the impulse responses are cumulated. Therefore, the exchange rate responses are measured in percent. 90% and 95% confidence intervals are based on HAC-robust standard errors. T_p, T_c, T_o denote the number of policy event days, control days, and other days, respectively.

To examine whether we find a delayed overshooting puzzle at longer horizons, Figure 3 shows the baseline results for a horizon of up to 50 working days. For the target shock, the exchange rate responses remain around zero at longer horizons. The point estimate is more persistent for the path and term premium shocks, but the responses are quite imprecisely estimated. This suggests that there is no delayed overshooting puzzle using daily data. But, given the large estimation uncertainty, other patterns are also possible, in line with Faust et al. (2003).

We also examined the impulse response of bilateral exchange rates, where a separate model was estimated for every currency pair (see Figure 4). Overall, the responses are similar but less precisely estimated than for the nominal effective exchange rate. This is expected as bilateral exchange rates are more noisy than the trade-weighted average. We also find that the responses to the target shock are mostly less persistent than those to the path and term premium shocks. In particular, the target shock responses reverse to zero after a few working days.

3.5 Weak instruments tests

Heteroscedasticity-based and high-frequency identification schemes use different approaches to deal with background noise and other contaminating shocks. However, both approaches may suffer from weak instruments problems. We, therefore, use a HAR test for weak instruments in the presence of multiple endogenous regressors and instruments proposed by Lewis and Mertens (2022). This exercise allows us to verify whether the heteroscedasticity-based instruments perform better or worse concerning the weak instruments problem than their high-frequency counterparts. In addition, they provide evidence on whether there are multiple dimensions of monetary policy in the first place. Table 3 shows the results for three exercises. The first column shows results only based on the instrument associated with a target shock, where the 3M interest rate is the endogenous variable. The second column corresponds to the second row of the model in Equation (1), where we have two endogenous regressors (the 3M and 2Y interest rates) and two instruments. Finally, the third column shows results for the third equation with three endogenous regressors (3M and 2Y interest rate and the 10Y-2Y term spread) and three instruments.

The first panel shows that the heteroscedasticity-based target instrument passes the weak instruments test. The test statistics are below the critical values for the path and term premium shocks.[34] This does not imply that the heteroscedasticity-based identification scheme performs worse, as the second panel shows. We find a qualitatively similar pattern for the high-frequency surprises. One possibility is that both approaches yield weak instruments, but because they capture different variation, combining them may alleviate the problem. The last panel shows that even when combining the corresponding instruments, identification remains weak for the path and term premium shocks.[35]

[34] Wright (2012) also does not reject the null hypothesis of one monetary policy dimension at the ELB.

[35] It would be interesting to examine whether a heteroscedasticity-based identification scheme using intraday changes would perform better because this would more strongly reduce background noise. We leave this as an interesting avenue for future research.

Tab. 3: Weak instruments tests

	Target	Path	Term premium
Test statistic	43.9	19.3	4.9
Critical value	23.1	27.3	31.2

	Swanson FFR	Swanson path	Swanson LSAP
Test statistic	23.2	16.7	10.5
Critical value	23.1	24.1	27.5

	Joint target/FFR	Joint path	Joint term premium/LSAP
Test statistic	34.0	14.1	12.1
Critical value	22.4	21.5	21.7

Notes: Weak instruments tests allowing for heteroscedasticity and autocorrelation with multiple endogenous regressors (Lewis and Mertens, 2022). The endogenous regressors are the 3M rate, 2Y rate, and 10Y - 2Y spread, respectively. The tests shown in the first, second, and third columns are based on one, two, and three instruments. We impose the recursive zero restrictions used to disentangle the multiple dimensions of monetary policy shocks with heteroscedasticity. The first panel shows the results for the heteroscedasticity-based instruments. The second panel shows the results based on the high-frequency surprises by Swanson (2021). The last panel includes the instruments from both the heteroscedasticity-based and high-frequency approaches. The significance level is set to 5% and the tolerance level to 10%.

3.6 Robustness tests

We conducted a range of robustness tests.[36] The results are summarized in Table 4, which shows correlations between the target, path, and term premium shocks with the corresponding high-frequency surprises, and Table 5, which shows the corresponding weak instruments tests.[37]

Tab. 4: Robustness tests: Correlation with high-frequency surprises

	Target	Path	Term premium
Baseline	0.66	0.56	0.50
FFR instead of 3M rate	0.46	0.46	0.50
Main variables with $P = 4$ lags	0.66	0.56	0.49
Additional control variables with $P = 2$ lags	0.66	0.55	0.50
Only use term structure data	0.58	0.50	0.51
Include speeches as events	0.66	0.56	0.48
Exclude contaminating events	0.74	0.55	0.19
Tuesday and Wednesday as control	0.66	0.58	0.49
Exclude unscheduled meetings	0.50	0.52	0.55
Effective lower bound period	0.03	0.29	0.48
Swanson (2021) estimation sample	0.67	0.61	0.48

Notes: Correlation between the target, path, and term premium shocks extracted using the Kalman filter and the corresponding high-frequency surprises by Swanson (2021).

We change the model specification in various ways. We use the FFR instead of the 3M interest rate to identify the target surprise. While the correlation with the high-frequency surprise falls for the target and path shocks, it remains unchanged for the term premium shock. Then we estimate specifications with four lags of the main dependent variables (3M rate, 2Y rate, 10Y-2Y term spread, effective exchange rate) and two lags of the main variables and additional control variables (corporate bond spread, stock price index, oil prices, commodity prices) as controls. The results remain virtually unchanged, suggesting that there was little serial correlation in the first differences of the financial market variables. Then, we follow Bu et al. (2021) and only use the term structure data by Gürkaynak et al. (2007) to extract the shocks. The correlation falls relative to the baseline, suggesting that the responses of stock prices and other financial market

[36] The additional data is discussed in Appendix C.
[37] For brevity, we do not report the impulse response functions.

variables comprise useful information to extract the unobserved monetary policy shocks.[38] Finally, we include relevant speeches identified with a topic-modeling approach as policy events.[39] The results remain similar.[40]

Tab. 5: Robustness tests: weak instruments tests

	Target	Path	Term premium
Baseline	43.9	19.3	4.9
Critical value	23.1	27.3	31.2
FFR instead of 3M rate	48.7	25.4	4.9
Critical value	23.1	26.9	31.2
Main variables with $P = 4$ lags	47.3	20.0	4.9
Critical value	23.1	27.1	31.2
Additional control variables with $P = 2$ lags	42.5	19.6	4.9
Critical value	23.1	26.9	31.0
Only use term structure data	43.9	19.3	4.8
Critical value	23.1	27.3	31.2
Include speeches as events	33.1	18.0	4.3
Critical value	23.1	27.8	31.2
Exclude contaminating events	68.8	11.7	4.1
Critical value	23.1	25.7	27.4
Tuesday and Wednesday as control	50.6	19.6	6.5
Critical value	23.1	28.2	32.5
Exclude unscheduled meetings	1.2	0.8	0.7
Critical value	23.1	25.6	31.6
Effective lower bound period	0.0	0.4	0.4
Critical value	23.1	25.3	30.8
Swanson (2021) estimation sample	43.5	16.3	8.4
Critical value	23.1	27.5	31.4

Notes: Tests for weak instruments allowing for heteroscedasticity and autocorrelation with multiple endogenous regressors (Lewis and Mertens, 2022). The endogenous regressors are the 3M rate, 2Y rate, and 10Y - 2Y spread, respectively. The tests in the first, second, and third columns are based on one, two, and three instruments. We impose the recursive zero restrictions used to disentangle the multiple dimensions of monetary policy shocks with heteroscedasticity. For each specification, the first row shows the test statistic, and the second row shows the critical value. The significance level is set to 5% and the tolerance level to 10%.

We then examine the robustness of the results when excluding various periods from the policy event and control days. We remove contaminating events because it is debatable that other

[38] This is in line with Boehm and Kroner (2024), who find that non-interest rate financial market variables respond strongly to monetary policy announcements.

[39] See Appendix D for more details. As Swanson and Jayawickrema (2023) and Bauer and Swanson (2022) suggest, speeches by central bank representatives may also affect financial market variables.

[40] The high-frequency surprises by Swanson (2021) only comprise FOMC announcements but no speeches. We did not compare our results with the speeches by Swanson and Jayawickrema (2023) because their dataset was not publicly available at the time of writing.

shocks are randomly distributed over control and policy event days. We exclude speeches identified with the topic model, important data releases, minutes released by the FOMC, policy announcements by the Bank of England and the European Central Bank, and various crises.[41] While the correlation for the target shock increases, it remains unchanged for the path shock and falls for the term premium shock. Then, we only include Tuesdays and Wednesdays as control days because most FOMC announcements (78%) occurred on one of these weekdays. The correlations remain virtually unchanged. Finally, we estimate the shocks on the same sample as Swanson (2021) and find no relevant difference.

There are more important differences for two specific robustness tests. If we exclude unscheduled meetings, the test statistics for the weak instruments tests fall below the critical values. The reason is that we predominantly remove events before 1994 when FOMC meetings were not publicly announced (see, e.g., Swanson and Jayawickrema, 2023). This is in line with the observation that the importance of target shocks declined over time. We also find weak instruments when estimating the model only at the ELB.[42] The correlation of the target shock with the high-frequency surprises falls, suggesting that most movements during this period were driven by other events or noise. By contrast, the correlation with the path and term premium shocks remains higher, suggesting that monetary policy primarily operated through these dimensions at the ELB. However, given the weak instruments tests, we must take these results with a grain of salt.

Then, we perform a placebo test by randomly selecting event periods (the same number as in the baseline, conditional on not being an actual FOMC announcement). Figure 11 in Appendix E shows the impulse responses (blue solid lines) jointly with the baseline responses (red dotted lines). The identifying assumptions constrain the 3M rate, 2Y rate, and 10Y-2Y term spread to increase by 25 bp. Otherwise, the responses are largely statistically insignificant.

One advantage of our procedure is to use the extracted shocks to estimate the causal effects on low-frequency macroeconomic variables (see, e.g., Stock and Watson, 2018).[43] We perform weak instruments tests and estimate impulse responses in a monthly IV-LP model, including the 3M, 2Y interest rates, and the 10Y-2Y term spread, as well as industrial production, consumer prices, and the nominal effective exchange rate. The shocks are aggregated to

[41] Specifically, we exclude 9/11, the global financial crisis, the outbreak of the Gulf War, New Year's Day (because of financial market volatility), and the Covid-19 crisis.

[42] We assume that interest rates were at the effective lower bound between 16 Dec 2008 and 16 Dec 2015 and between 16 Mar 2020 and 17 Mar 2022. Allowing the column of Ψ associated with the short-term interest rate to change at the ELB would be an interesting extension. We leave this as an avenue for future research.

[43] Figure 9 in Appendix E shows that the shocks extracted with the Kalman filter are highly correlated with the high-frequency surprises by Swanson (2021) at monthly frequency as well.

monthly frequency as the sum of non-missing values. We set the monthly shock series to zero for months without an event. We associate the three shocks with the same variables as in the daily model and impose the recursive zero restrictions. The model is estimated in first (log-)differences, including 12 lags of all dependent variables.[44]

Table 9 in Appendix E shows the weak instruments tests. As in the daily model, the instruments for the path and term premium shocks do not pass the weak instruments tests for both identification schemes. The results are more favorable for the heteroscedasticity-based target shock. Overall, regarding the weak instruments tests, the heteroscedasticity-based approach performs slightly better than the high-frequency surprises.

Figure 12 in Appendix E shows the cumulative monthly impulse responses. As only the target shock passes the weak instruments tests, we should interpret the other responses carefully. But we report them for the sake of completeness. Focussing on the heteroscedasticity-based identification scheme (blue solid lines), industrial production temporarily declines after a target tightening of 25 basis points. Meanwhile, the CPI shows a delayed response, falling by about 0.5 percent after 10 months. Interestingly, the exchange rate falls by about 3 percent, but only after about three years.[45] Turning to the path shock, we also find a decline in industrial production, the CPI, and an exchange rate appreciation. The decline of industrial production is more pronounced and more persistent. For the term premium shock, we do not find statistically significant responses. The responses to the target and path shocks are qualitatively similar to those based on high-frequency surprises. The results differ more strongly for the term premium (LSAP) shock. However, as seen before, this dimension has not passed the weak instruments tests.

Another advantage is that we can apply the approach to countries and periods with missing high-frequency data. Therefore, we extend the sample period to 1986. This is the longest period for which we were able to obtain all financial market data to extract the shocks. As events, we constructed a more experimental data set starting in 1982 with 48 scheduled FOMC meetings and 19 discount rate changes that we use as event days. Table 10 in Appendix E shows that the target shock still passes the weak instruments test. However, all test statistics are lower. Regarding the extracted shock series, Figure 13 in Appendix E shows that the correlations with

[44] The additional data sources are discussed in Appendix C. The financial market data are aggregated to monthly frequency using the last available daily observation of the month. For the lag length, we follow the existing literature (Bauer and Swanson, 2022, Gertler and Karadi, 2015, Ramey, 2016). We also estimated specifications, including the excess bond premium by Gilchrist and Zakrajšek (2012), without the recursive zero restrictions and varying the lag length. The results remain qualitatively similar. Most impulse responses were less precisely estimated, however.

[45] It is puzzling that we do not find a delayed response in the daily data, while we find a stronger delay in the monthly data. Examining the cause of this difference may be an interesting avenue for future research.

the baseline series are very high. Therefore, the approach allows us to extend the shock series to earlier sample periods.[46]

4 Concluding remarks

In this paper, we propose to combine a heteroscedasticity-based identification scheme with recursive zero restrictions along the term structure of interest rates to identify multiple orthogonal monetary policy shocks. We then show how to estimate daily dynamic causal effects by modifying the IV approach by Rigobon and Sack (2004). In addition, we show how to predict the unobserved monetary policy shocks using the Kalman filter.

Our shocks are substantially correlated with the multi-dimensional high-frequency surprises by Swanson (2021), even though we use a different identification scheme. Both approaches show similar effects on the exchange rate and other variables. This suggests that our two-step approach is a valid alternative for countries and periods where high-frequency data is missing or the exact intraday timing of monetary policy announcements is unknown.

A monetary policy tightening, independent of the specific dimension, generally appreciates the US dollar exchange rate. However, there is a difference between daily and monthly models. While there is no delayed overshooting puzzle in the daily data, which may be due to estimation uncertainty, the delayed response appears in monthly local projection estimates. In addition, the monthly response is more delayed for the target than for the path shock. This contrasts with monthly SVAR models that often find no delayed overshooting on modern data samples (see, e.g., Kim et al., 2017).

Our approach can be extended in various ways. First, we detected weak instruments problems for path and term premium shocks. This may be alleviated by removing more background noise. Therefore, using our approach with intraday (e.g. close – open) rather than daily data would be interesting. In addition, it would be interesting to allow the impact matrix to change over time, for example, imposing that the short-term interest rate does not respond to a target shock at the ELB or that the term premium shock is present only after 2008.[47] Finally, it would be interesting to apply the approach to longer sample periods, where monetary policy events are known, but high-frequency data and the exact intraday timing of the announcements are unknown.

[46] It would be an interesting avenue for future research to assemble more financial market data and events to estimate monetary policy shocks further back in time.

[47] For example, Cloyne et al. (2023) explore time-varying differences in the fiscal multiplier in a local projections framework.

References

Altavilla, C., Brugnolini, L., Gürkaynak, R. S., Motto, R., and Ragusa, G. (2019). Measuring euro area monetary policy. *Journal of Monetary Economics*, 108:162–179, DOI: 10.1016/j.jmoneco.2019.08.016.

Bacchetta, P. and van Wincoop, E. (2021). Puzzling exchange rate dynamics and delayed portfolio adjustment. *Journal of International Economics*, 131:103460, DOI: 10.1016/j.jinteco.2021.103460.

Bauer, M. D. and Swanson, E. T. (2022). A reassessment of monetary policy surprises and high-frequency identification. In Eichenbaum, M., Hurst, E., and Ramey, V. A., editors, *NBER Macroeconomics Annual 2022, volume 37*. University of Chicago Press.

Bauer, M. D. and Swanson, E. T. (2023). An alternative explanation for the "Fed information effect". *American Economic Review*, 113(3):664–700, DOI: 10.1257/aer.20201220.

Benigno, G. (2004). Real exchange rate persistence and monetary policy rules. *Journal of Monetary Economics*, 51(3):473–502, DOI: 10.1016/j.jmoneco.2003.06.001.

Bjørnland, H. C. (2009). Monetary policy and exchange rate overshooting: Dornbusch was right after all. *Journal of International Economics*, 79(1):64–77, DOI: 10.1016/j.jinteco.2009.06.003.

Blei, D. M. and Lafferty, J. D. (2007). A correlated topic model of Science. *The Annals of Applied Statistics*, 1(1), ISSN: 1932-6157, DOI: 10.1214/07-AOAS114.

Boehm, C. and Kroner, T. N. (2024). Monetary policy without moving interest rates: The Fed non-yield shock. NBER Working Papers 32636, National Bureau of Economic Research.

Braun, R., Miranda-Agrippino, S., and Saha, T. (2024). Measuring monetary policy in the UK: The UK monetary policy event-study database. *Journal of Monetary Economics*, page in press, DOI: 10.1016/j.jmoneco.2024.103645.

Brennan, C. M., Jacobson, M. M., Matthes, C., and Walker, T. B. (2024). Monetary policy shocks: Data or methods? Finance and Economics Discussion Series 2024-001, Federal Reserve Board, DOI: 10.17016/FEDS.2024.011.

Bu, C., Rogers, J., and Wu, W. (2021). A unified measure of Fed monetary policy shocks. *Journal of Monetary Economics*, 118:331–349, DOI: 10.1016/j.jmoneco.2020.11.002.

Canetg, F. and Kaufmann, D. (2022). Overnight rate and signalling effects of central bank bills. *European Economic Review*, 143:104060, DOI: 10.1016/j.euroecorev.2022.104060.

Cloyne, J., Jordà, O., and Taylor, A. M. (2023). State-dependent local projections: Understanding impulse response heterogeneity. Working Paper 30971, National Bureau of Economic Research, DOI: 10.3386/w30971.

Dornbusch, R. (1976). Expectations and exchange rate dynamics. *Journal of Political Economy*, 84(6):1161–1176, DOI: 10.1086/260506.

Eichenbaum, M. and Evans, C. L. (1995). Some empirical evidence on the effects of shocks to monetary policy on exchange rates. *The Quarterly Journal of Economics*, 110(4):975–1009, DOI: 10.2307/2946646.

Fama, E. F. and MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, 81(3):607–636, DOI: 10.1086/260061.

Faust, J., Rogers, J., Swanson, E., and Wright, J. (2003). Identifying the effects of monetary policy shocks on exchange rates using high frequency data. *Journal of the European Economic Association*, 1(5):1031–1057, DOI: 10.1162/154247603770383389.

Georgiadis, G. and Jarociński, M. (2023). Global spillovers from multi-dimensional US monetary policy. Working Paper Series 2881, European Central Bank.

Gertler, M. and Karadi, P. (2015). Monetary policy surprises, credit costs, and economic activity. *American Economic Journal: Macroeconomics*, 7(1):44–76, DOI: 10.1257/mac.20130329.

Gilchrist, S. and Zakrajšek, E. (2012). Credit Spreads and Business Cycle Fluctuations. *American Economic Review*, 102(4):1692–1720, ISSN: 0002-8282, DOI: 10.1257/aer.102.4.1692, Retrieved from https://pubs.aeaweb.org/doi/10.1257/aer.102.4.1692.

Gürkaynak, R. S., Kara, A. H., Kısacıkoğlu, B., and Lee, S. S. (2021). Monetary policy surprises and exchange rate behavior. *Journal of International Economics*, 130:103443, DOI: 10.1016/j.jinteco.2021.103443.

Gürkaynak, R. S., Sack, B., and Swanson, E. (2005). Do actions speak louder than words? The response of asset prices to monetary policy actions and statements. *International Journal of Central Banking*, 1(1).

Gürkaynak, R. S., Sack, B., and Wright, J. H. (2007). The U.S. Treasury yield curve: 1961 to the present. *Journal of Monetary Economics*, 54(8):2291–2304, DOI: 10.1016/j.jmoneco.2007.06.029.

Hamilton, J. D. (1994). *Time Series Analysis*. Princeton University Press.

Jarociński, M. (2024). Estimating the Fed's unconventional policy shocks. *Journal of Monetary Economics*, 144:103548, DOI: 10.1016/j.jmoneco.2024.01.001.

Jordà, O. (2005). Estimation and inference of impulse responses by local projections. *American Economic Review*, 95(1):161–182, DOI: 10.1257/0002828053828518.

Kearns, J. and Manners, P. (2006). The impact of monetary policy on the exchange rate: A study using intraday data. *International Journal of Central Banking*, 2(4).

Kim, S.-H., Moon, S., and Velasco, C. (2017). Delayed overshooting: Is it an '80s puzzle? *Journal of Political Economy*, 125(5):1570–1598, DOI: 10.1086/693372.

Kuttner, K. N. (2001). Monetary policy surprises and interest rates: Evidence from the Fed funds futures market. *Journal of Monetary Economics*, 47(3):523–544, DOI: 10.1016/S0304-3932(01)00055-1.

Lewis, D. J. (2019). Announcement-specific decompositions of unconventional monetary policy shocks and their macroeconomic effects. Staff Reports 891, Federal Reserve Bank of New York.

Lewis, D. J. (2022). Robust inference in models identified via heteroskedasticity. *The Review of Economics and Statistics*, 104(3):510–524, DOI: 10.1162/rest_a_00963.

Lewis, D. J. and Mertens, K. (2022). A robust test for weak instruments with multiple endogenous regressors. Staff Reports 1020, Federal Reserve Bank of New York.

Miranda-Agrippino, S. and Ricco, G. (2021). The transmission of monetary policy shocks. *American Economic Journal: Macroeconomics*, 13(3):74–107, DOI: 10.1257/mac.20180124.

Montiel Olea, J. L., Plagborg-Møller, M., Qian, E., and Wolf, C. K. (2024). Double robustness of local projections and some unpleasant varithmetic. Working Paper 32495, National Bureau of Economic Research, DOI: 10.3386/w32495.

Müller, G. J., Wolf, M., and Hettig, T. (2023). Delayed overshooting: The case for information rigidities. *mimeo*, Retrieved from mwolfunisg.github.io/website/assets/files/DelayedOvershooting_2023Feb.pdf.

Nakamura, E. and Steinsson, J. (2018). High-frequency identification of monetary non-neutrality: The information effect. *The Quarterly Journal of Economics*, 133(3):1283–1330, DOI: 10.1093/qje/qjy004.

Newey, W. and West, K. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3):703–708.

Ramey, V. A. (2016). Macroeconomic shocks and their propagation. Working Paper 21978, National Bureau of Economic Research, DOI: 10.3386/w21978.

Ricco, G., Degasperi, R., and Hong, S. (2020). The global transmission of U.S. monetary policy. CEPR Discussion Papers 14533, C.E.P.R.

Rigobon, R. (2003). Identification through heteroskedasticity. *The Review of Economics and Statistics*, 85(4):777–792.

Rigobon, R. and Sack, B. (2004). The impact of monetary policy on asset prices. *Journal of Monetary Economics*, 51(8):1553–1575, DOI: 10.1016/j.jmoneco.2004.02.004.

Roberts, M. E., Stewart, B. M., and Airoldi, E. M. (2016). A model of text for experimentation in the social sciences. *Journal of the American Statistical Association*, 111(515):988–1003, DOI: 10.1080/01621459.2016.1141684.

Roberts, M. E., Stewart, B. M., and Tingley, D. (2019). stm: An R package for structural topic models. *Journal of Statistical Software*, 91(2), DOI: 10.18637/jss.v091.i02.

Rogoff, K. (2002). Dornbusch's overshooting model after twenty-five years. IMF Working Papers 2002/039, International Monetary Fund.

Schlaak, T., Rieth, M., and Podstawski, M. (2023). Monetary policy, external instruments, and heteroskedasticity. *Quantitative Economics*, 14(1):161–200, DOI: 10.3982/QE1511.

Schmitt-Grohé, S. and Uribe, M. (2022). The effects of permanent monetary shocks on exchange rates and uncovered interest rate differentials. *Journal of International Economics*, 135:103560, DOI: 10.1016/j.jinteco.2021.103560.

Scholl, A. and Uhlig, H. (2008). New evidence on the puzzles: Results from agnostic identification on monetary policy and exchange rates. *Journal of International Economics*, 76(1):1–13, DOI: 10.1016/j.jinteco.2008.02.005.

Stock, J. H. and Watson, M. W. (2018). Identification and estimation of dynamic causal effects in macroeconomics using external instruments. *The Economic Journal*, 128(610):917–948, DOI: 10.1111/ecoj.12593.

Swanson, E. T. (2021). Measuring the effects of Federal Reserve forward guidance and asset purchases on financial markets. *Journal of Monetary Economics*, 118:32–53, DOI: 10.1016/j.jmoneco.2020.09.003.

Swanson, E. T. (2023a). The importance of Fed Chair speeches as a monetary policy tool. *AEA Papers and Proceedings*, 113:394–400, DOI: 10.1257/pandp.20231073.

Swanson, E. T. (2023b). The macroeconomic effects of the Federal Reserve's conventional and unconventional monetary policies. Working Paper 31603, National Bureau of Economic Research, DOI: 10.3386/w31603.

Swanson, E. T. and Jayawickrema, V. (2023). Speeches by the Fed Chair are more important than FOMC announcements: An improved high-frequency measure of U.S. monetary policy shocks. *mimeo*.

Wright, J. H. (2012). What does monetary policy do to long-term interest rates at the zero lower bound? *The Economic Journal*, 122(564):F447–F466.

Zhu, L. (2023). Let the market speak: Using interest rates to identify the Fed information effect. Working paper, SSRN, DOI: 10.2139/ssrn.4035869.

Appendix

A Extension daily model

The model can be extended to include control variables, such as lags of the dependent variables, in the information set. Let the model comprise M control variables and a vector of constant terms (α):

$$\begin{aligned} y_t = &\ \alpha + \sum_{l=1}^{L} \Phi_l x_{t-l} + \Psi \varepsilon_t + \Gamma v_t \quad \text{for } t \in P \\ y_t = &\ \alpha + \sum_{l=1}^{L} \Phi_l x_{t-l} + \Gamma v_t \quad \text{for } t \in C \end{aligned} \tag{3}$$

where Φ_l are $N \times M$ matrices of coefficients for every lag $l = 1, \ldots, L$. Then, we simply add constant terms and x_{t-l}, for $l = 1, \ldots, L$ as additional regressors in the IV estimation. The construction of the instruments remains unchanged.

Following Jordà (2005), we can also estimate dynamic effects by iterating the dependent variable forward:

$$\begin{aligned} y_{t+h} = &\ \alpha_h + \sum_{l=1}^{L} \Phi_l^{(h)} x_{t-l} + \sum_{n=0}^{h} \Psi^{(h-n)} \varepsilon_{t+n} + \Gamma^{(h-n)} v_{t+n} \quad \text{for } t \in P \\ y_{t+h} = &\ \alpha_h + \sum_{l=1}^{L} \Phi_l^{(h)} x_{t-l} + \sum_{n=0}^{h} \Gamma^{(h-n)} v_{t+n} \quad \text{for } t \in C \end{aligned} \tag{4}$$

where $\Psi^{(h)}$ and $\Gamma^{(h)}$ are the impulse response functions after h periods and $\Phi_l^{(h)}$ are coefficients on the control variables, which differ for every horizon h.[48]

The error term in the IV estimation includes future monetary policy and other shocks. The exclusion restriction is still valid as the instruments are only affected by current shocks, not by future shocks.[49] We can use the same first stages and then replace the dependent variables in the second stage with y_{t+h} to estimate the impulse response after h periods. For the same reason, we can estimate cumulative responses by using $\sum_{n=0}^{h} y_{t+n}$ as the dependent variable with the same instruments.

[48]For example, if the data generating process is a VAR(1) we have that $y_{t+h} = \alpha_h + \Phi^h y_{t-1} + \sum_{n=0}^{h} \Phi^{h-n} \Psi \varepsilon_{t+n} + \Phi^{h-n} \Gamma v_{t+n}$, with $\Psi^{(h)} \equiv \Phi^h \Psi$ and $\Gamma^{(h)} \equiv \Phi^h \Gamma$.

[49]However, the error term is autocorrelated, such that it is important to use a HAC-consistent variance estimator (see Newey and West, 1987).

B Simulation study

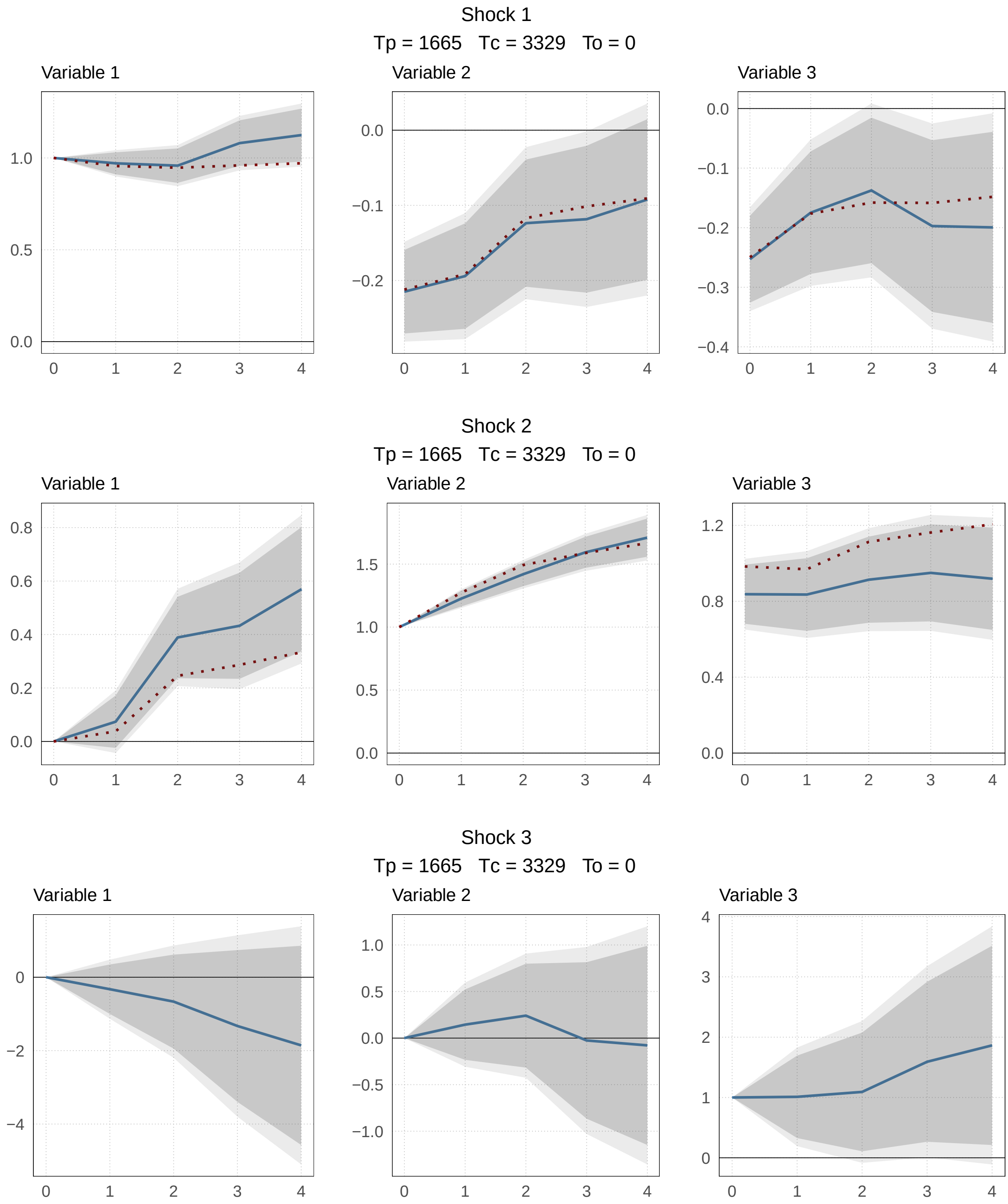
A simulation study illustrates that we can estimate the impulse responses of a multi-dimensional monetary policy shock with IV. We simulate 5'000 observations for $N = 3$ variables using a VAR with $P = 2$ lags. $R = 3$ i.i.d. structural shocks occur on all periods and $E = 2$ i.i.d. structural shocks occur every third period. Both impact matrices are lower triangular. The specific values of the impact matrices and VAR coefficients are drawn randomly, subject to the constraint that the VAR is stationary.

Figure 5 shows that we can estimate the impulse responses of the three variables to the two shocks (first two rows). The red dotted line shows the actual impulse response. The impact effect is accurately estimated, while there are somewhat larger deviations and wider confidence intervals at longer horizons. We also estimate the impulse response to a non-existing third dimension (third row). The confidence intervals are wide and mostly include zero. For the third variable, the normalization leads to a significant response in the first period. This suggests that if we estimate the dynamic causal effect to a non-existing dimension, we expect significant responses of the variables on which we impose the normalization, which corresponds to the assumption in heteroscedasticity-based identification schemes that the shock increases the variance of the variable on event days. Meanwhile, the other variables are not significantly affected.

Can we detect more formally whether we estimate the dynamic causal effects to a non-existing dimension? Wright (2012) proposes a bootstrap-based test on the difference of the variance-covariance matrix to test whether there is a single monetary policy dimension at the effective lower bound. Alternatively, note that the instrument we construct for the non-existing third dimension should be weak, as shocks 1 and 2 already explain the variation of variable 3. Table 6 indeed shows that the test for weak instruments in the presence of multiple endogenous regressors suggested by Lewis and Mertens (2022) does not reject the null hypothesis of weak instruments for the third equation, where we have three endogenous regressors and three heteroscedasticity-based instruments. In contrast, we reject the null for equations 1 and 2.

We also conducted a simulation study to test our strategy to extract the shocks using the Kalman filter. The data have been simulated with the same structure as above, except that we simulate separate data sets for $N = 3, 10, 20$ to test whether increasing the number of variables increases the prediction's accuracy. In this exercise, we set $R = N$. Figure 6 shows the first 100 observations and predictions of the structural shocks, respectively, where every

Fig. 5: Impulse responses for simulated data $E = 2$



Notes: Cumulative impulse responses two-dimensional shock occurring every third period estimated using IV-LP. The red dotted lines give the true impulse responses. 90% and 95% confidence intervals are based on HAC-robust standard errors. T_p, T_c, T_o denote the number of policy event days, control days, and other days, respectively.

Tab. 6: Tests for weak instruments for simulated data $E = 2$

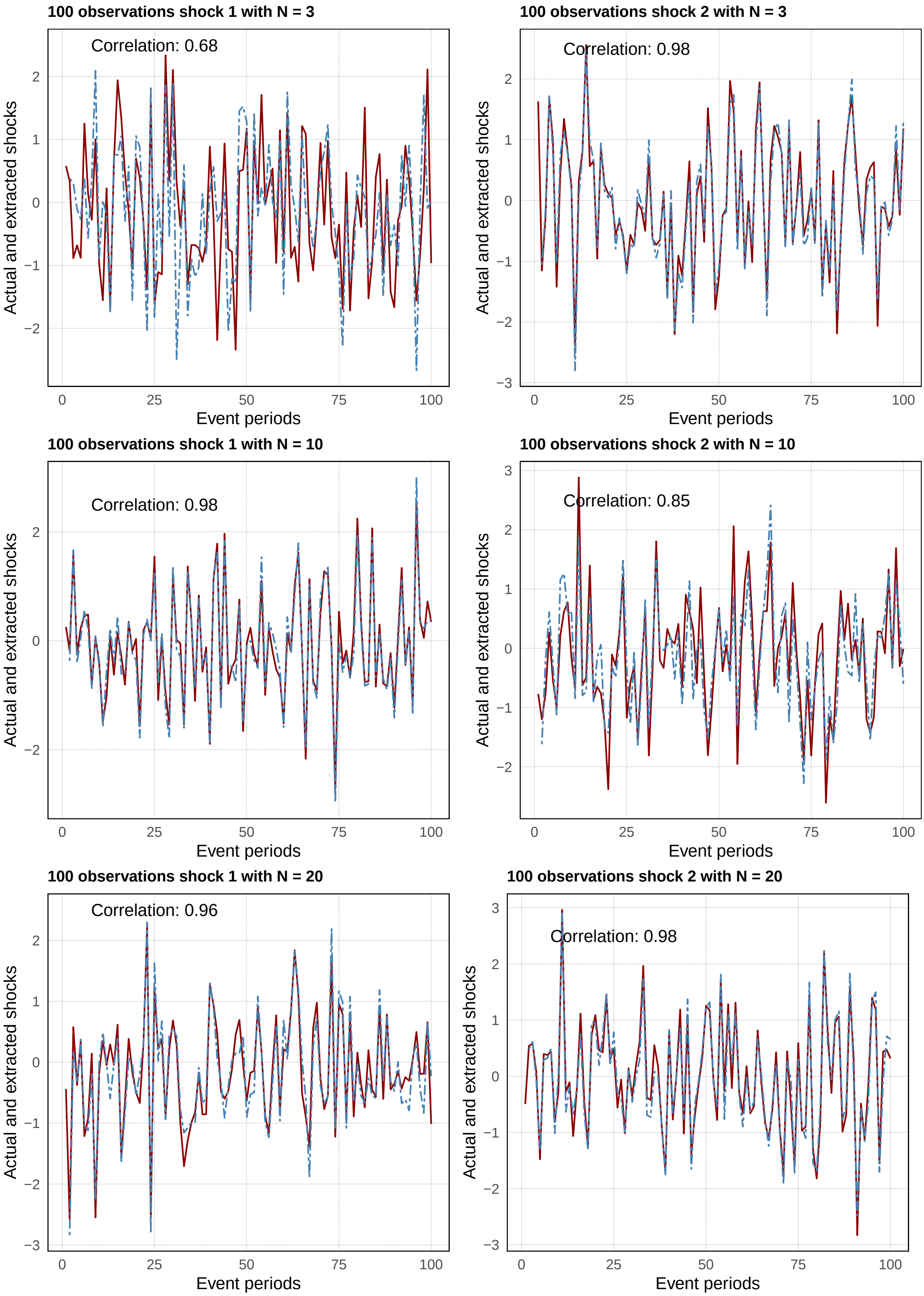
	Equation 1	Equation 2	Equation 3
Test statistic	425.1	261.6	0.3
Critical value	23.1	23.4	23.3

Notes: Tests for weak instruments allowing for heteroscedasticity and autocorrelation with multiple endogenous regressors (Lewis and Mertens, 2022). The endogenous regressors are the first, second, and third variables in the system. The tests shown in the first, second, and third columns are based on one, two, and three instruments. We impose the recursive zero restrictions used to disentangle the multiple dimensions of monetary policy shocks with heteroscedasticity. The significance level is set to 5% and the tolerance level to 10%.

column stands for a different underlying structural shock, and every row varies the number of variables (N). We can relatively accurately extract the underlying shocks with three variables. However, increasing the number of variables increases the accuracy of the prediction. This may be related to the fact that the response to the structural shocks varies across variables in the simulation, providing independent information on the underlying shocks. Of course, if every variable responds identically to the structural shocks, adding a variable will not improve the prediction.

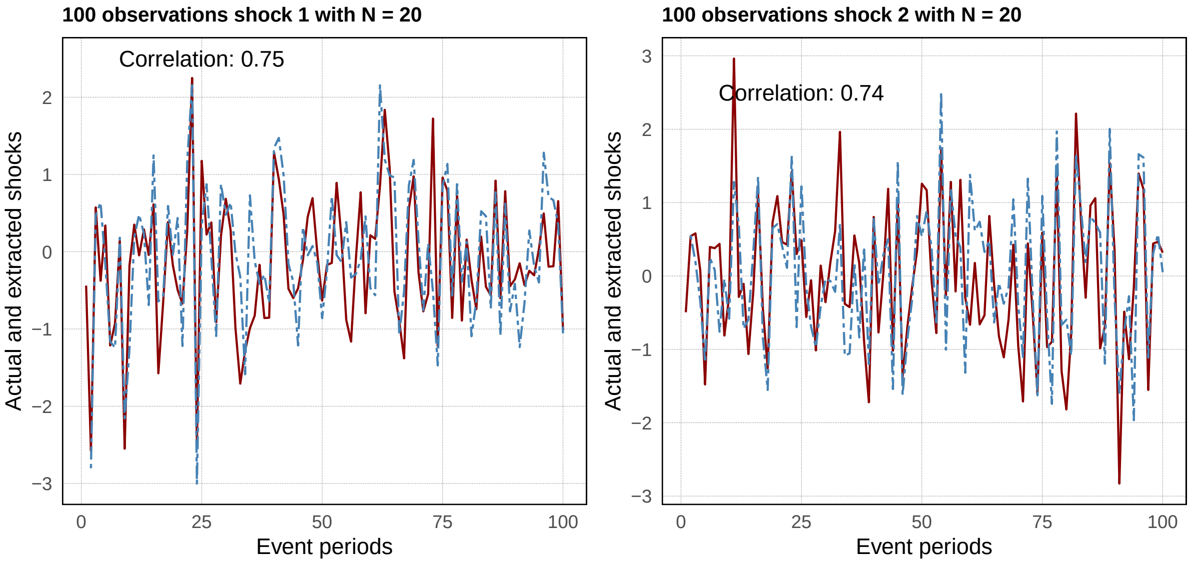
Finally, we follow Bu et al. (2021) and use Fama and MacBeth (1973) regressions to extract the shocks with $N = 20$. Note that in these simulations, the orthogonality condition is satisfied because we randomly draw the coefficients of the impact matrices. However, the approach does not deliver the linear minimum MSE prediction of the shocks. Figure 7 shows that the correlation with the actual shocks is lower than the Kalman-filter predictions.

Fig. 6: Simulated and extracted shocks for simulated data $E = 2$ using the Kalman filter



Notes: Actual (red solid line) and extracted (blue dashed line) structural shocks using the Kalman filter.

Fig. 7: Simulated and extracted shocks $E = 2$ using Fama-MacBeth regressions



Notes: Actual (red solid line) and extracted (blue dashed line) structural shocks using Fama-MacBeth regressions.

C Data

Tables 7 and 8 summarize the sources for the time series, policy events, and control days. The baseline data set covers the period 1988 – 2022. We construct a more experimental data set for robustness tests over a longer sample period.

Tab. 7: Time series

Category	**Source**	**Variants**	**Time stamp**	**Comments**
Treasury bill yields	Board of Governors	3M, 6M, 1Y, 2Y, 3Y, 5Y, 7Y, 30Y	4pm EST	www.federalreserve.gov/releases/h15/. We do not use the 20Y yield as it comprises missing data
Treasury bill yields	Gürkaynak et al. (2007)	1Y to 30Y	End of day	www.federalreserve.gov/data/nominal-yield-curve.htm
Federal Funds Rate	Board of Governors		4pm EST	www.federalreserve.gov/releases/h15/
Exchange rates	Board of Governors	NEER, USD/CHF, USD/JPY, USD/GBP, USD/EUR	Noon EST	www.federalreserve.gov/releases/h10/. For the nominal effective exchange rate, we linked the discontinued series with FRED identifier DTWEXM with DTWEXAFEGS. For the USD/EUR exchange rate, we linked the USD/DEM with the USD/EUR exchange rate using the official changeover exchange rate from www.eu-info.de/euro-waehrungsunion/5007/5222/5170/.
Stock prices	S&P Dow Jones, NASDAQ OMX	S&P 500, NASDAQ	4pm EST	de.tradingview.com/symbols/SPX/. FRED variable keys: SP500, NASDAQCOM
Corporate bond spreads	Moody's	AAA, BAA	Unclear	FRED variable keys: DAAA, DBAAA. We computed the spreads as the difference to the 10Y government bond yield
Commodity price index	Dow Jones, Bloomberg	Eikon Datastream	Close	We link the Dow Jones spot commodity price index with the Bloomberg spot commodity price index (BCOM)
Oil price index	US Energy Information Administration	Crude oil WTI	Close	FRED variable key: DCOILWTICO
Industrial production	Board of Governors			FRED variable key: INDPRO
CPI	BLS			FRED variable key: CPIAUCSL

A few comments are in order. Besides a nominal effective exchange rate, we also examine the USD exchange rate vis-à-vis the CHF, GBP, JPY, CAD, and EUR. Before the euro-changeover, we used the USD/DEM exchange rate. The USD/DEM is transformed to a hypothetical USD/EUR using the official euro-changeover exchange rate. All exchange rates are defined as one USD in terms of foreign currency. A decrease in the exchange rate is an appreciation of the USD.

Tab. 8: Events

Category	Source	Start	Comments
FOMC announcements	1982 to 1987: www.federalreserve.gov/monetarypolicy/fomc_historical_year.htm, 1988 to 2019: Bauer and Swanson (2022), from 2020: www.federalreserve.gov/monetarypolicy/fomccalendars.htm	1982	Longer samples possible
Discount rate changes	Monetary Policy and Open Market Operations 1982 - 1989, FRBNY Quarterly Review, 1983 - 1990, www.newyorkfed.org/research/quarterly_review/75th.html	1982	Longer samples possible
Speeches and Congressional Testimony	Up to 1996: alfred.stlouisfed.org/, from 1997: www.federalreserve.gov/newsevents/speeches.htm	1982	Longer samples possible
FOMC minutes	www.federalreserve.gov/monetarypolicy/fomccalendars.htm	1988	
ECB decisions	Altavilla et al. (2019)	1999	
BoE decisions	Braun et al. (2024)	1997	
CPI releases	www.bls.gov/bls/news-release/cpi.htm, alfred.stlouisfed.org/	1982	Longer samples possible
PPI releases	www.bls.gov/bls/news-release/ppi.htm	1994	
Employment situation releases	www.bls.gov/bls/news-release/empsit.htm	1994	
Employment cost releases	www.bls.gov/bls/news-release/eci.htm	1982	Longer samples possible
GDP releases	www.bea.gov/index.php/news/archive?field_related_product_target_id=All&created_1=All&title=gross%20domestic%20product&page=0	1996	Includes first, second and third releases
Industrial production releases	www.federalreserve.gov/releases/g17/release_dates.htm	1982	Longer samples possible

Notes: Prior to 1994, the FOMC did not explicitly announce its target for the FFR, but implemented changes in its target via open market operations. These open market operations were conducted at 11:30 am the next morning (see, e.g., Swanson and Jayawickrema, 2023). Therefore, we used the day after an FOMC meeting as the day of the event from 1982 to 1987.

In a robustness test, we also use dates of relevant speeches and testimony before Congress by the Chair and the Vice-Chair of the Federal Reserve. Recent work by Swanson (2023a) and Swanson and Jayawickrema (2023) shows that these speeches are an important source of variation in US monetary policy. In contrast to Swanson and Jayawickrema (2023) who read the newspaper the following day to judge whether the speech had implications for monetary policy, we use a more data-driven approach to identify relevant speeches. We use a topic modeling approach, identifying transcripts mainly concerning monetary policy decisions. A detailed description of how relevant articles are identified can be found in Appendix D. We aim to exclude transcripts that refer mainly to the state of the economy or regulatory changes. In total, we add 81 speeches to our event dataset.[50]

As a robustness test, we assembled data over a longer sample. Between 1982 and 1988, we identified 17 additional speeches, 48 scheduled FOMC meetings, and 19 discount rate changes that we used as event days. However, because some financial market data is only available on a shorter sample, the sample for the robustness tests starts in 1986.

For the monthly model, we aggregate the financial market variables (3M, 2Y interest rates, 10Y-2Y term spread, and nominal effective exchange rate) using the last observation of the month. Then we add two monthly macroeconomic variables (consumer price index and industrial production). The sources are given in Table 7.

[50] Two of them take place on the same date as a FOMC announcement.

D Identification of relevant speeches

Swanson (2023a) and Swanson and Jayawickrema (2023) show that speeches and Congressional testimony (henceforward just speeches) of the Chair and the Vice-Chair of the Federal Reserve Board are an important source of variation in US monetary policy. Therefore, we augment the event dataset with relevant speeches. In contrast to Swanson and Jayawickrema (2023) who read the newspaper the next morning to judge whether the speech had implications for monetary policy, we use a more data-driven approach to identify relevant speeches.

To identify speeches of the Chair and the Vice-Chair of the Federal Reserve, we apply a Correlated Topic Model (CTM).[51] We downloaded all speeches by officials of the Federal Reserve from 1996 onwards from the website of the Federal Reserve.[52] Speeches before 1996 are collected from ALFRED. [53] We then identify and link together words that often appear together (e.g., interest rate or Federal Reserve) by calculating bigrams (contiguous sequences of two words). Finally, we apply the CTM after cleaning the corpus from stopwords, numbers, and punctuation.

The CTM was initially developed by Blei and Lafferty (2007). Here, we use the algorithm described in Roberts et al. (2016, 2019). We estimate a Structural Topic Model (STM), which reduces to a fast implementation of the CTM if estimated without covariates.[54] The CTM is a statistical model used to analyze large sets of documents. It assumes that each document in the collection comprises a mixture of different topics, and each topic is a probability distribution over the words in the vocabulary. It is superior in this context to other topic modeling approaches, such as Latent Dirichlet Allocation (LDA), because it explicitly models the correlations between the topics, which may be important for understanding the underlying structure of the data. For classifying speeches of Federal Reserve officials, there may be topics frequently discussed together (such as inflation and monetary policy) or strongly influencing each other.

We follow the notation by Roberts et al. (2016) denoting documents using the index $d \in \{1, \dots, D\}$ and words (or positions within the documents) using the index $n \in \{1, \dots, N\}$. Each word in a document, represented as $w_{d,n}$, is an instance of distinct words drawn from a

[51] We follow Swanson and Jayawickrema (2023) and focus on the most influential members of the FOMC: the Federal Reserve Board Chair and the Federal Reserve Board Vice Chair. However, to estimate the topic model, we used speeches given by Board Governors.

[52] www.federalreserve.gov/newsevents/speeches.htm

[53] alfred.stlouisfed.org

[54] See Roberts et al. (2016, 2019) for more details on the STM.

vocabulary that is indexed by $v \in \{1, \ldots, V\}$. Additionally, the model assumes the selection of a certain number of topics, K, which are indexed by $k \in \{1, \ldots, K\}$.

The CTM is a generative model assuming that each document d, given the number of topics K and observed words $w_{d,n}$, is generated in the following way:

$$\begin{aligned} \eta_d &\sim \mathcal{N}_{K-1}(\mu, \Sigma) \\ \theta_{d,k} &= \frac{\exp(\eta_{d,k})}{\sum_{i=1}^{K} \exp(\eta_{d,i})} \end{aligned}$$

where η_d is the latent topic proportion vector for document d, transformed to the simplex via a logistic function to get θ_d. $\eta_{d,K}$ is fixed to zero to identify the model. μ is the mean vector, and Σ is the covariance matrix (capturing topic correlations) of the topic proportion. Given the topic proportion vector, θ_d, for each word, indexed by n, within document d, a topic indicator is sampled from

$$z_{d,n} \sim \text{Multinomial}_K(\theta_d)$$

whose positive component indicates the topic associated with that particular position. Conditional on such a topic indicator, a word is sampled from

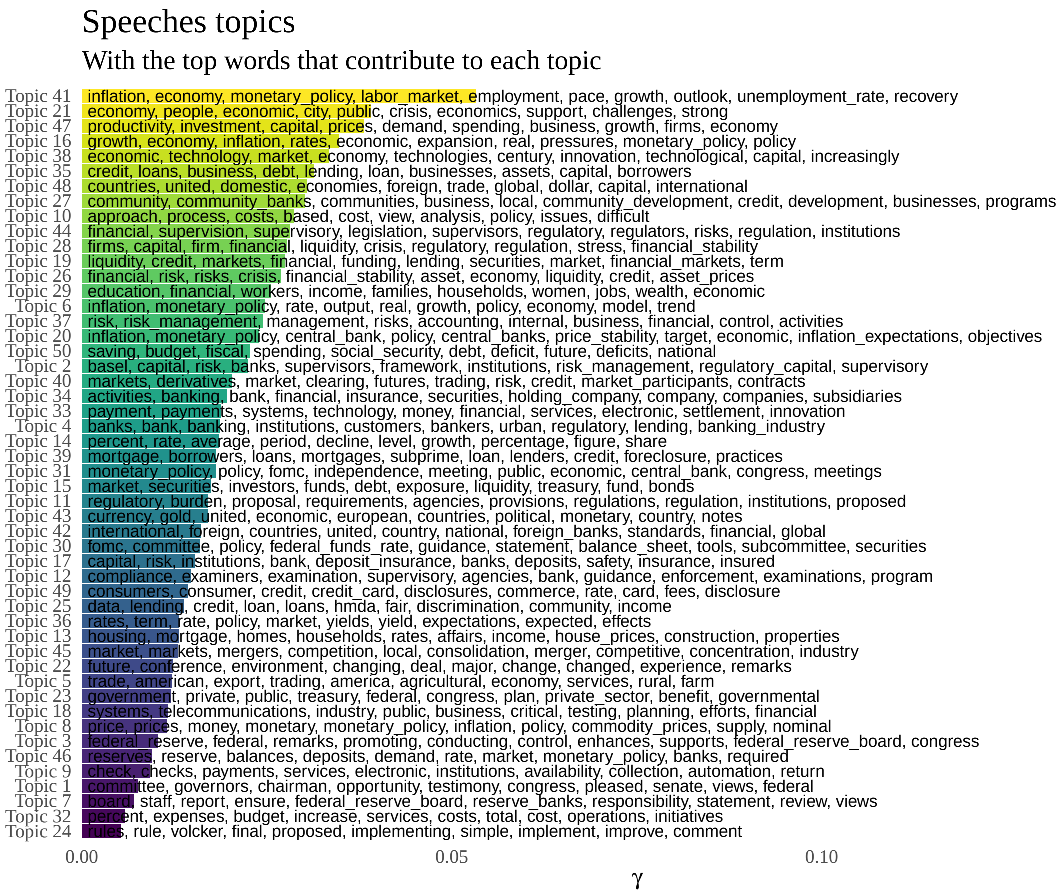
$$w_{d,n} \sim \text{Multinomial}_V(\beta_{z_{d,n}})$$

where V is the size of the vocabulary and β is the $K \times V$ matrix representing the distributions of terms in the vocabulary corresponding to the K topics.

The objects of interest in a CTM include the distributions of topics within documents (θ_d), the distributions of words across topics (β), the topic assignments for each word ($z_{d,n}$), and the parameters (μ, Σ) of the logistic normal distribution. Estimating these components allows us to understand the thematic structure in a text corpus. However, inference in a CTM is challenging due to the non-conjugate nature of the logistic normal and multinomial distributions. We use an approximate variational EM algorithm based on a Laplace approximation developed by Roberts et al. (2016).

Following Roberts et al. (2019) we set the number of topics to $K = 50$. Figure 8 illustrates the identified topics in terms of their frequency of occurrence within the text corpus and the words that most accurately describe them. From a human standpoint, the top words are perceived as coherent and meaningful, resulting in the interpretability of the topics. Therefore, we use the top words to identify the topics associated with monetary policy. We explicitly choose only those topics that are directly related to monetary policy. According to our judgment, six topics can be associated with monetary policy: 6, 20, 30, 31, 36, 46. Accordingly, we identify and include 81 speeches in the event dataset.

Fig. 8: Topic prevalence with the top words that contribute to topics



E Additional results

Tab. 9: Tests for weak instruments monthly data

	Target	Path	Term premium
Test statistic	44.5	2.4	1.4
Critical value	23.1	26.9	28.6

	Swanson FFR	Swanson path	Swanson LSAP
Test statistic	13.0	0.3	0.2
Critical value	23.1	24.1	25.7

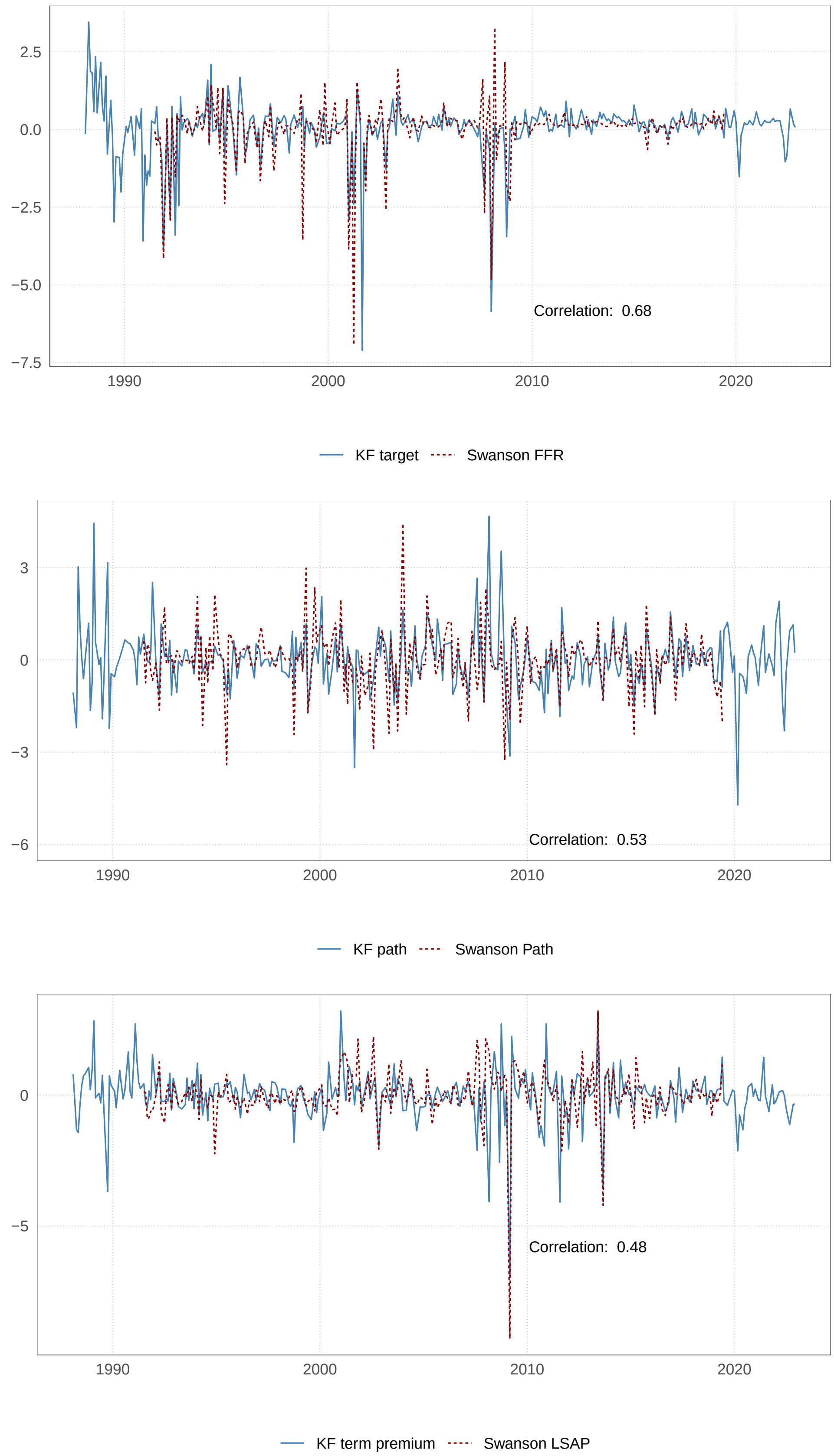
Notes: Tests for weak instruments allowing for heteroscedasticity and autocorrelation with multiple endogenous regressors (Lewis and Mertens, 2022) based on monthly data. The endogenous regressors are the 3M rate, 2Y rate, and 10Y - 2Y spread, respectively. The tests in the first, second, and third columns are based on one, two, and three instruments. We impose the recursive zero restrictions used to disentangle the multiple dimensions of monetary policy shocks with heteroscedasticity. The first panel shows the results for the heteroscedasticity-based instruments. The second panel shows the results based on the high-frequency surprises by Swanson (2021). The significance level is set to 5% and the tolerance level to 10%.

Tab. 10: Tests for weak instruments starting in 1986

	Target	Path	Term premium
Test statistic	30.8	16.6	2.7
Critical value	23.1	27.7	31.2

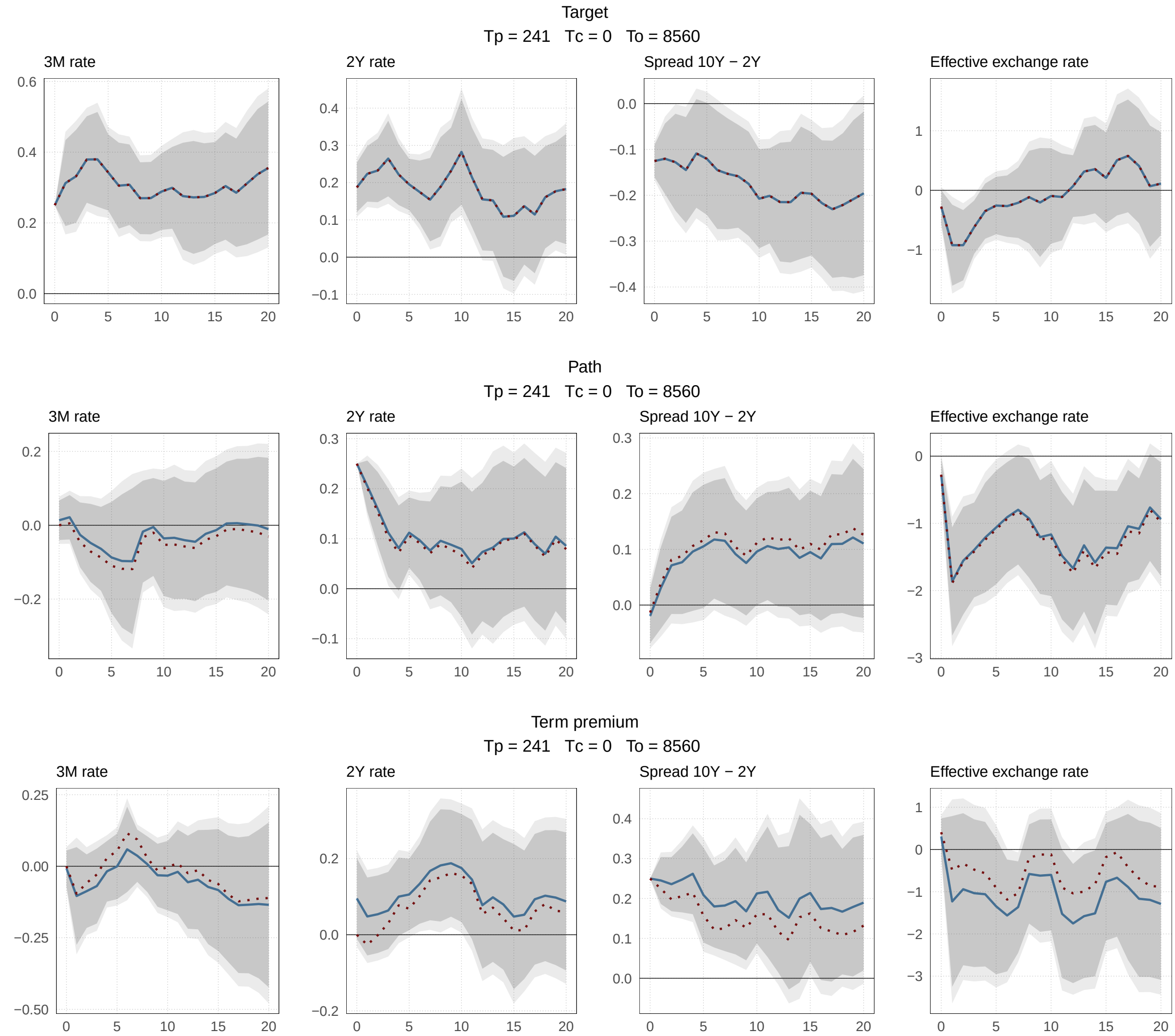
Notes: Tests for weak instruments allowing for heteroscedasticity and autocorrelation with multiple endogenous regressors (Lewis and Mertens, 2022) using data starting in 1986. The endogenous regressors are the 3M rate, 2Y rate, and 10Y - 2Y spread, respectively. The tests in the first, second, and third columns are based on one, two and three instruments. We impose the recursive zero restrictions used to disentangle the multiple dimensions of monetary policy shocks with heteroscedasticity. The significance level is set to 5% and the tolerance level to 10%.

Fig. 9: Heteroscedasticity-based and high-frequency shocks (monthly frequency)



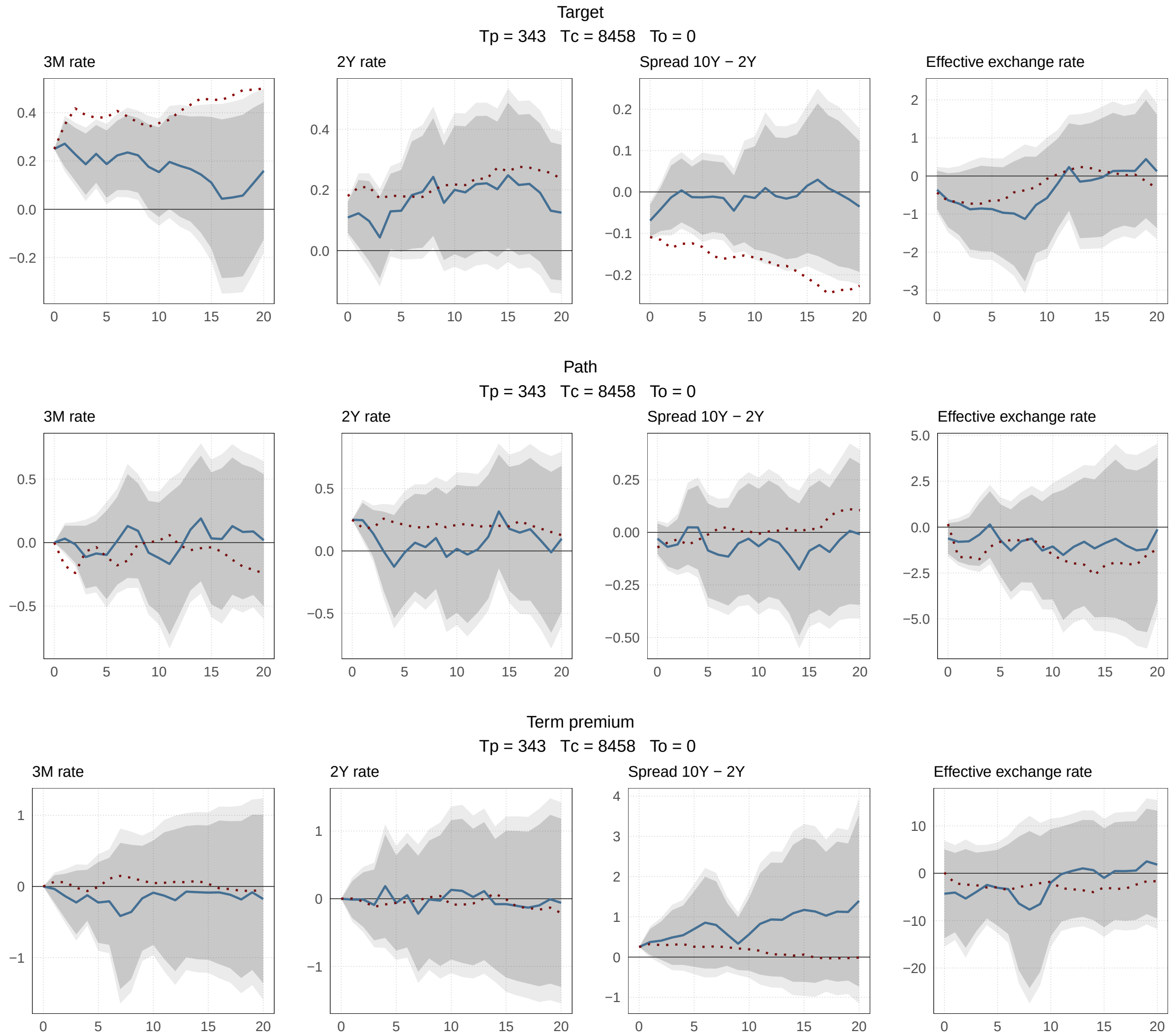
Notes: Monthly sum of the daily shocks extracted using the Kalman filter (KF) and high-frequency surprises by Swanson (2021). All series normalized to a mean of zero and a standard deviation of one.

Fig. 10: Impulse responses to high-frequency surprises with and without zero restrictions



Notes: Impulse responses to high-frequency monetary policy surprises by Swanson (2021). The responses are normalized to a 25 bp increase in the 3M rate, 2Y rate, and 10Y - 2Y spread, respectively. The blue solid line shows the unrestricted responses. The red dotted lines additionally impose the recursive zero restrictions. The horizontal axis measures working days (excluding weekends and holidays). The models are estimated in first (log-)differences, but the impulse responses are cumulated. Therefore, all interest rate responses are measured in percentage points and the exchange rate responses are measured in percent. 90% and 95% confidence intervals are based on HAC-robust standard errors. T_p, T_c, T_o denote the number of policy event days, control days, and other days, respectively.

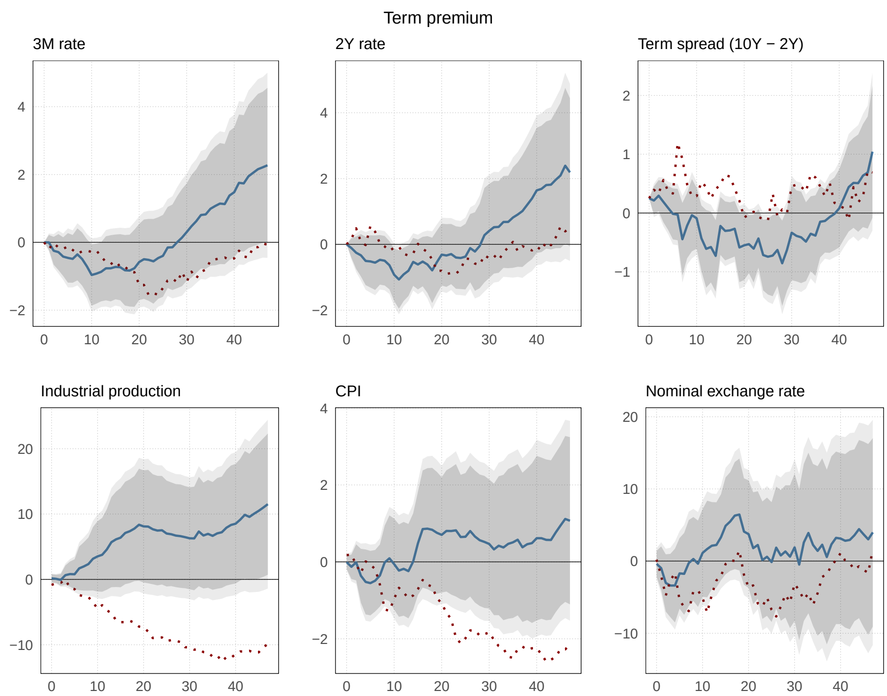
Fig. 11: Heteroscedasticity-based impulse responses using random event dates as placebo test



Notes: Impulse responses using randomly selected event dates. The responses are normalized to a 25 bp increase in the 3M rate, 2Y rate, and 10Y - 2Y spread, respectively. The blue solid lines show the responses to the placebo events. The red dotted lines show the responses from our baseline specification. The horizontal axis measures working days (excluding weekends and holidays). The models are estimated in first (log-)differences, but the impulse responses are cumulated. Therefore, all interest rate responses are measured in percentage points and the exchange rate responses are measured in percent. 90% and 95% confidence intervals are based on HAC-robust standard errors. T_p, T_c, T_o denote the number of policy event days, control days, and other days, respectively.

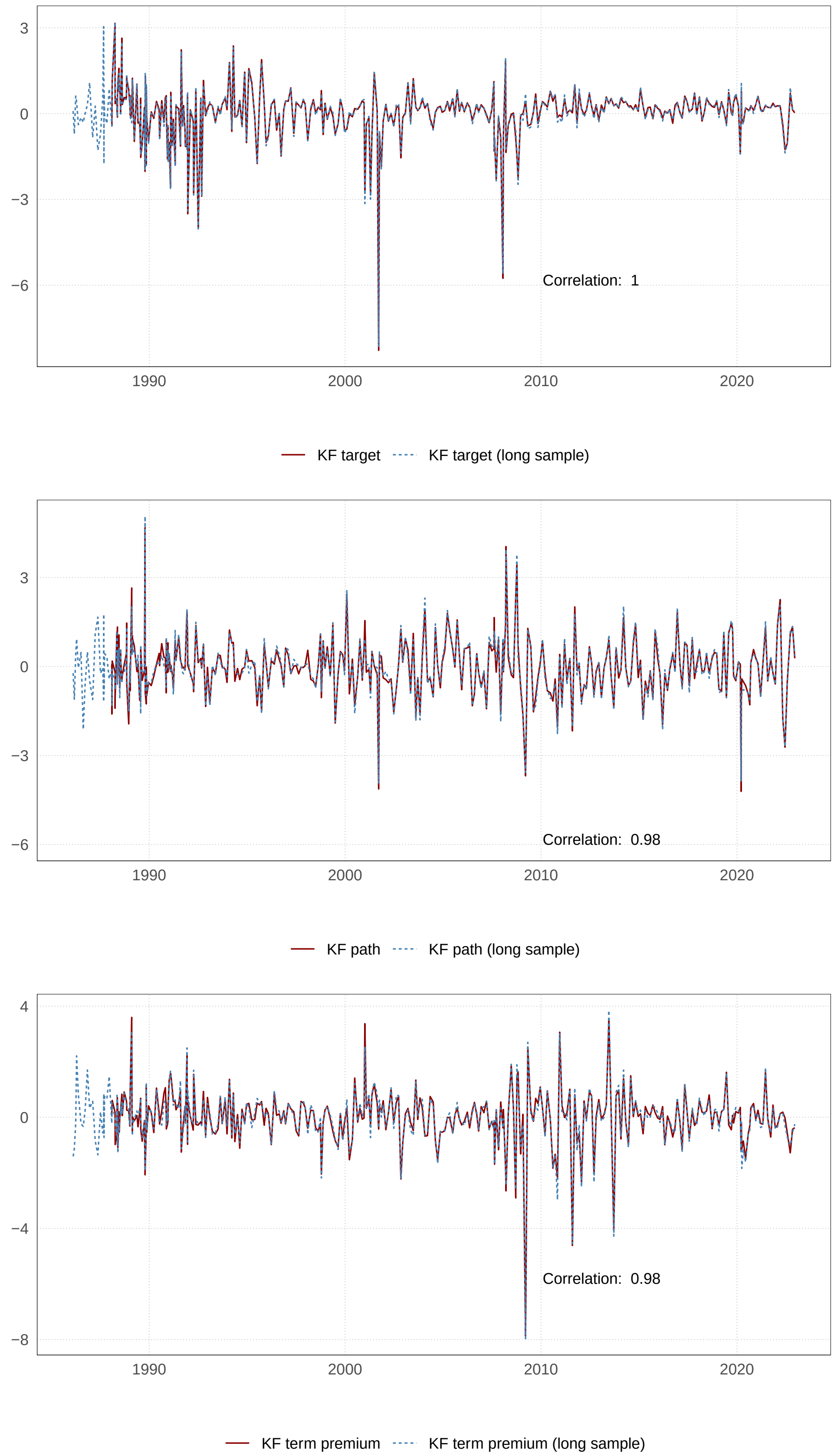
Fig. 12: Monthly impulse responses (continued on next page)

Fig. 12: Monthly impulse responses (continued from previous page)



Notes: Monthly impulse responses to monetary policy shocks (target, path, and term premium). The responses are normalized to a 25 bp increase in the 3M rate, 2Y rate, and 10Y - 2Y spread, respectively. The blue solid line shows the IV estimates using the heteroscedasticity-based shocks extracted using the Kalman filter. The red dotted lines are the IV estimates to the high-frequency surprises by Swanson (2021). In both models, we impose the recursive zero restrictions on the monthly impact response. The models are estimated in first (log-)differences, but the impulse responses are cumulated. Therefore, all interest rate responses are measured in percentage points and the other responses are measured in percent. 90% and 95% confidence intervals are based on HAC-robust standard errors.

Fig. 13: Heteroscedasticity-based shocks starting in 1986



Notes: Daily shocks extracted using the Kalman filter (KF) on a sample starting in 1988 and a more experimental sample starting in 1986. All series normalized to a mean of zero and a standard deviation of one.