Identification of Systematic Monetary Policy[*]

Lukas Hack Klodiana Istrefi Matthias Meier

This version: March 2026
First Version: January 2022

Abstract

We propose a novel identification design to estimate the effects of systematic monetary policy on the propagation of macroeconomic shocks. The design combines (i) a time-varying measure of systematic monetary policy based on the historical composition of hawks and doves in the FOMC with (ii) an instrument that leverages the FOMC rotation of voting rights. We apply our design to government spending shocks. We find that a dovish FOMC supports the expansionary effects of higher spending by delaying policy rate hikes, leading to large fiscal multipliers. GDP does not expand when the FOMC is hawkish, but inflation expectations are contained.

Keywords: Systematic monetary policy, FOMC, rotation, government spending.
JEL Codes: E32, E52, E62, E63, H56.

---

[*] An earlier version of this paper was titled "Systematic Monetary Policy and the Effects of Government Spending." We thank Regis Barnichon, Francesco Fusari, Anna Matzner, Karel Mertens, Alexander Meyer-Gohde, Filippo Palotti, and Josha van Spronsen for insightful discussions, and Klaus Adam, Antoine Camous, Antonio Ciccone, John Cochrane, Davide Debortoli, Refet Gürkaynak, Peter Karadi, Michael McMahon, Evi Pappa, Morten Ravn, Jon Steinsson and Christian Wolf, as well as participants at various seminars and conferences for helpful comments. Lukas Hack and Matthias Meier acknowledge funding from the German Research Foundation (DFG) through CRC TR 224 (Project C05). Matthias Meier acknowledges funding from DFG Projects 550469881/550470072 and from the UniCredit & Universities Foundation. The views expressed in this article are solely the responsibility of the authors, and should not be interpreted as reflecting the views of the ECB or the Eurosystem.
Lukas Hack (ETH Zurich and University of Mannheim): hack@kof.ethz.ch. Klodiana Istrefi (European Central Bank and CEPR): klodiana.istrefi@ecb.europa.eu. Matthias Meier (University of Mannheim and CEPR): m.meier@uni-mannheim.de.

1 Introduction

Monetary policy is not random but a purposeful response to macroeconomic conditions. This response represents systematic monetary policy. Fundamentally, the systematic response reflects the preferences of the policymakers, e.g., concerning price stability and employment, which change over time as the policymakers change. As a consequence, the effects of macroeconomic shocks differ across time, depending on systematic monetary policy. In theory, systematic monetary policy is well-known to be important for the propagation of macroeconomic shocks. However, there is no direct evidence on the causal effects of systematic monetary policy in the U.S.[1]

The main contribution of this paper is an identification design to estimate the causal effects of the Federal Reserve's systematic monetary policy on the propagation of macroeconomic shocks. We use historical fluctuations in the composition of hawks and doves in the Federal Open Market Committee (FOMC) to measure time variation in systematic monetary policy. To address the concern that these fluctuations are endogenous to economic and political developments, we propose an instrument that exploits the mechanical rotation of voting rights in the FOMC. To the best of our knowledge, our FOMC rotation instrument is the first instrument for systematic monetary policy.

We then apply the identification design to address a classical question in macroeconomics: How do the effects of fiscal policy depend on the response of monetary policy? This question is deemed crucial in the policy (e.g., Blinder, 2022) and academic debate (e.g., Woodford, 2011; Farhi and Werning, 2016). However, the debate lacks causal evidence. Providing causal evidence is the second contribution of this paper. We show that the Federal Reserve's systematic monetary policy has a significant effect on the GDP response to fiscal policy. When the FOMC is dovish, it delays tightening in response to an expansionary fiscal spending shock, which supports the expansion of GDP. Conversely, GDP does not expand but rather contracts under a hawkish FOMC that tightens faster and more aggressively. Fiscal multipliers are between two and three when the FOMC is dovish and indistinguishable from zero when it is hawkish.

We measure time variation in systematic U.S. monetary policy building on the narrative classification of FOMC members by Istrefi (2019), which uses news archives to classify members

---

[1] A vast empirical literature estimates the effects of monetary policy shocks (e.g., the pioneering work by Romer and Romer, 1989; Bernanke and Blinder, 1992; Cochrane and Piazzesi, 2002). These shocks are commonly understood as deviations from a policy rule, whereas most policy variation is due to systematic monetary policy, i.e., the rule itself. While evidence on monetary policy shocks may be informative about the effects of systematic monetary policy under certain assumptions (e.g., McKay and Wolf, 2023), we propose to directly estimate the causal effects of systematic monetary policy.

of the FOMC as hawks and doves for the period 1960 to 2023. Hawks are more concerned about inflation, while doves are more concerned about supporting employment and growth. Our measure of systematic monetary policy is the aggregate Hawk-Dove balance for each FOMC meeting. The Hawk-Dove balance is an appealing measure of systematic monetary policy because it parsimoniously summarizes the aggressiveness of the FOMC towards fulfilling one or the other leg of the dual mandate, without having to specify a policy reaction function or the policy tools.

Identifying the causal effects of systematic monetary policy, independent of how it is measured, is challenging because of endogeneity. For example, systematic monetary policy may change in response to unemployment or inflation (Davig and Leeper, 2008). Similarly, the appointment of central bankers can depend on economic and political circumstances, e.g., as documented for the Nixon administration (Abrams, 2006). We discuss this identification challenge through the lens of a New Keynesian model in which the coefficients of the monetary policy rule fluctuate in response to macroeconomic shocks. The model dynamics can be represented as a state-dependent local projection. The OLS estimates of the local projection will fail to identify the causal effects of systematic monetary policy because they are contaminated by unobserved shocks that change the monetary policy rule. Instead, we show that an instrument that captures exogenous variation in systematic monetary policy achieves identification.

We construct an instrument that exploits exogenous variation in the Hawk-Dove balance arising from the FOMC rotation of voting rights – an annual mechanical scheme that shuffles four out of twelve voting rights among eleven Federal Reserve Bank (FRB) presidents. The mechanical nature of the rotation renders it orthogonal to economic and political circumstances, and we provide direct evidence that our instrument is indeed uncorrelated with contemporaneous and lagged macroeconomic indicators and shocks. The rotation is also relevant: its correlation with the overall Hawk-Dove balance is 0.63, and the Fed watchers in the media track it closely. We further validate our measure of systematic monetary policy and the rotation instrument. We show that the federal funds rate (FFR) responds substantially more strongly to high inflation forecasts when the FOMC is hawkish, whereas the response to output gap forecasts is only weakly amplified. This suggests that our Hawk-Dove balance captures meaningful time variation in systematic monetary policy. These findings hold when instrumenting the Hawk-Dove balance by the rotation instrument but disappear for the corresponding OLS estimates, underscoring the practical relevance of our instrumental variable.

Our identification design combines the measure of systematic monetary policy and the instrument in a state-dependent local projection that can be applied to any macroeconomic shock of interest. Specifically, we regress an outcome of interest on the shock, the shock interacted with the Hawk-Dove balance, the Hawk-Dove balance in levels, and possibly further controls. The instrument vector is given by the vector of regressors when replacing the Hawk-Dove balance with the FOMC rotation instrument. This local projection is in line with the approximate dynamics of a broad class of DSGE models. However, different from DSGE models with time-varying systematic monetary policy, our design identifies the effects of systematic monetary policy imposing few structural assumptions. Our design allows studying the effects of policy counterfactuals, implemented through variation in the Hawk-Dove balance.

The main application of our identification design is to study the effects of government spending shocks in the U.S. We focus on the military spending shocks in Ramey (2011) and Ramey and Zubairy (2018) for the period 1960–2014. We find that the real GDP response depends significantly on systematic monetary policy.[2] The GDP response to an expansionary shock increases in the share of dovish FOMC members and decreases in the share of hawks. When the Hawk-Dove balance exceeds the sample average by two doves (roughly one standard deviation), quarterly GDP increases by up to 0.7% in response to the shock, which is expected to raise cumulative military spending by 1% of GDP over the next five years. Conversely, quarterly GDP falls by up to 0.3% when the Hawk-Dove balance exceeds the sample average by two hawks. In contrast to the IV estimates, OLS underestimates the dependence of the GDP response on systematic monetary policy at short horizons but overestimates it at longer horizons.

A common metric for assessing the effectiveness of fiscal spending is the spending multiplier, the dollar increase of real GDP per additional dollar of real government spending. We estimate the two- and four-year cumulative spending multipliers and find strong dependence on systematic monetary policy. While multipliers under a hawkish FOMC are typically insignificant with point estimates at or below 0, we find that dovish multipliers are between 2 and 3 and statistically significant. Moreover, the average multipliers are larger and much more precisely estimated when accounting for systematic monetary policy compared to a linear model that omits this state dependence. These results are robust to various modeling choices, as we show in an extensive sensitivity analysis.

We further inspect the mechanism behind the FOMC-dependent effects of spending shocks. We show that the FFR rises under a hawkish FOMC. Under a dovish FOMC, the FFR initially falls and rises only with substantial delay. When the Hawk-Dove balance exceeds the sample average by two hawks, the FFR starts to increase within one year after the

---

[2] In the post-Korean War sample, Ramey (2011) finds that these shocks have weak explanatory power for contemporaneous government spending. In contrast, we show that the shocks have statistically significant dynamic effects on government spending when accounting for time-varying systematic monetary policy.

shock, and increases by up to 52 basis points at a two-year horizon. Conversely, when the FOMC is dovish, the FFR falls and remains below the pre-shock level for more than two years after the shock, and then sharply rises toward a 47 basis point increase three years after the shock. The different interest rate responses are consistent with the fiscal multiplier estimates across hawkish and dovish FOMCs. Moreover, we find that hawkish policy is more successful in containing inflation (expectations). We further provide a calibrated New Keynesian model that rationalizes our evidence. Interest rate smoothing in the Taylor rule is a key model feature to turn short-lived differences in systematic monetary policy (through the rotation) into more persistent differences in interest rates, inflation, and GDP. Furthermore, we corroborate the plausibility of our results and the broad relevance of systematic monetary policy by applying our identification design to two additional shocks, a TFP shock and a broad demand shock.

Finally, we complement our quantitative analysis with narrative evidence from the historical records of the FOMC meetings. The records reveal that FOMC members and staff frequently discuss changes in military government spending, their potential impact on the economy and inflation, and the FOMC's policy response. We provide case studies of two important military spending buildup events in the 1960s, associated with the U.S. Space Program and the Vietnam War. The case studies suggest that the hawkish FOMC in the former event indeed tightened faster, whereas the dovish FOMC in the latter event delayed action.

Relation to literature. This paper contributes to a literature that aims to identify the effects of systematic monetary policy on the propagation of macroeconomic shocks. Closely related are McKay and Wolf (2023) and Barnichon and Mesters (2023) who use multiple monetary policy (news) shocks to estimate the effects of counterfactual monetary policy rules.[3] Under the assumption that systematic monetary policy affects private agents only through changes in the policy instrument, their approach allows identifying the effects of counterfactual interest rate paths. Instead, our approach leverages historical variation in systematic monetary policy. We estimate the causal state dependency with respect to systematic monetary policy, which allows constructing monetary policy counterfactuals. In particular, we can implement counterfactual FFR responses through shifts in the Hawk-Dove balance. Relative to McKay and Wolf (2023) and Barnichon and Mesters (2023), our approach avoids potential problems related to the identification and size of monetary policy shocks. The related literature highlights various identification problems for monetary policy shocks (Ramey, 2016). In fact, time variation in systematic monetary policy further chal-

---

[3] McKay and Wolf (2023) focuses on constructing policy counterfactuals, whereas Barnichon and Mesters (2023) studies optimal policy. Relatedly, Wolf (2023) uses the approach of McKay and Wolf (2023) to provide fiscal policy shock counterfactuals for a strict inflation-targeting central bank.

lenges identification (Bauer and Swanson, 2023; Hack et al., 2024). In addition, the effects of monetary policy shocks are typically small, particularly in recent decades. This may restrict the analysis to more modest policy counterfactuals to avoid extrapolation errors.

A closely related, earlier literature constructs monetary policy counterfactuals via monetary policy shocks (e.g., Bernanke et al., 1997; Kilian and Lewis, 2011). Yet, this approach is subject to the Lucas critique (Sargent, 1979). Our identification design is not subject to the Lucas critique because we explicitly model and estimate how the dynamics depend on systematic monetary policy. Another closely related paper is Cloyne et al. (2023), which leverages time-invariant cross-country differences in the policy rate response to fiscal shocks to estimate the role of systematic monetary policy on the propagation of fiscal consolidation shocks. Whereas Cloyne et al. (2023) leverages cross-country differences, we leverage exogenous historical variation in U.S. systematic monetary policy.

An alternative approach is to estimate time variation in systematic monetary policy through non-linear DSGE models (e.g., Davig and Leeper, 2007; Fernández-Villaverde et al., 2010; Bianchi, 2013), Taylor-rule regressions (e.g., Bauer et al., 2024), and VAR models (e.g., Primiceri, 2005) with systematic monetary policy modeled as a latent variable. The rotation of FOMC voting rights is often cited as a motivation for modeling such time variation. However, the empirical implementation does not directly model the rotation. In contrast, the identification strategy often involves assuming persistent processes for systematic monetary policy, which contradict the more short-lived fluctuations caused by the rotation. In contrast, our identification design explicitly leverages the rotation. Another common identifying assumption in the literature is that fluctuations in systematic monetary policy are exogenous. Instead, our instrumental variable approach addresses the potential endogeneity.[4]

We build on Istrefi (2019) for the classification of individual FOMC members as hawks and doves. Istrefi (2019) shows that these preferences match with narratives on monetary policy, preferred interest rates, dissents, and forecasts of FOMC members. Bordo and Istrefi (2023) study the origins of these preferences, linking them to early-life experiences and education. While Istrefi (2019) and Bordo and Istrefi (2023) introduce the Hawk-Dove balance, our key contribution is to construct an instrument for the Hawk-Dove balance that exploits the rotation. We further go beyond Istrefi (2019) and Bordo and Istrefi (2023) by combining the Hawk-Dove balance with the rotation instrument to propose a novel identification design, which we apply to study the effects of macroeconomic shocks. While the rotation of voting rights is a well-known institutional feature, our paper is the first to operationalize it for

---

[4] DSGE models that do allow for endogeneity include Davig and Leeper (2008), Barthélemy and Marx (2017), and Chang et al. (2021). In fact, the latter shows that imposing exogeneity is restrictive: macroeconomic shocks account for sizable variation in the monetary policy rule in their estimated DSGE model.

causal identification.

Finally, our paper relates to a large empirical literature that estimates the government spending multiplier. Most empirical estimates find an average fiscal spending multiplier between 0.5 and 1.5 (e.g., Blanchard and Perotti, 2002; Mountford and Uhlig, 2009; Ramey, 2011). Our findings show that the average fiscal spending multiplier may be downward-biased and substantially less precisely estimated when not accounting for time-varying systematic monetary policy. Further closely related are recent papers that study the effects of government spending shocks at the zero lower bound (e.g., Ramey and Zubairy, 2018; Miyamoto et al., 2018). Zero lower bound episodes are endogenous to the business cycle which means the estimates may reflect monetary policy but also the shocks leading to it. Instead, we isolate the causal effects of monetary policy on the propagation of fiscal policy. Another related paper is Nakamura and Steinsson (2014), which estimates relative regional multipliers that difference out the response of monetary policy. Our paper also relates to recent papers that estimate state-dependencies of the multiplier, e.g., depending on the economy being in a recession (Auerbach and Gorodnichenko, 2012; Jordà and Taylor, 2016; Ramey and Zubairy, 2018; Ghassibe and Zanetti, 2022); sign of the shock (Barnichon et al., 2022); exchange-rate regime, trade openness, and public debt (Ilzetzki et al., 2013); foreign holdings of debt (Broner et al., 2022); and tax progressivity (Ferriere and Navarro, 2025). Compared to this literature, our analysis tackles the endogeneity problem of the state variable. The state we consider captures the monetary policy reaction, and our results highlight the importance of fiscal-monetary interaction for macroeconomic stabilization and the role of who decides monetary policy.

The paper is organized as follows: Section 2 presents a New Keynesian model to illustrate the identification challenge. Section 3 introduces our identification design. Section 4 presents the main empirical results on fiscal shocks. Section 5 provides evidence on the mechanism. Section 6 provides a narrative of the FOMC records in the 1960s. Section 7 concludes.

2 Identification challenge

In this section, we present a stylized non-linear New Keynesian model in which systematic monetary policy may fluctuate endogenously. We use the model to expound the challenge of empirically identifying the effects of systematic monetary policy on the propagation of macroeconomic shocks.

A New Keynesian model. The model is a textbook New Keynesian model (e.g., Galí, 2015) except for a monetary policy rule with time-varying coefficients. Households choose consumption, labor and bond holdings to maximize $E_0 \sum_{t=0}^{\infty} \beta^t \left(\log C_t - N_t^{1+\varphi} \right)$ subject to budget constraints. Intermediate good firms produce variety goods using $Y_{it} = x_t^a N_{it}$ where x_t^a is exogenous productivity. The price of a variety good can be reset with a constant probability $1 - \theta$. Final good firms produce the final good $Y_t = \left(\int_0^1 Y_{it}^{(\epsilon-1)/\epsilon} di \right)^{\epsilon/(\epsilon-1)}$. A fiscal policy authority finances government spending $G_t = \gamma Y x_t^s$ with lump-sum taxes where $\gamma \in [0, 1)$, Y is steady-state output, and x_t^s denotes exogenous changes in fiscal spending. Goods market clearing requires $Y_t = C_t + G_t$. The exogenous variables follow stable AR(1) processes $\log x_t^k = \rho_k \log x_{t-1}^k + \varepsilon_t^k$ with $\varepsilon_t^k \sim (0, \sigma_k^2)$ for $k = a, s$ respectively. A monetary policy rule closes the model. Letting lowercase letters denote (log) deviations from the steady state, the monetary authority sets nominal interest rates i_t according to

$$i_t = \widetilde{\phi}_t \pi_t, \tag{2.1}$$

where $\widetilde{\phi}_t$ is *systematic monetary policy* which fluctuates according to a stable AR(1)

$$\phi_t = \rho_\phi \phi_{t-1} + \zeta^s \varepsilon_t^s + \zeta^a \varepsilon_t^a + \eta_t, \tag{2.2}$$

where $\widetilde{\phi}_t = \phi + \phi_t$ and $\phi \in (1, \infty)$ denotes the unconditional mean of $\widetilde{\phi}_t$. Importantly, we allow systematic monetary policy to be endogenous, as ϕ_t may respond to macroeconomic shocks $(\varepsilon_t^s, \varepsilon_t^a)$.[5,6] Such endogeneity creates an empirical identification challenge as we discuss toward the end of this section. In addition, we allow for exogenous changes in systematic monetary policy, captured by the exogenous policy shifter η_t.[7] We assume that ε_t^s, ε_t^a, and η_t are mutually independent and identically distributed over time. Accounting for the effects of systematic monetary policy ϕ_t, the approximate equilibrium dynamics of GDP are given by

$$y_t = a + b_s x_t^s + b_a x_t^a + c_s x_t^s \phi_t + c_a x_t^a \phi_t + d\phi_t, \tag{2.3}$$

where a, b_s, b_a, c_s, c_a, d are coefficients that depend on the deep structural parameters of the model. Appendix A.1 provides details on the derivation.

---

[5] For DSGE models with exogenous changes in the Taylor rule coefficients, see Davig and Leeper (2007) and Bianchi (2013); for endogenous changes, see Davig and Leeper (2008) and Barthélemy and Marx (2017).

[6] Hence, ϕ_t may indirectly depend on endogenous variables (e.g., inflation) because the shocks $(\varepsilon_t^s, \varepsilon_t^a)$ move these endogenous variables and ϕ_t at the same time if (ζ^s, ζ^a) are non-zero.

[7] The exogenous shifter to systematic monetary policy, η_t, is conceptually distinct from monetary policy shocks. While η_t changes the slope of the policy rule (2.1), a monetary policy shock changes the intercept. For further discussion on the distinction between monetary policy shocks and systematic monetary policy, see Hack et al. (2024).

Identification challenge. We next discuss the challenge of identifying the effects of systematic monetary policy from a regression when y_t is generated by (2.3). Without loss of generality, we focus our discussion on the fiscal spending shock. Consider an econometrician who observes $\{y_t, \varepsilon_t^s, \phi_t\}$, and estimates the state-dependent local projection

$$y_{t+h} = \alpha^h + \beta^h \varepsilon_t^s + \gamma^h \varepsilon_t^s \phi_t + \delta^h \phi_t + v_{t+h}^h, \tag{2.4}$$

for $h = 0, \ldots, H$ forecast horizons. For $h = 0$, the residual v_{t+h}^h contains lagged spending shocks, contemporaneous and lagged technology shocks, and the interaction of these shocks with ϕ_t. For $h > 0$, the residual further contains shocks (ε_t^s, ε_t^a) and policy shifter (η_t) occurring between t and $t + h$. The estimands in (2.4) are

$$\beta^h = b_s(\rho_s)^h \ , \qquad \gamma^h = c_s(\rho_s\rho_\phi)^h \ , \qquad \delta^h = d(\rho_\phi)^h \ . \tag{2.5}$$

Both β^h, the average effect of the spending shock, and γ^h, the differential effect associated with ϕ_t, diminish in the forecast horizon h.

We next ask whether the OLS estimates of $(\beta^h, \gamma^h, \delta^h)$ are consistent, i.e., whether they asymptotically recover the estimands in (2.5).[8] Consistency holds under the strong exogeneity assumption $\zeta^s = \zeta^a = 0$, that is, if ϕ_t is independent of the macroeconomic shocks. In contrast, if ϕ_t correlates with at least one of the shocks, the OLS estimates do *not* consistently estimate $(\beta^h, \gamma^h, \delta^h)$.[9] If, for example, ϕ_t responds to a spending shock, the OLS estimator will be contaminated by the response of GDP to the spending shock.

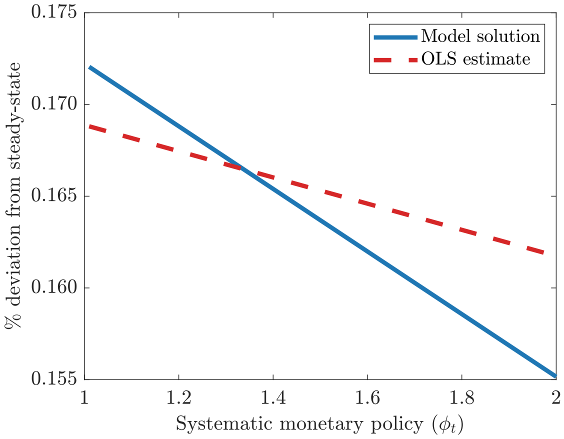
Now suppose the econometrician observes an instrument ϕ_t^{IV} that is correlated with ϕ_t (relevance), but uncorrelated with all past, contemporaneous, and future macroeconomic shocks $(\varepsilon_t^s, \varepsilon_t^a)$ and all past and future policy shifters η_t (exogeneity). Hence, the instrument is correlated with the contemporaneous policy shifter η_t. Consider the IV estimates of $(\beta^h, \gamma^h, \delta^h)$ when using $(\varepsilon_t^s, \varepsilon_t^s\phi_t^{IV}, \phi_t^{IV})$ as instrument vector for the regressors $(\varepsilon_t^s, \varepsilon_t^s\phi_t, \phi_t)$. The IV estimator consistently estimates $(\beta^h, \gamma^h, \delta^h)$, even when ϕ_t fluctuates endogenously in response to macroeconomic shocks $(\zeta^a, \zeta^s \neq 0)$. For further details, see Appendix A.2. This result guides the remainder of our paper in which we propose an instrument for systematic monetary policy and use it to estimate the causal effects of systematic monetary policy.

---

[8] We explicitly include δ^h in the vector of coefficients because including the (endogenous) control variable ϕ_t in the regression is important for identification, as ϕ_t is correlated with ε_t^s and $\varepsilon_t^s\phi_t$ in general.

[9] If the econometrician observes and includes *all* shocks and corresponding interaction terms in the regression according to equation (2.3), then the OLS estimates will be consistent without the exogeneity assumption. In practice, this is infeasible as many shocks are (partially) unobserved.

Figure 1: GDP response and systematic monetary policy



Notes: The solid line shows the model solution for the GDP response to a spending shock as a function of systematic monetary policy (ϕ_t), i.e., $b_s + c_s\phi_t$, with b_s and c_s given by (2.6) and the parametrization: $\beta = 0.99$, $\theta = 0.75$, $\epsilon = 9$, $\varphi = 2$, $\gamma = 0.2$, $\bar{\phi} = 1.5$, $\zeta^s = 1$, $\zeta^a = 0.25$, $\sigma_s = \sigma_a = 1$. The dashed line shows the OLS estimate $\hat{\beta}^0 + \hat{\gamma}^0\phi_t$ based on a regression of (2.4) when the terms in v^h_{t+h} are unobserved. The estimands are $\beta^0 = b_s = 0.164$ and $\gamma^0 = c_s = -0.017$, and the large-sample OLS estimates are $\hat{\beta}^0 = 0.164$ and $\hat{\gamma}^0 = -0.002$.

Illustration. To illustrate the effects of systematic monetary policy and the identification challenge, we focus on a special case in which $\rho_s = \rho_a = \rho_\phi = 0$. The GDP response to the fiscal spending shock ε^s_t, and its dependence on ϕ_t, is described by

$$b_s = \gamma\,(1 + \lambda\phi)\,\omega^{-1}\;, \qquad c_s = -\gamma(1-\gamma)\lambda\varphi\omega^{-2}\;, \tag{2.6}$$

where $\omega = 1 + \lambda\,(\varphi(1-\gamma) + 1)\,\phi$, $\lambda = (1-\theta)(1-\beta\theta)/\theta$. Since $b_s > 0$ and $c_s < 0$ (under standard parameter restrictions), the GDP response falls in the strength of the monetary policy reaction to inflation. This is the monetary offset (e.g., Woodford, 2011; Christiano et al., 2011).

The solid line in Figure 1 illustrates the monetary offset. The dashed line illustrates the OLS bias in the estimated GDP response to the spending shock. In our example, the OLS estimate strongly understates the role of systematic monetary policy.

Generalization to a broader class of models. The stylized New Keynesian model is not only useful to discuss the identification challenge, it also provides a structural justification for estimating a state-dependent local projection. Importantly, such a state-dependent local projection arises in a much broader class of models. In a companion note to this paper (Hack et al., 2026), we show that the class of models nests many models in the DSGE literature, including models with endogenous persistence and with heterogeneous agents, while allowing for time-varying systematic monetary policy.

3 Identification design

In this section, we propose (and validate) an identification design to study how systematic monetary policy in the U.S. shapes the propagation of macroeconomic shocks. Our identification design relies on three elements: (i) a measure of systematic monetary policy, (ii) an instrument for systematic monetary policy, and (iii) a state-dependent local projection regression that combines (i) and (ii) to tackle the identification challenge discussed in the preceding section.

3.1 Hawk-Dove balance in the FOMC

In the following, we build on the classification of FOMC members into hawks and doves by Istrefi (2019) and argue that the associated Hawk-Dove balance captures well variation in systematic monetary policy over time.

The FOMC. U.S. monetary policy is decided by the FOMC, which consists of 12 (voting) members: the seven members of the Board of Governors, including the Federal Reserve Chair, the president of the Federal Reserve Bank (FRB) of New York, and four of the remaining 11 FRB presidents, who serve one-year terms on a mechanically rotating basis. The other seven FRB presidents are (non-voting) FOMC participants. The FOMC's voting structure was established by the Banking Act of 1935 with the intention to balance centralized authority with regional representation and to reinforce the Federal Reserve's political independence through a more decentralized appointment process of FRB presidents (Fed, 2013).

Our analysis focuses on voting members, consistent with the literature on committee-based central bank decision-making (e.g., Belden, 1989; Blinder, 2007; Riboni and Ruge-Murcia, 2010, 2022; Bordo and Istrefi, 2023). Voting matters *prima facie* given that the composition of the FOMC receives substantial attention from financial markets and the media. In addition, we show that FRB presidents obtaining voting rights through the rotation participate significantly more actively in FOMC meetings, especially during policy deliberations, see Table D.1 in Appendix D. This evidence supports the view that voting rights matter, consistent with prior literature finding that the rotation matters for FRB presidents' public communication (Ehrmann et al., 2022), for predicting their policy decisions (Fos and Xu, 2025), and for the public reception of policy decisions (Madeira and Madeira, 2019). This motivates our focus on voting FOMC members.

Individual policy preferences. We measure time variation in (aggregate) systematic monetary policy based on the policy preferences of individual FOMC members. We use the

Istrefi (2019) classification of FOMC members as hawks and doves for the period 1960-2023 to measure individual preferences. Underlying the classification are articles from various U.S. newspapers and Fed watchers. Istrefi (2019) uses these articles to classify individual FOMC members as hawks or doves for each FOMC meeting, based on the news available up to that point.[10] This produces a panel tracking members over time at FOMC meeting frequency. The classification captures differences in policymakers' leanings (preferences) towards the two sides of the Fed's dual mandate. Hawks are perceived to be more concerned with inflation, while doves are more concerned with employment and growth.[11] Through the lens of our model in Section 2, we can think about hawks as preferring a larger inflation coefficient ϕ_t than doves.

Of the 149 FOMC members between 1960 and 2023, 132 are classified as hawks or doves, while 17 cannot be classified due to sparse early coverage or recent appointments. Most (98) maintain a consistent type as a hawk or dove, and the 34 who switch camps do so rarely—just 1.7% of member-meeting pairs—with swings evenly distributed across time and direction. Istrefi (2019) shows that these preferences match well with individual policy tendencies, as expressed by preferred interest rates, forecasting patterns, and dissents. In addition, Bordo and Istrefi (2023) shows that educational background and early-life experiences predict whether a FOMC member is a hawk or a dove. The long-lasting effect of early experiences suggests that we capture preferences that have stable meaning over time, which is consistent with the very few swings in our sample.

Hawk-Dove balance. We next aggregate individual FOMC members' preferences into an (aggregate) Hawk-Dove balance for each meeting (cf. Istrefi, 2019), reflecting that committee-based monetary policy results from the aggregation of individual preferences.[12]

First, we map the qualitative Hawk-Dove classification into a symmetric numerical scale, defining $Hawk_{i\tau}$ as the policy preference of member i at meeting τ. Our scale accounts for the view that swingers, i.e., FOMC members who switched from hawk to dove, or vice versa, are often considered more moderate, "middle-of-the-road" FOMC members, relative to consistent hawks or doves. We assign $Hawk_{i\tau} = 0$ when the policy preference of the

---

[10] Istrefi (2019) covers 1960–2014. We use the extension until 2023 by Bordo and Istrefi (2023).

[11] A typical example of a newspaper quote used to categorize a hawk reads: *Volcker leans toward tight-money policies and high interest rates to retard inflation*, New York Times, 2 May 1975. For a dove: *The weakness of Treasury prices and higher yields was seen reflecting the view that Bernanke will be 'pro-growth' and perhaps less hawkish on inflation, said John Roberts, managing director at Barclays Capital in New York*, Dow Jones Capital Markets Report, 24 October 2005.

[12] Relatedly, Blinder (1999) writes: *While serving on the FOMC, I was vividly reminded of a few things all of us probably know about committees: that they laboriously aggregate individual preferences; that they need to be led; that they tend to adopt compromise positions on difficult questions; and–perhaps because of all of the above–that they tend to be inertial.*

FOMC member is (yet) unknown.

$$Hawk_{i\tau} = \begin{cases} +1 & \textit{Consistent hawk} \\ +\frac{1}{2} & \textit{Swinging hawk} \\ 0 & \textit{Preference unknown} \\ -\frac{1}{2} & \textit{Swinging dove} \\ -1 & \textit{Consistent dove} \end{cases} \tag{3.1}$$

Second, we aggregate the individual policy preferences to the Hawk-Dove balance

$$Hawk_{\tau} = \frac{1}{|\mathcal{M}_{\tau}|} \sum_{i \in \mathcal{M}_{\tau}} Hawk_{i\tau} \ , \tag{3.2}$$

where $\mathcal{M}_{\tau}$ denotes the set of FOMC members at meeting τ.[13] The Hawk-Dove balance in (3.2) is the arithmetic average across individual preferences, which is our preferred aggregation as it conforms well with the consensual mode in which the FOMC typically operates.[14,15] We aggregate $Hawk_{\tau}$ from meeting frequency to quarterly frequency ($Hawk_t$ in quarter t) as the average balance in the first month of the quarter. If the first month is without a meeting, we use the first preceding month with a meeting.

We present the time series of the Hawk-Dove balance from 1960 to 2023 as the solid line in Figure 2. There is considerable variation in this balance, featuring both hawkish and dovish majorities. The variation reflects the turnover of rotating and non-rotating FOMC members, and changes in policy preferences of incumbent FOMC members. We discuss the importance of these components for $Hawk_t$ fluctuations in Section 3.2.

Hawk-Dove balance as measure of systematic monetary policy. We next argue that $Hawk_t$ is a good proxy for time variation in systematic monetary policy. First, the Hawk-Dove balance matches well with narratives of monetary policy in the U.S. (Istrefi,

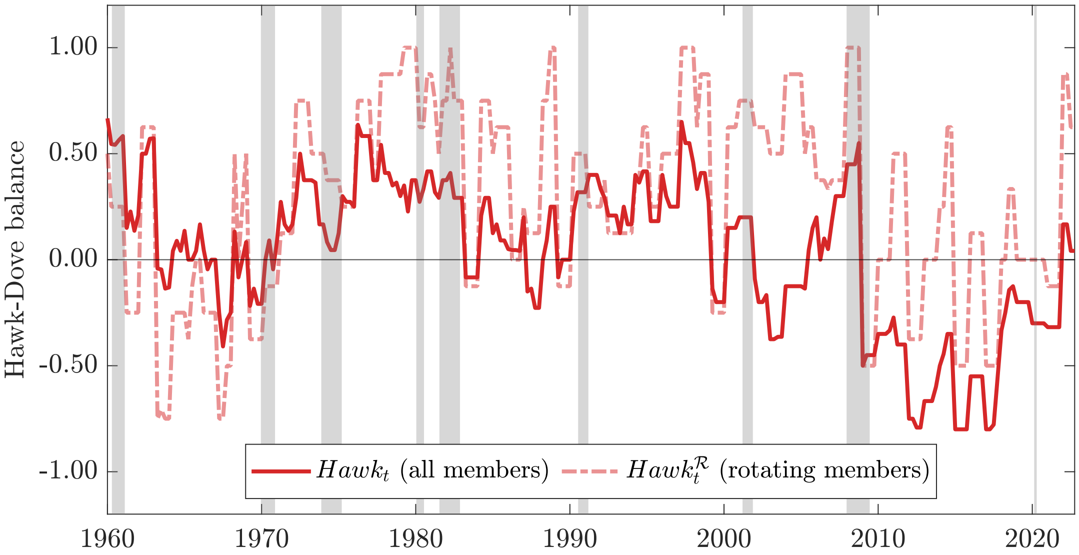
---

[13] A full FOMC consists of $|\mathcal{M}_{\tau}| = 12$ members but $|\mathcal{M}_{\tau}|$ is occasionally below 12 because of absent members or vacant positions. If a substitute temporarily replaces an absent FOMC member, we assume the substitute acts in the interest of the original FOMC member and assign the same policy preference, see Appendix B. This assumption affects less than 1% of all observations and is not important for our results.

[14] Riboni and Ruge-Murcia (2010) argues that a consensus model fits actual policy decisions of the Federal Reserve. In addition, Riboni and Ruge-Murcia (2022) provides evidence suggesting that policy proposals of the Fed Chair are the result of a compromise, reflecting a balance of power within the FOMC. Our empirical findings are robust to aggregating by the median or with a higher weight on the Fed Chair, see Section 4.5.

[15] Cieslak et al. (2023) constructs a Hawk-Dove score based on the language in FOMC meeting transcripts. While our measure captures FOMC members' preferences related to the Fed's dual mandate, their measure captures sentiments regarding the current direction of policy changes. Furthermore, Ferguson et al. (2023) classifies central bank governors in 80 countries as hawks and doves, with respect to financial sector support.

Figure 2: Hawk-Dove balance in the FOMC



Notes: The solid red line shows the quarterly time series of the aggregate Hawk-Dove balance of the FOMC ($Hawk_t$) from 1960 until 2023. The dashed red line shows the aggregate Hawk-Dove balance of the subgroup of rotating FRB presidents with voting right in period t, the FOMC rotation instrument ($Hawk_t^{\mathcal{R}}$). Grey bars indicate NBER dated recessions.

2019). For example, the dovish leaning of $Hawk_t$ in the mid-1960s coincides with the FOMC delaying anti-inflationary action (Meltzer, 2005). The hawkish majorities in the 1970s might be surprising given the high inflation rates in this period. Yet it is consistent with monetary policy being misguided by an underestimated natural rate of unemployment (DeLong, 1997) and persistence of inflation (Primiceri, 2006). In particular, Orphanides (2004) argues that for the periods before and after Paul Volcker's appointment in 1979, policy was broadly similar and consistent with a strong reaction to Greenbook inflation forecasts. Similarly, Coibion and Gorodnichenko (2011) finds no statistically significant difference in the Fed's response to inflation between the pre- and post-1979 period.[16] During the 1980s, the less hawkish FOMC reflects nominations of dovish Board members by President Reagan. In addition, it is consistent with the imperfect credibility of hawkish policy during the Volcker disinflation, as observed in persistently elevated long-term interest rates (indicative of inflation expectations) in this period (Goodfriend and King, 2005).

Second, observed fluctuations in $Hawk_t$ are strongly predictive of the FFR. In Hack et al. (2024), we show that a large share of variation in the FFR can be explained by interactions between $Hawk_t$ and various Greenbook forecasts, beyond the predictive content of the fore-

---

[16] Prominent proponents of the view that there was a structural break in the response to inflation include Taylor (1999) and Clarida et al. (2000). If monetary policy did respond more aggressively after 1979, that would imply measurement error in $Hawk_t$. However, such measurement error does not necessarily affect our findings to the extent that the IV estimates mitigate the potential bias arising from measurement error.

casts in levels. Importantly, fluctuations in the (non-interacted) level of $Hawk_t$ explain little variation in the FFR. These findings reinforce the view that $Hawk_t$ captures variation in the slope, rather than the intercept, of the policy rule.

Our approach of measuring systematic policy via $Hawk_t$ has several advantages over alternative approaches (e.g., Clarida et al., 2000; Coibion and Gorodnichenko, 2011; Bauer et al., 2024). Importantly, we do not have to specify a particular reaction function, nor do we need to restrict the analysis to specific policy instruments or communication strategies. We further avoid the well-known identification issues that plague the estimation of monetary policy rules (Cochrane, 2011; Carvalho et al., 2021). Independent of the policy tool or policy rule, our measure reflects the aggressiveness of the FOMC towards fulfilling one or the other leg of the dual mandate. The Hawk-Dove balance reflects public beliefs, in real time, about monetary policymakers. In contrast, ex-post estimates of systematic monetary policy may inadvertently use ex-post information not available at the time of the policy decision, potentially giving rise to misleading conclusions (Orphanides, 2003).

3.2 FOMC Rotation Instrument

We next propose and discuss a novel FOMC rotation instrument that allows us to identify the effects of systematic monetary policy when monetary policy is endogenous to the state of the economy (cf. Section 2).

Potential endogeneity. In general, systematic monetary policy may fluctuate for a broad range of reasons, among them macroeconomic conditions and political pressure (e.g., as highlighted in the Nixon tapes, see Abrams (2006)). The former may reflect various combinations of macroeconomic shocks. For example, the Federal Reserve may become more dovish in response to high unemployment, or more hawkish in response to high inflation (cf. Davig and Leeper, 2008). The latter may coincide with fiscal policy shocks (e.g., the U.S. President may aim to raise government spending while promoting an accommodative (dovish) monetary policy response).[17] The New Keynesian model in Section 2 illustrates that an empirical measure of systematic monetary policy afflicted by such endogeneity does not permit causal identification via OLS.

To understand how our measure of systematic monetary policy may reflect such endogeneity, we first categorize the reasons why it varies over time. $Hawk_t$ can vary because of new appointments (following resignations or term completions), the annual rotation of voting rights among FRB presidents, and swings in preferences by FOMC members across time.

---

[17] Relatedly, President Trump pressured the Fed (Bianchi et al., 2023; Camous and Matveev, 2021).

Table 1: Summary statistics

	Mean	Median	SD	Autocorr	Corr	Min	Max	T
$Hawk_t$	0.04	0.09	0.35	0.66	-	-0.80	0.67	252
$Hawk_t^{\mathcal{R}}$	0.28	0.33	0.46	0.22	0.63	-0.75	1.00	252

Notes: This table shows summary statistics for the quarterly time series from 1960 until 2023. $Hawk_t$ is the average Hawk-Dove balance of the FOMC. $Hawk_t^{\mathcal{R}}$ is the FOMC rotation instrument. Autocorr refers to the year-over-year autocorrelation. Corr refers to the correlation with $Hawk_t$.

Due to its mechanical nature, changes in $Hawk_t$ through the rotation of voting rights are exogenous to macroeconomic shocks. However, changes in $Hawk_t$ through new appointments and swings may be endogenous.

FOMC rotation instrument. To address the endogeneity of the Hawk-Dove balance, we propose an instrument that leverages exogenous variation in $Hawk_t$ that arises from the annual FOMC rotation. Each year, four FOMC memberships rotate among eleven FRB presidents following a mechanical scheme that has been in place since the early 1940s. According to the scheme, some FRB presidents become FOMC members every second year (Cleveland and Chicago) and others every third year (Philadelphia, Richmond, Boston, Dallas, Atlanta, St. Louis, Minneapolis, San Francisco and Kansas City). As the rotation of voting rights is independent of the state of the economy, it induces exogenous variation in $Hawk_t$. To leverage the variation from the FOMC rotation, we propose a novel instrument, which we refer to as the FOMC rotation instrument. Formally, the instrument is given by

$$Hawk_\tau^{\mathcal{R}} = \frac{1}{|\mathcal{R}_\tau|} \sum_{i \in \mathcal{R}_\tau} Hawk_{i\tau} \ , \tag{3.3}$$

where $\mathcal{R}_\tau$ denotes the set of rotating FOMC members at FOMC meeting τ.[18] We aggregate the FOMC rotation instrument to quarterly frequency analogously to $Hawk_t$.

In Figure 2, the dashed line presents the FOMC rotation instrument over time. On average, $Hawk_t^{\mathcal{R}}$ is more hawkish than $Hawk_t$, reflecting the fact that FRB presidents tend to be more hawkish than governors (Istrefi, 2019; Bordo and Istrefi, 2023). Both series display sizable variation over time, but fluctuations in $Hawk_t^{\mathcal{R}}$ are more short-lived, with a year-over-year autocorrelation of 0.22 compared to 0.66 for $Hawk_t$, see Table 1.

---

[18] A full set of rotating members consists of $|\mathcal{R}_\tau| = 4$ members. In our sample, $|\mathcal{R}_\tau| = 4$ for 625 out of 634 FOMC meetings and $|\mathcal{R}_\tau| = 3$ for the remaining nine meetings because of an absent member.

Relevance of instrument. Our instrument $Hawk_t^{\mathcal{R}}$ aggregates the policy preferences of one-third of the FOMC members, capturing a significant share of the variation in $Hawk_t$, with a correlation of 0.63 between the two.[19] We further study the explanatory power of the FOMC rotation instrument via a stylized first-stage regression by projecting $Hawk_t$ on $Hawk_t^{\mathcal{R}}$ and a constant. Applying the weak instrument test from Montiel Olea and Pflueger (2013) yields an effective F-statistic of 43.38 which is above 37.42, the critical value for rejecting a relative weak instrument bias exceeding 5%.[20] This suggests that the instrument satisfies the relevance condition. A more thorough assessment of instrument strength for our main results is delegated to Section 4.4.

Finally, the rotation is considered important by Fed watchers in the media. Each year before the rotation, they discuss its implications for monetary policy. A typical media discussion, here an article in The New York Times from January 1, 2011, reads as follows:

> *As the Federal Reserve debates whether to scale back, continue or expand its \$600 billion effort to nurse the economic recovery, four men will have a newly prominent role in influencing the central bank's path. The four men are presidents of regional Fed banks, and under an arcane system that dates to the Depression, they will become voting members in 2011 on the Federal Open Market Committee, [...] the change in voting composition is likely to give the committee a somewhat more hawkish cast. This could amplify anxieties about unforeseen effects of Bernanke's policies [...]. Two of the four new voters are viewed as hawkish on inflation, meaning that they tend to be more worried about unleashing future inflation than they are about reducing unemployment in the short run.*

Exogeneity of instrument. We next argue that variation in $Hawk_t^{\mathcal{R}}$ is quasi-exogenous. We provide four reasons: First, at the core of the argument is the fact that the rotation scheme is mechanical, following a time-invariant rotation rule, and therefore unrelated to macroeconomic shocks.[21] Second, new appointments of FRB presidents are relatively infrequent and unlikely to be influenced by political pressure from the federal government. FRB

---

[19] We further provide a decomposition of $Hawk_t$ into intensive margin changes of incumbent FOMC members' policy preferences and extensive margin changes in the composition of the FOMC due to appointments and the rotation, see Appendix C for details. The rotation accounts for 61% of the variance in yearly changes of $Hawk_t$, the turnover of non-rotating FOMC members for about a quarter of the variance, and the remainder is due to preference changes of incumbent FOMC members and various covariance terms.

[20] We also reject the null of the weak instrument bias exceeding 5% when adding four lags of $Hawk_t$ and $Hawk_t^{\mathcal{R}}$ to control for serial correlation in both variables. In either case, we use Newey-West standard errors with automatic bandwidth selection.

[21] For example, the large drop in $Hawk_t^{\mathcal{R}}$ during the Great Recession reflects the rotation. We observe four consistent hawks rotating out and three consistent doves and one hawk rotating in; all appointed before the Great Recession and none having switched preferences, making this drop in $Hawk_t^{\mathcal{R}}$ credibly exogenous.

presidents are nominated by the Board of Directors of the respective Federal Reserve district and confirmed by the Board of Governors. The directors are to represent the financial institutions and the broader public in the district.[22] In contrast, members of the Board of Governors (including the Fed Chair) are nominated by the U.S. President and confirmed by the Senate, rendering their appointment process more prone to be influenced by macroeconomic shocks. Furthermore, the average effective tenure of an FRB president is eleven years, but only seven years for a governor in our sample. Relatedly, Bordo and Istrefi (2023) shows that different from governors, there is no correlation between the preferences of the FRB presidents and the U.S. president's party at the time of their appointment. In addition, to the extent that some regional FRBs have persistent leanings toward either the dovish or the hawkish camp, e.g., hawks in Cleveland and doves in San Francisco, the appointment of new FRB presidents is even less likely to be affected by macroeconomic conditions.

Third, we argue that swings are a negligible threat to the exogeneity of our instrument. For rotating FOMC members, swings occur only in 1.3% of member-meeting pairs.[23] In addition, we find that swings account for a negligible fraction of the variance of the rotation instrument.[24] In addition, among the few swings that did happen, some do not appear linked to the state of the economy.[25]

Fourth, we address residual concerns related to appointments and swings by testing the predictability of $Hawk_t^{\mathcal{R}}$ by macroeconomic indicators and shocks. We project $Hawk_t^{\mathcal{R}}$ on the contemporaneous value and four quarterly lags of a given predictor. The set of predictors includes conventional macroeconomic indicators such as real GDP, inflation, and interest rates. Similarly, we consider a rich set of empirical shock measures from the literature, e.g., shocks to productivity, taxation, and oil supply. The complete list of predictors and the estimation results are presented in Figure D.1 in Appendix D. The results show that there is no meaningful predictability of our instrument. Out of the 120 tested coefficients, we find that only 5 are statistically significant, below the Type I error rate of 5% that one should

---

[22] Formally, the Board of Governors approves the appointments of FRB presidents. In the words of former Governor Kevin Warsh *it would be reasonably unprecedented in modern times, for the Reserve Bank's preferred choice not to ultimately be accepted by the Board of Governors* (Bordo, 2016).

[23] Specifically, in 2,527 member-meeting observations, we observe only 33 swings.

[24] We decompose $Hawk_t^{\mathcal{R}}$ into swings and extensive margin changes due to the rotation and appointments. The rotation accounts for 93% of the variance in yearly changes of $Hawk_t^{\mathcal{R}}$, appointments for 7%, and swings account for a negligible fraction of the variance, see Appendix C for details.

[25] Bordo and Istrefi (2023) discuss three major swing waves in the FOMC during 1960-2014. The first wave is a hawkish wave influenced by inflation dynamics in the late 1960s to early 1970s. The second wave is a hawkish swing in the early 1990s, related to the discussion on inflation targeting inspired by the announcements of the Reserve Bank of New Zealand and Bank of Canada. Finally, the third swing wave is a dovish one in the late 1990s, following a new understanding of the economy.

expect under exogeneity of the instrument.[26,27]

Overall, the above arguments support the validity of our FOMC rotation instrument for identifying the causal effects of systematic monetary policy. To the best of our knowledge, this paper is the first to propose an instrument for systematic monetary policy. We believe this is a substantial contribution to the literature which opens up myriad research questions.

3.3 Validation exercise for $Hawk_t$ and $Hawk_t^{\mathcal{R}}$

We provide a validation exercise for $Hawk_t$ and $Hawk_t^{\mathcal{R}}$ with four key takeaways. First, it supports the case that $Hawk_t$ and $Hawk_t^{\mathcal{R}}$ capture time variation in systematic monetary policy (the slope of the policy rule). Second, the results aid the interpretation of our Hawk-Dove balances, suggesting we primarily capture variation in the Fed's response to inflation. Third, the results reject the alternative hypothesis that our measure captures conventional monetary policy shocks (shifting the intercept of the policy rule). Finally, the results highlight the importance of accounting for endogeneity in systematic monetary policy.

We regress the FFR on forecasts of inflation and output gap, allowing for state dependence with respect to $Hawk_t$. Formally, we estimate

$$\begin{aligned} FFR_{t+h} =& \alpha^h + \beta_\pi^h \hat{\pi}_t + \beta_y^h \hat{y}_t + \gamma_\pi^h \hat{\pi}_t (Hawk_t - \overline{Hawk}) + \gamma_y^h \hat{y}_t (Hawk_t - \overline{Hawk}) \\ &+ \delta^h (Hawk_t - \overline{Hawk}) + \zeta^h Z_{t-1} + v_{t+h}^h, \end{aligned} \tag{3.4}$$

for $h = 0, 1, ..., H$ forecast horizons. FFR_t denotes the effective FFR, $\hat{\pi}_t$ and $\hat{y}_t$ denote the (demeaned) Greenbook forecasts of inflation and the output gap, $\overline{Hawk}$ denotes the arithmetic sample mean of $Hawk_t$, and Z_{t-1} includes four lags of the FFR and four lags of both Greenbook forecasts.[28] The sample runs from 1969Q1 through 2008Q4.[29]

The coefficients of interest are β_π^h, β_y^h, γ_π^h, γ_y^h. The former two capture the change in the FFR following an increase in Greenbook forecasts given a Hawk-Dove balance equal to its sample average. The latter two capture the differential change in the FFR following an increase in Greenbook forecasts and in the Hawk-Dove balance. We estimate equation (3.4) by OLS or IV using $Hawk_t^{\mathcal{R}}$ as instrument.[30] All estimates are normalized to correspond to an increase in the Greenbook forecasts of inflation and output gap by one percentage point, respectively,

---

[26] When repeating this exercise using the overall Hawk-Dove balance, we find more statistically significant predictors, in line with the above-discussed endogeneity concerns, see Figure D.2.

[27] Our main results are robust to residualizing $Hawk_t^{\mathcal{R}}$ with respect to the few significant macroeconomic indicators and shocks, see Section 4.5.

[28] The forecast horizon of regressors $\hat{\pi}_t$ and $\hat{y}_t$ is three-quarters ahead. See Appendix B for further details.

[29] The sample start is the first quarter for which both Greenbook forecasts are available based on Coibion and Gorodnichenko (2011). The sample is chosen to end before the ZLB episode.

[30] The instrument vector corresponds to the vector of regressors when replacing $Hawk_t$ with $Hawk_t^{\mathcal{R}}$.

and an increase in the Hawk-Dove balance by two hawks.

Figure 3 shows the estimates together with 68% and 95% confidence bands using Newey-West standard errors. Panels (a) and (b) show the IV estimates of β_π^h and β_y^h. Given an average Hawk-Dove balance, we find that the FFR responds more strongly to inflation forecasts than to output gap forecasts. This is in line with related estimates in the literature (e.g., Coibion and Gorodnichenko, 2011) and common parametrizations of Taylor rules. A one percentage point higher inflation forecast roughly raises the FFR by one percentage point at horizons between one and two years.

The most important results of the validation exercise are presented in panels (c) and (d). High inflation forecasts in periods of a hawkish FOMC are associated with a strongly amplified increase in the FFR. The IV estimates of γ_π^h are significantly different from zero at the 5% level for horizons up to three years. Relative to β_π^h, the estimated magnitudes of γ_π^h mean that a one percentage point higher inflation forecast raises the FFR by approximately two percentage points (instead of one) at horizons between one and two years if two more hawks are in the FOMC. In contrast, the differential response of the FFR to output gap forecasts, γ_y^h, is relatively small and significant only for the first year. In stark contrast to the IV estimates, the OLS estimates of γ_π^h and γ_y^h are mostly insignificant, suggesting that accounting for the endogeneity of $Hawk_t$ is crucial. This reinforces the importance of our instrumental variable.

Finally, panel (e) shows the estimated FFR response to $Hawk_t$ (not interacted with Greenbook forecasts). The estimates are close to zero and statistically insignificant, suggesting $Hawk_t$ does not capture variation in the intercept of the policy rule (monetary policy shocks).

3.4 Local projection framework

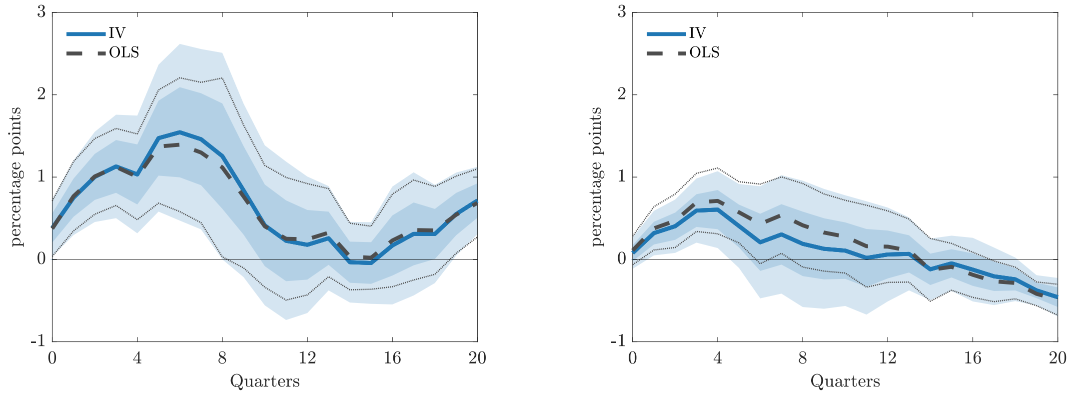
The final element of our identification design is an estimation approach to characterize how the propagation of macroeconomic shocks is causally shaped by systematic monetary policy. We propose to combine $Hawk_t$ and $Hawk_t^{\mathcal{R}}$ in a state-dependent local projection that is consistent with the New Keynesian model discussed in Section 2.

We propose to regress an outcome of interest x_{t+h} on a shock of interest ε_t^s, the interaction of the shock with the Hawk-Dove balance $Hawk_t$, as well as $Hawk_t$ in levels and a vector of additional control variables Z_{t-1}. Formally, the state-dependent local projection is given by

$$x_{t+h} = \alpha^h + \beta^h \varepsilon_t^s + \gamma^h \varepsilon_t^s (Hawk_t - \overline{Hawk}) + \delta^h (Hawk_t - \overline{Hawk}) + \zeta^h Z_{t-1} + v_{t+h}^h, \quad (3.5)$$

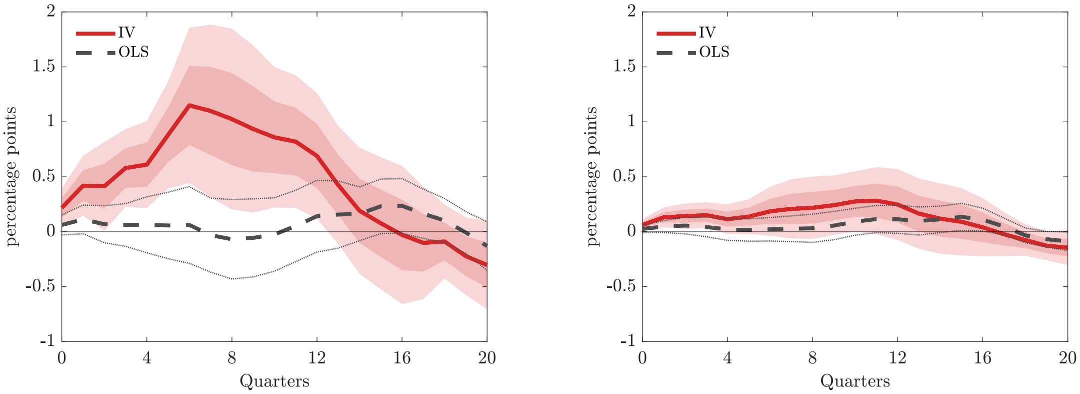
for $h = 0, \ldots, H$ forecast horizons. Equation (3.5) resembles (3.4), but with the important distinction that we consider the response to shocks instead of forecasts. To address the

Figure 3: FFR response to Greenbook forecasts



(a) Average response to $\hat{\pi}$ (β_π^h)

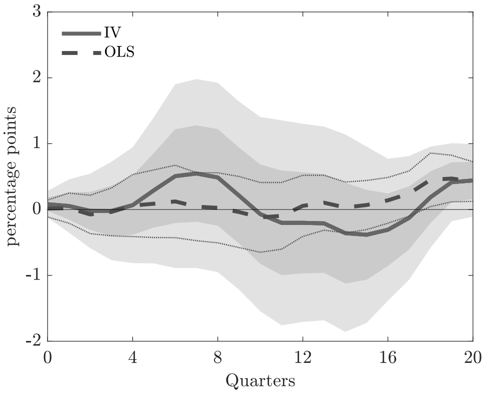
(b) Average response to $\hat{y}$ (β_y^h)



(c) Differential response to $\hat{\pi}$ (γ_π^h)

(d) Differential response to $\hat{y}$ (γ_y^h)

(e) Response to $Hawk_t$ (δ^h)



Notes: The figure shows the dynamic behavior of the FFR following (demeaned) Greenbook forecasts of inflation ($\hat{\pi}$) and the output gap ($\hat{y}$) that are one percentage point above their respective sample averages, conditional on systematic monetary policy ($Hawk_t$). We show IV and OLS estimates based on (3.4) as solid and dashed lines, respectively. The coefficient β_i^h ($i = \hat{\pi},\ \hat{y}$) captures the responses when $Hawk_t$ equals its sample average, γ_i^h the differential responses when $Hawk_t$ exceeds the sample average by two hawks, and δ^h the response to $Hawk_t$ exceeding the sample average by two hawks conditional on the Greenbook forecasts. The shaded areas indicate 68% and 95% confidence bands of the IV estimates using Newey-West standard errors. The thin gray lines indicate the corresponding 95% confidence bands of the OLS estimates.

potential endogeneity of $Hawk_t$, we use the instrument vector

$$q_t = \left[\ 1,\ \varepsilon_t^s,\ \varepsilon_t^s \left(Hawk_t^{\mathcal{R}} - \overline{Hawk}^{\mathcal{R}}\right),\ \left(Hawk_t^{\mathcal{R}} - \overline{Hawk}^{\mathcal{R}}\right),\ Z_{t-1}\ \right] \tag{3.6}$$

for the regressors in (3.5). The two key coefficients in (3.5) are β^h and γ^h, which capture the average response, when the Hawk-Dove balance equals its sample average, and the differential response, when the FOMC is more hawkish than the sample average.[31]

Based on Section 2, the IV estimator is consistent if the instrument $Hawk_t^{\mathcal{R}}$ is orthogonal to all macroeconomic shocks (both observed shocks ε_t^s and other unobserved shocks) at all lags and leads. In the next section, we discuss whether the identifying assumptions are satisfied in the context of a government spending shock.

In general, this framework can be used to study the propagation of any shock through systematic U.S. monetary policy. Our framework permits revisiting a range of important empirical questions, such as the role of systematic monetary policy for the effects of oil-related shocks (e.g., Bernanke et al., 1997; Kilian and Lewis, 2011), technology shocks (e.g., Galí et al., 2003), news shocks (e.g., Barsky and Sims, 2011), fiscal spending shocks (e.g., Ramey and Zubairy, 2018), and tax shocks (e.g., Romer and Romer, 2010). Moreover, our framework allows the estimation of a new set of moments that can be used to discipline structural models with time variation in systematic monetary policy, such as regime-switching models (e.g., Davig and Leeper, 2007; Bianchi, 2013; Bianchi and Ilut, 2017).

4 Government spending and monetary policy

In this section, we use our identification design to estimate how the effects of U.S. government spending shocks depend on systematic monetary policy. We find that a hawkish FOMC significantly dampens the expansionary effects of increased government spending on GDP, while a dovish FOMC supports it. Relatedly, we find sizeable differences in the fiscal multiplier depending on the hawkishness or dovishness of the FOMC. We further provide evidence on the strength of our instrument and perform an extensive sensitivity analysis.

4.1 Data and identifying assumptions

We next discuss the data (in addition to $Hawk_t$ and $Hawk_t^{\mathcal{R}}$) and the identifying assumptions for our analysis of government spending shocks.

---

[31] Formally, we define state-dependent impulse responses as $\mathbb{E}\left[x_{t+h}|\varepsilon_t^s = \varepsilon, Hawk_t = Hawk\right] - \mathbb{E}\left[x_{t+h}|\varepsilon_t^s = 0, Hawk_t = Hawk\right] = \left[\beta^h + \gamma^h\left(Hawk - \overline{Hawk}\right)\right]\varepsilon$, where both expectations additionally condition on the control vector Z_{t-1}.

Data. Our analysis uses the local projection framework in equations (3.5)-(3.6). Our baseline shock of interest, ε_t^s in (3.5), is the military spending shock constructed by Ramey (2011) and Ramey and Zubairy (2018), based on a narrative approach to identify surprise build-ups (or build-downs) in U.S. military spending. The shock is constructed as the present value of expected changes in real defense spending over the next years, typically up to a horizon of five years, and expressed relative to real potential GDP. The two primary outcome variables of interest, x_{t+h} in (3.5), are real GDP and real government spending, both expressed relative to real potential GDP. The vector of control variables, Z_{t-1} in (3.5), includes four lags of real GDP and real government spending, both relative to potential output, and four lags of the fiscal spending shock. If we restrict $\gamma^h = \delta^h = 0$, our specification of (3.5) corresponds to equation (1) of Ramey and Zubairy (2018). This facilitates the comparability of our results with the literature.

Our baseline sample covers the period from 1960Q1 to 2014Q4, which is the longest possible sample for which the Hawk-Dove balance and the fiscal spending shocks are available. Our sample includes important military spending shocks, e.g., the Vietnam War, the Carter-Reagan military buildup, and 9/11. On the other hand, our sample excludes WWII and the Korean War which are important events in Ramey (2011) and Ramey and Zubairy (2018).[32] However, we consider it desirable to exclude these events from our analysis because monetary policy was less autonomous from fiscal policy prior to the Treasury-Fed Accord in 1951.

Identifying assumptions. Two key identifying assumptions need to be satisfied to assign causal interpretation to the estimates of β^h and γ^h in (3.5). The first assumption is that the FOMC rotation instrument is orthogonal to all macroeconomic shocks, at all leads and lags. This assumption is strongly supported by the arguments and evidence in Section 3.2.[33]

The second assumption is that military spending shocks are random shocks. In particular, the distribution of military spending shocks may not depend on systematic monetary policy. According to Ramey and Zubairy (2018), military spending shocks are unanticipated changes in spending plans triggered by geopolitical events and therefore exogenous to the economy. This argument similarly applies when conditioning on systematic monetary policy. We provide three additional arguments for the shocks being independent of systematic mone-

---

[32] Ramey (2011) shows that excluding the Korean War renders military spending shocks a weak instrument for contemporaneous government spending. In general, it is not surprising that the military spending shock is a weak instrument for contemporaneous government spending because the shock largely pertains to future spending. Therefore, we do not use the military spending shock as an instrument but as a shock in our local projection framework (3.5). We find a significant dynamic government spending response, see Section 4.2.

[33] In the presence of other time-varying state-dependencies, e.g., time variation in the slope of the Phillips curve, our instrument must also be orthogonal with respect to such variation. Given the mechanical nature of the rotation and the fact that fluctuations in $Hawk_t^{\mathcal{R}}$ are relatively short-lived and uncorrelated with various business cycle indicators, our instrument plausibly also satisfies this condition.

tary policy: (i) the response of military spending to the shock does not significantly depend on systematic monetary policy (see Section 4.2); (ii) the news quotes used to construct military spending shocks do not mention monetary policy, the Federal Reserve, or the FOMC in our sample, see the supplementary appendix to Ramey and Zubairy (2018); and (iii) $Hawk_t$ does not predict spending shocks. Regressing military spending shocks ε^s_{t+h} on $Hawk_t$, and using $Hawk^{\mathcal{R}}_t$ as an instrument, we find no significant effects of the $Hawk_t$ on contemporaneous or future military spending shocks, see Figure E.1 in Appendix E.

4.2 Responses of GDP and government spending

We next present our empirical estimates of the causal effects of systematic monetary policy on the responses of real GDP and real government spending to fiscal spending shocks. We find that expansionary spending shocks raise GDP more strongly when the FOMC is dovish.

IV estimates. Figure 4 shows the IV estimates of the responses of real GDP and real government spending (G) to a military spending shock conditional on systematic monetary policy ($Hawk_t$). The estimates are based on the local projection framework (3.5)-(3.6) as specified in Section 4.1. The solid lines show the point estimates and the shaded areas indicate 68% and 95% confidence bands using Newey-West standard errors.[34] All estimates of β^h and γ^h are normalized to correspond to an expansionary shock that raises the expected present discounted value of future military spending by one percent of potential GDP.

Panels (a) and (b) show the estimates of β^h for GDP and G, which capture the responses when $Hawk_t$ equals its sample average. The average responses of both GDP and G are positive and significantly different from zero at most horizons beyond the first year. Both responses build up gradually and exceed 0.14% for GDP and 0.09% for total G after one year. The response of military G (dashed line) resembles total G, meaning the expansion of total G primarily reflects higher military G.[35]

Panels (c) and (d) show the estimates of γ^h, which capture the differential responses of GDP and G when increasing the Hawk-Dove balance by two hawks.[36] Importantly, the GDP response is lower after a fiscal expansion when the FOMC is more hawkish. This effect is statistically significant at the 5% level until three years after the shock. The estimated

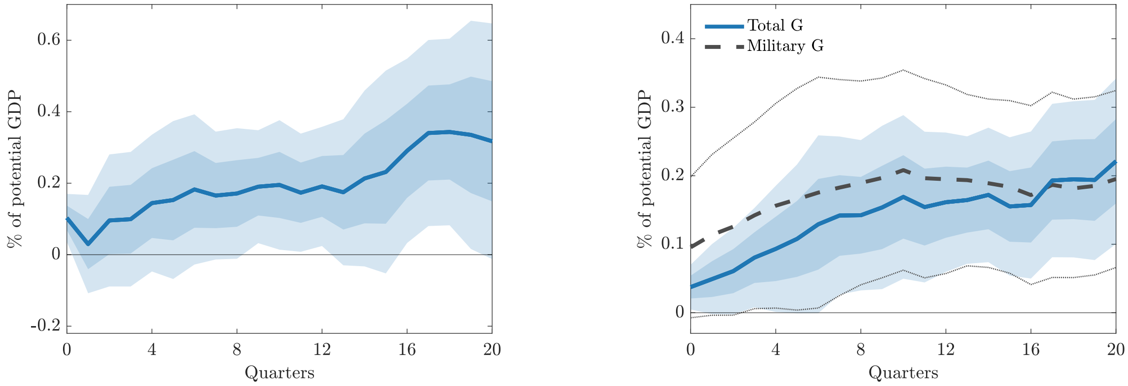
---

[34] For the Newey-West standard errors, we set the bandwidth to $h+1$, where h is the horizon in (3.5). A truncation parameter rule (Lazarus et al., 2018) or automatic bandwidth selection leads to similar results.

[35] Military G is defined relative to real potential GDP, analogous to total G, see Appendix B for details.

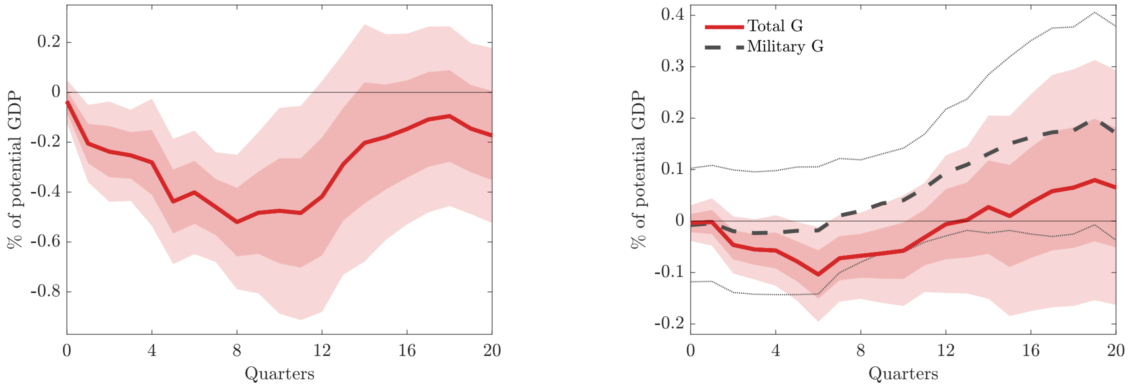
[36] Specifically, γ^h is scaled to capture an increase in $(Hawk_t - \overline{Hawk})$ of 2/12. This means, for example, that two FOMC members with unknown preferences are replaced by two consistent hawks, or that two FOMC members swing from dovish to hawkish. An increase in $Hawk_t$ by 2/12 slightly exceeds one standard deviation of the change in $Hawk_t$, which is 0.15.

Figure 4: Responses to spending shocks conditional on monetary policy



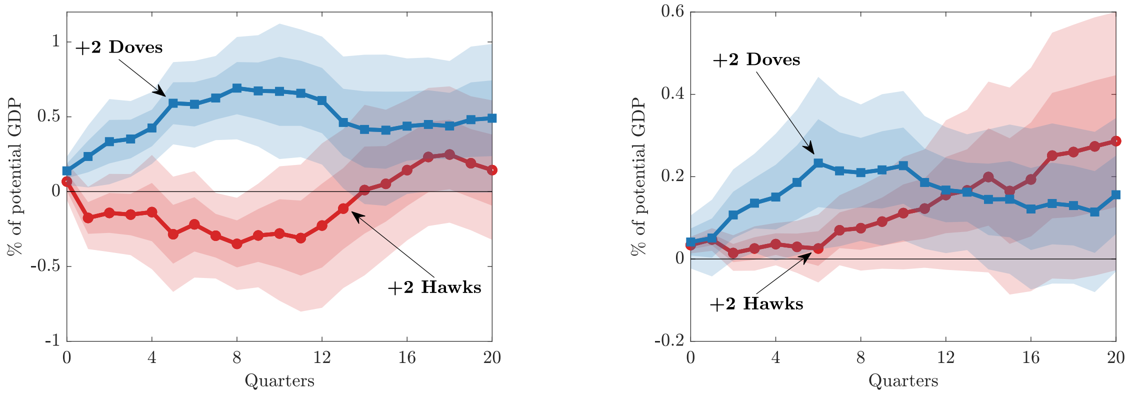
(a) Average GDP response (β^h)

(b) Average G response (β^h)



(c) Differential GDP response (γ^h)

(d) Differential G response (γ^h)



(e) State-dependent GDP response ($\beta^h \pm \gamma^h$)

(f) State-dependent G response ($\beta^h \pm \gamma^h$)

Notes: The figure shows responses of real GDP and real government spending (G), separately for total G and military G, to an expansionary military spending shock, corresponding to one percent of potential GDP, conditional on systematic monetary policy ($Hawk_t$). We show IV estimates based on the local projection framework (3.5)-(3.6) as specified in Section 4.1. The β^h captures the responses when $Hawk_t$ equals its sample average. The γ^h captures the differential responses when $Hawk_t$ exceeds the sample average by two hawks. The $\beta^h \pm \gamma^h$ shows the state-dependent responses when $Hawk_t$ exceeds the sample average either by two hawks (+2 Hawks) or by two doves (+2 Doves). The shaded areas indicate 68% and 95% confidence bands using Newey-West standard errors. The dotted lines indicate 95% confidence bands for military G.

magnitudes are sizable. Between two and three years after the shock, the GDP response is around 0.4% lower under a more hawkish FOMC. Conversely, the GDP response is 0.4% higher when there are two more doves in the FOMC. The differential response of (total) government spending (G) is also negative at horizons until three years after the shock, albeit smaller in absolute terms and less significant. Focusing on military G, the estimated differential response is insignificant at all horizons. This result supports our identifying assumption that the military spending shock does not depend on $Hawk_t^{\mathcal{R}}$. The difference between the estimates for total and military G implies a negative differential response of non-military G. This response is unsurprising in an environment of tighter monetary policy and constitutes a part of the transmission of systematic monetary policy.

Panels (e) and (f) of Figure 4 show $\beta^h \pm \gamma^h$, the state-dependent responses when $Hawk_t$ exceeds the sample average either by two hawks (+2 Hawks) or by two doves (+2 Doves). The GDP response strongly varies between the dovish and the hawkish FOMC. The dovish FOMC supports the GDP expansion while the hawkish FOMC undoes the GDP expansion. Quantitatively, GDP increases by up to 0.69% under the dovish FOMC, but falls by up to 0.35% under the hawkish FOMC. The former response is highly statistically significant, whereas the latter response is less precisely estimated.

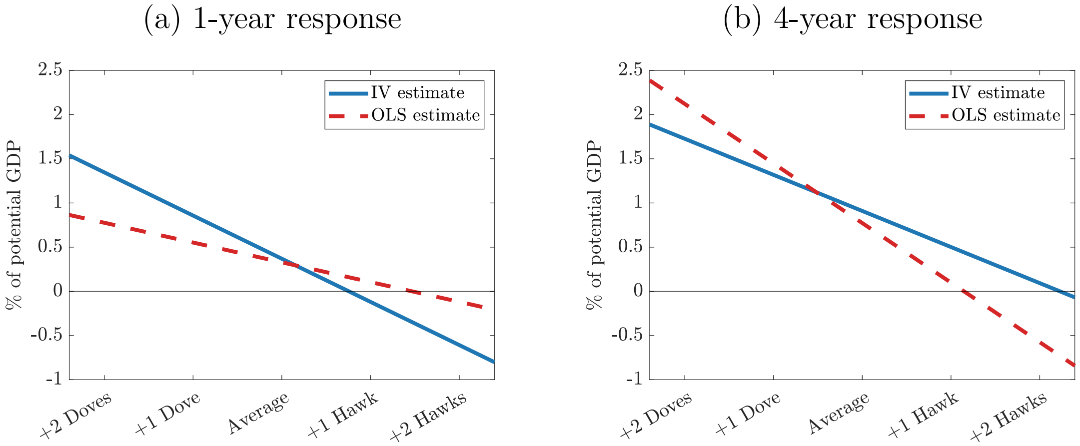
Overall, our evidence suggests that the monetary offset of fiscal spending shocks is not a constant feature of monetary policy but varies strongly with the Hawk-Dove balance in the FOMC. In contrast to the GDP response, government spending displays smaller and less significant differences in the state-dependent responses.

Comparison with OLS. We compare our IV estimates with OLS estimates. Figure 5 shows the response of GDP as a function of the FOMC's Hawk-Dove balance in the first and fourth year after the shock. In the first year, the OLS estimates substantially understate the dependence of the GDP response on the Hawk-Dove balance. In contrast, the OLS estimates overstate this dependence in the fourth year. This comparison suggests that ignoring the endogeneity of $Hawk_t$ leads to biased conclusions about the role of systematic monetary policy for fiscal spending shocks.

4.3 Fiscal spending multiplier

A key object for the design and evaluation of fiscal policies is the fiscal spending multiplier. We use our framework to estimate how the fiscal spending multiplier depends on the hawkishness of the FOMC. We find that a dovish FOMC leads to substantially larger multipliers, relative to an average or a more hawkish FOMC composition.

Figure 5: GDP responses for OLS and IV


(a) 1-year response

(b) 4-year response

Notes: The figure shows the yearly real GDP response to an expansionary military spending shock, corresponding to one percent of potential GDP, conditional on systematic monetary policy ($Hawk_t$). We show IV and OLS estimates based on the local projection framework (3.5)-(3.6) as specified in Section 4.1. The displayed estimates are computed as $\sum_{h=H-3}^{H}[\beta^h + \gamma^h(Hawk_t - \overline{Hawk_t})]$ for H =4 quarters in Panel (a) and $H = 16$ quarters in Panel (b).

Definition and estimation. The multiplier is defined as the dollar amount by which GDP increases per dollar increase in fiscal spending (both in real terms). A common procedure is to compute the multiplier as the cumulative response of GDP to a spending shock divided by the cumulative response of government spending to the same shock over some horizons of interest (e.g., Mountford and Uhlig, 2009; Ramey and Zubairy, 2018). To study how systematic monetary policy shapes the fiscal multiplier, we define the monetary policy-dependent fiscal multiplier as

$$FM^H(\chi) = \frac{\sum_{h=0}^{H}\left(\beta_{\text{GDP}}^h + \gamma_{\text{GDP}}^h \chi\right)}{\sum_{h=0}^{H}\left(\beta_{\text{G}}^h + \gamma_{\text{G}}^h \chi\right)} \tag{4.1}$$

where H is the forecast horizon, β_i^h and γ_i^h are the average and differential responses of outcome $i \in \{\text{GDP}, \text{G}\}$ to a spending shock, and χ indicates some level of the Hawk-Dove balance in deviation from the sample mean $(Hawk_t - \overline{Hawk})$.[37] We estimate the responses for cumulative GDP and government spending jointly by seemingly unrelated regressions, see Appendix F.2. This allows us to compute standard errors that account for serial correlation and the cross-correlation between the numerator and denominator of (4.1).[38]

---

[37] Alternatively, one could discount future horizons in (4.1). For common discount rates, this will have a minor impact on our estimated fiscal multipliers.

[38] Our baseline inference procedure for the fiscal multiplier uses the Delta method in conjunction with Driscoll-Kraay standard errors. We further provide Anderson-Rubin type confidence sets that are robust to weak instruments and to the denominator of the multiplier being close to zero, see Section 4.4.

Results. Table 2 presents the IV estimates of the fiscal spending multipliers $FM^H(\chi)$ for both a two-year and a four-year horizon. For an average Hawk-Dove balance, $\chi = 0$, the cumulative spending multiplier is 1.3–1.4 for the two horizons and significantly different from zero at the 10% level. Analogous to Figure 4, we consider a range of χ from $-2/12$ to $+2/12$. As the FOMC becomes more dovish than average, the multiplier increases from 1.3 to 2.3 for one additional dove ($\chi = -1/12$), and to 3 for two additional doves ($\chi = -2/12$). The difference between the average and the dovish multipliers is similar across the two horizons. Moreover, the difference is statistically significant at the 5% level for the four-year horizon, see Table E.1 in Appendix E. Conversely, as the FOMC becomes more hawkish, the multiplier $FM^H(\chi)$ drops to zero or below and is insignificantly different from zero. The differences in $FM^H(\chi)$ across χ are mainly driven by differences in the cumulative GDP response rather than the G response. The differences in the GDP response across χ are larger in magnitude and more significant, see Table E.1. This result is analogous to the findings in Figure 4.

Comparison with linear model. We estimate how the fiscal spending multiplier depends on systematic monetary policy, whereas much of the related literature has estimated a single 'average' fiscal spending multiplier (e.g., Blanchard and Perotti, 2002; Ramey, 2016). To compare our results with this tradition in the literature, we estimate an average fiscal spending multiplier in a linear version of our framework when restricting $\gamma^h = \delta^h = 0$. The resulting fiscal multiplier is given by $\widetilde{FM}^H = (\sum_{h=0}^{H} \beta^h_{\text{GDP}})/(\sum_{h=0}^{H} \beta^h_{\text{G}})$ and the estimates are presented in the last column of Table 2. We find average multipliers of about 0.85 at both horizons. While this estimate is relatively close to the multiplier estimates in Ramey and Zubairy (2018), which range from 0.66 to 0.71 (see their Table 1), it is substantially below the multiplier of 1.3 for an average FOMC composition ($FM^H(0)$) in our baseline model. In addition, the standard errors for the multiplier in the linear model are substantially larger than the standard errors of $FM^H(0)$. This comparison suggests that accounting for systematic monetary policy is important for the magnitude and precision of multiplier estimates.[39]

4.4 Weak instruments and robust inference

A common concern with IV estimates is the strength of the instrument. We provide evidence supporting the strength of our instruments, including weak instrument tests, reinforcing the contribution of our identification design. Finally, we provide robust inference for the estimated responses and fiscal multipliers.

---

[39] One can formally show that the estimator in the linear local projection has larger asymptotic variance than the corresponding average estimator in the state-dependent local projection.

Table 2: Government spending multipliers and monetary policy

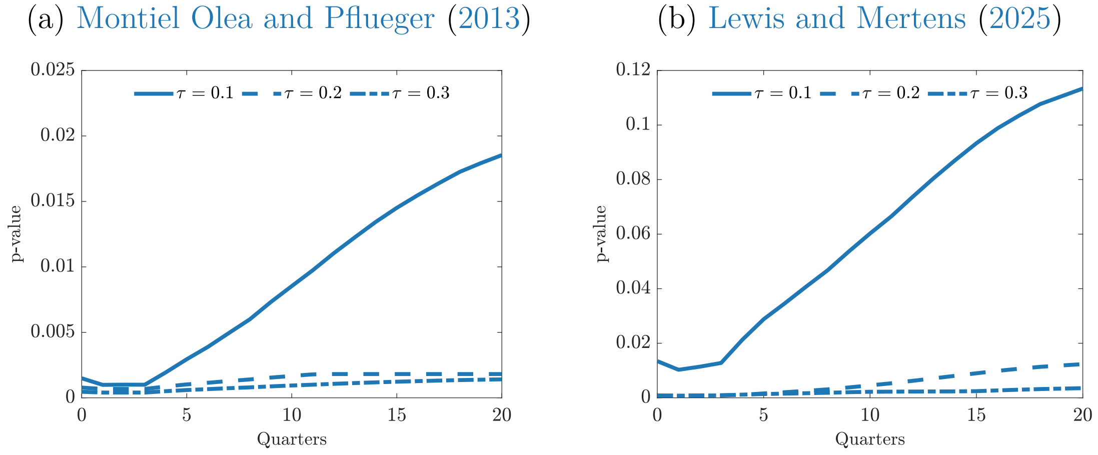
	Baseline model					Linear
Outcome	+2 Hawks	+1 Hawk	Average	+1 Dove	+2 Doves	model
			Two-year horizon			
Multiplier	-4.700	-0.448	1.358	2.357	2.991	0.860
	(5.011)	(1.396)	(0.702)	(0.921)	(1.225)	(1.427)
GDP (*cum*)	-1.682	-0.269	1.144	2.557	3.970	0.616
	(0.990)	(0.771)	(0.653)	(0.690)	(0.861)	(1.057)
G (*cum*)	0.358	0.600	0.843	1.085	1.327	0.716
	(0.249)	(0.300)	(0.395)	(0.510)	(0.635)	(0.338)
			Four-year horizon			
Multiplier	-1.723	0.030	1.316	2.300	3.077	0.838
	(2.515)	(0.830)	(0.466)	(0.790)	(1.135)	(1.449)
GDP (*cum*)	-2.693	0.056	2.804	5.552	8.301	1.494
	(2.472)	(1.547)	(0.845)	(1.039)	(1.868)	(2.747)
G (*cum*)	1.562	1.846	2.130	2.414	2.698	1.782
	(1.006)	(0.806)	(0.745)	(0.852)	(1.079)	(0.689)

Notes: The table shows IV estimates of the cumulative fiscal spending multipliers $FM^H(\chi)$ in equation (4.1) for $H = 8$ (top panel) and $H = 16$ quarters (bottom panel), as well as the cumulative GDP response (numerator of $FM^H(\chi)$) and the cumulative G response (denominator of $FM^H(\chi)$). The coefficients are estimated using a cumulative version of the local projection framework (3.5)-(3.6) as specified in Section 4.1. For our baseline model, the columns present different states of the Hawk-Dove balance between "+2 Hawks" ($\chi = +2/12$), "Average" ($\chi = 0$), and "+2 Doves" ($\chi = -2/12$). The linear model in the last column presents the estimates when we restrict $\gamma^h = \delta^h = 0$ in the local projection (3.5). Driscoll-Kraay standard errors are in parenthesis, see Appendix **F** for details.

First-stage results. Our local projection framework (3.5) contains two endogenous regressors, $\varepsilon_t^s(Hawk_t - \overline{Hawk})$ and $(Hawk_t - \overline{Hawk})$. In the two associated first-stage regressions, we find that the instrumental variable $\varepsilon_t^s(Hawk_t^{\mathcal{R}} - \overline{Hawk}^{\mathcal{R}})$ has a positive effect on the endogenous variable $\varepsilon_t^s(Hawk_t - \overline{Hawk})$ that is significant at the one percent level. Similarly, $(Hawk_t^{\mathcal{R}} - \overline{Hawk}^{\mathcal{R}})$ has a positive and highly significant effect on $(Hawk_t - \overline{Hawk})$. In both regressions, the R^2 increases by about 0.4 when including the instruments as regressors. Taken together, these results suggest that our instruments are strong (Bound et al., 1995).

Weak instrument tests. We use three statistical tests to assess the strength of our instrument more formally. First, we use the Montiel Olea and Pflueger (2013) test of weak instruments, which is popular in time series settings because it is robust to autocorrelation

Figure 6: Weak instrument tests


(a) Montiel Olea and Pflueger (2013)

(b) Lewis and Mertens (2025)

Notes: The figure shows p-values for rejecting the null of weak instruments for the responses of real GDP, based on the local projection framework (3.5)-(3.6) as specified in Section 4.1. The Montiel Olea and Pflueger (2013) test evaluates the null of the bias in γ^h exceeding a threshold τ. Similarly, the Lewis and Mertens (2025) test evaluates the null of the ℓ^2 norm of the bias in γ^h and δ^h exceeding a threshold τ. For the former, the endogenous regressor $Hawk_t$ is not tested but directly replaced by its first stage fitted value. The critical values and associated p-values are based on Newey-West standard errors.

and heteroskedasticity. Formally, we test whether the relative weak instrument bias for the IV estimates of γ^h exceeds 10%, 20%, or 30%.[40] Panel (a) of Figure 6 shows the p-values of the weak instrument tests for the differential GDP response. At all horizons, even a relatively small 10% bias ($\tau = 0.1$) can be rejected at significance levels below 2%.

The second weak instrument test we apply was recently developed by Lewis and Mertens (2025) and generalizes Montiel Olea and Pflueger (2013) to allow for multiple endogenous regressors. We apply this test to jointly evaluate whether the average relative bias across γ^h and δ^h exceeds some threshold τ and report the results in Panel (b) of Figure 6. A small average bias of 10% can be rejected at significance levels below 10% for most horizons. Moreover, we can reject a bias of 20% at the two percent level for all horizons. For government spending, both tests lead to similar conclusions.

Lastly, we test for weak instruments via the reduced form of our regression framework. Following Chernozhukov and Hansen (2008), the hypothesis test of the reduced form estimates of γ^h against zero is equivalent to testing whether the instrument has zero relevance. Figure F.1 in the Appendix shows that the reduced-form estimates for γ^h are significant. To summarize, all three tests indicate that our instruments are not weak.

---

[40] We apply the test to γ^h because it is our main coefficient of interest (together with β^h), and because the Montiel Olea and Pflueger (2013) test can only be applied to a single endogenous regressor. For the other endogenous regressor, $(Hawk_t - \overline{Hawk})$ in levels, we estimate the first stage separately and plug in the fitted values in the second stage used to test the interaction term. If we alternatively replace the $Hawk_t$ level term by $Hawk_t^{\mathcal{R}}$ we obtain very similar results.

Robust inference for impulse responses. To address residual concerns about instrument strength, we further provide inference that is robust to weak instruments and allows for multiple endogenous regressors based on Andrews (2018). We find robust confidence sets for the differential GDP and G responses similar to our baseline intervals, see Figure F.2 in the Appendix. This provides additional support for the strength of our instruments.

Robust inference for fiscal multipliers. We provide Anderson and Rubin (1949) type inference for the fiscal multiplier, following Andrews et al. (2019). Importantly, the procedure is based on a test statistic with a limiting distribution that does not depend on the strength of the instruments and that does not depend on the denominator of the fiscal multiplier being non-zero. We provide details of the implementation in Appendix F.2. The robust confidence sets leave our conclusions about fiscal multipliers in Table 2 broadly unchanged, see Figure F.3. In particular, we estimate fiscal multipliers that are significantly different from zero at the 10% level when the FOMC is dovish. The hawkish multipliers are insignificant. Finally, the average multiplier is significantly different from zero at the 5% level, whereas the multiplier in the linear model remains insignificant.

4.5 Sensitivity analysis

In this section, we provide an extensive sensitivity analysis to assess the robustness of our baseline results. We investigate alternative Hawk-Dove balances, an alternative spending shock, varying sample periods, and the inclusion of additional control variables.

Alternative Hawk-Dove balances. We address potential concerns regarding the aggregation of individual policy preferences and the comparability of preferences over time. While our baseline $Hawk_t$ aggregates individual preferences by an unweighted arithmetic average, we consider four alternative aggregation schemes. First, we use the median policy preference across FOMC members. Second, we use an arithmetic average but double the weight of the Fed Chair. Third, we use the arithmetic average but do not distinguish between consistent and swinging FOMC members when defining $Hawk_{i\tau}$ in (3.1). Across the three alternative aggregations, we find responses and multipliers similar to the baseline, see Figure G.1 and Table G.1.[41] In a fourth alternative aggregation, we consider the role of strong majorities in the FOMC. We construct an alternative Hawk-Dove balance which equals -1 if $Hawk_t$ falls below the first quartile or tertile of the distribution of $Hawk_t$ over time, +1 above the highest quartile or tertile, and zero otherwise. Both specifications roughly align with the

---

[41] Our results are further robust to doubling only the weight of Chairs Paul Volcker and Alan Greenspan, both widely considered to be particularly influential chairs.

baseline responses and multipliers, see Figure G.2 and Table G.2.

Another potential concern is that the meaning of being a hawk or dove might have changed over time. To account for trends in the Hawk-Dove balance, we consider an alternative Hawk-Dove balance that subtracts from the baseline $Hawk_t$ its backward-looking 5, 10, or 15-year moving average. We obtain results similar to the baseline, see Figure G.6 and Table G.1.

Alternative FOMC rotation instruments. We further investigate alternative rotation instruments to address potential endogeneity concerns. First, to address concerns due to preference swings, we either allow swings in the instrument only with a time lag of 8 or 16 quarters or impose that preferences equal the average preference of an FRB president, rendering them time-invariant. The implied state-dependence of the responses and fiscal multipliers in Figure G.4 and Table G.1 are slightly muted compared to the baseline.[42] Second, to address the residual predictability of $Hawk_t^{\mathcal{R}}$ as discussed in Section 3.2, we orthogonalize $Hawk_t^{\mathcal{R}}$ with respect to those variables that display significant predictive power, as shown in Figure D.1. Using the purged rotation instrument barely changes our results, see Figure G.5 and Table G.1.

Alternative spending shocks. The Ramey and Zubairy (2018) shock has a high kurtosis, which naturally raises concerns about outlier sensitivity. To address this concern, we consider a shock series that is winsorized to eliminate excess kurtosis for the set of (non-zero) shocks, which requires winsorizing two negative and four positive shocks. Our results are quite robust to using the winsorized shocks, see Figure G.8 and Table G.1.

The Ramey and Zubairy (2018) shock is a specific type of spending shock. We investigate the external validity of our results by using an alternative fiscal spending shock, which is identified from a timing restriction on total government spending as suggested by Blanchard and Perotti (2002), henceforth BP. They assume that only government spending shocks can affect government spending contemporaneously. We find that GDP and G respond more swiftly compared to our baseline, in line with the nature of the BP shock, see Figure G.7. More importantly, we find that a hawkish FOMC significantly dampens the expansionary effect on GDP. The average fiscal multiplier is around 1.3 for the four-year horizon, see Table G.1, which is remarkably similar to our baseline multiplier. The fiscal multiplier ranges from 0.86 to 1.73 between the hawkish and dovish FOMC ($\chi = \pm 2/12$). While the variation in the multiplier is more compressed compared to the baseline, it is similarly

---

[42] The slightly muted state-dependence does not necessarily imply that our results are influenced by endogenous swings. The alternative rotation instrument also takes out variation from swings that are exogenous, see the discussion in Section 3.2.

significant.[43]

Great Recession and ZLB. Our baseline results are estimated using a sample from 1960Q1 to 2014Q4, which includes the Great Recession (GR) and the subsequent ZLB period. We investigate the sensitivity of our results on a sample that ends either in 2007Q4 or in 2008Q4. For both of these subsamples, our responses and multiplier estimates are very similar to the baseline, see Figure G.3 and Table G.1.

Additional (non-linear) control variables. Finally, we investigate the sensitivity of our results to adding potentially important co-variates to the baseline specification of our local projection framework. The additional control variables are short-term and long-term interest rates, inflation, and the primary surplus. While the estimates are similar to the baseline, we naturally give up some statistical power, see Figure G.9 and Table G.1. Nevertheless, we estimate dovish multipliers between 1.6 and 2.0, which substantially exceeds the average multiplier, consistent with our baseline results. We further add lags of $Hawk_t$, or we consider non-linear controls by including interactions of $Hawk_t$ with the control variables. The results are remarkably close to the baseline, see Figure G.10 and Table G.1.

5 Inspecting the mechanism

In this section, we inspect the mechanism behind our findings in Section 4. We show that in response to an expansionary spending shock, nominal and real interest rates rise, and inflation is dampened under a hawkish FOMC. Conversely, interest rates initially fall and rise only with substantial delay under a dovish FOMC. We further discuss our results in comparison with related estimates in the literature, rationalize our evidence in a calibrated New Keynesian model, and provide evidence for other macroeconomic shocks, similarly pointing to an important role of systematic monetary policy.

5.1 Interest rates and inflation

Nominal interest rates. We study the response of the FFR, as well as the annualized yield on 1-year and 10-year Treasury securities, to government spending shocks by using our local projection framework (3.5)-(3.6) with interest rates as outcome variable x_{t+h}. We

---

[43] The compressed variation in the multiplier is consistent with the FFR response to the BP shock. Initially, FFR significantly rises under a more hawkish FOMC, but the magnitude is smaller than for the baseline spending shocks. Starting two years after the shock, the differential FFR response flips sign and FFR is lower under a more hawkish FOMC. The different FFR responses across spending shocks likely reflect the fact that G rises only temporarily after the BP shock, but persistently after the military spending shock.

follow the specification in Section 4.1 but include four lags of the FFR, 1-year and 10-year Treasury yields, and CPI inflation as additional control variables to account for potential pre-trends in these outcomes.

Panels (a), (c), and (e) of Figure 7 show the IV estimates of β^h, the average response of the three nominal interest rates when $Hawk_t$ equals its sample average. The average FFR response appears muted in the first year, after which it gradually increases and reaches 31 basis points at horizons beyond two years. The average responses of the 1-year and 10-year yields feature similar shapes, albeit at lower magnitudes. Panels (b), (d), and (f) show the IV estimates of $\beta^h \pm \gamma^h$, the state-dependent interest rate responses when $Hawk_t$ exceeds the sample average either by two hawks (+2 Hawks) or by two doves (+2 Doves). All interest rates increase faster and more strongly under a hawkish FOMC. Compared to the average response, the peak in the FFR is reached one year earlier and is almost double in size (about 52 basis points). In contrast, under a dovish FOMC, the FFR falls for almost two years, and we find a reversion to a higher FFR only three years after the shock. We find similar differences for 1-year and 10-year Treasury yields, suggesting that the monetary regimes also differ in their effects on expected future policy at long horizons.

Importantly, starting three quarters after the shock, we find a large and statistically significant gap in the FFR response between the hawkish and dovish scenarios. We further find that the differences in the FFR response largely reflect those spending shocks that occur at the beginning of the year.[44] Hence, the (initial) difference in the FFR response may reflect the votes of those rotating FOMC members present at the time of the shock. The evidence is therefore consistent with the view that voting matters, in line with the rationale for our instrument. Moreover, the sluggish FFR response is consistent with the initial uncertainty surrounding the military spending shock and the gradually evolving macroeconomic effects of the shock, see Figure 4. Section 6 provides narrative evidence from the FOMC historical records suggesting that the FOMC actually delays action until some uncertainty about the spending plans and their potential effect on the economy and inflation is resolved.

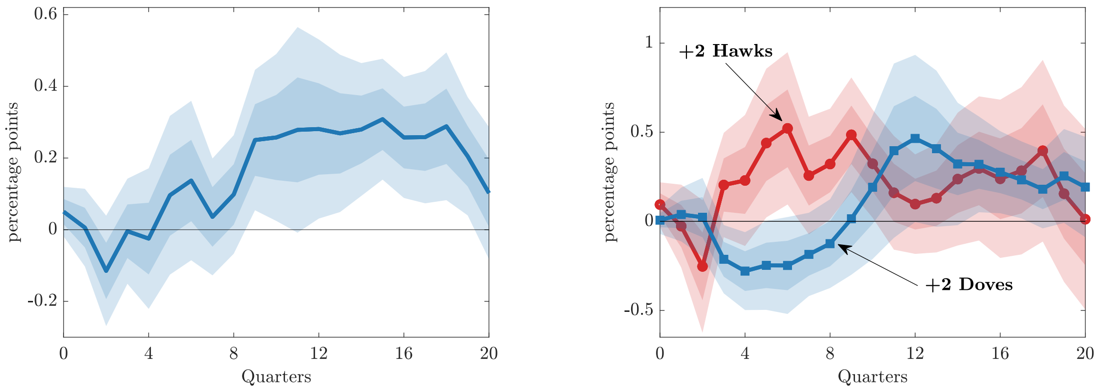
On a more general note, the estimated FFR responses can be used to construct *conventional* policy counterfactuals (as in McKay and Wolf, 2023). In particular, we can implement counterfactual FFR responses through shifts in the Hawk-Dove balance, and construct the associated effects on other outcomes, e.g., the fiscal multiplier, as we do in Section 5.2.

Inflation rates. We further assess the effects of the military spending shocks on inflation expectations, CPI core inflation (excluding food and energy prices), and CPI headline infla-

---

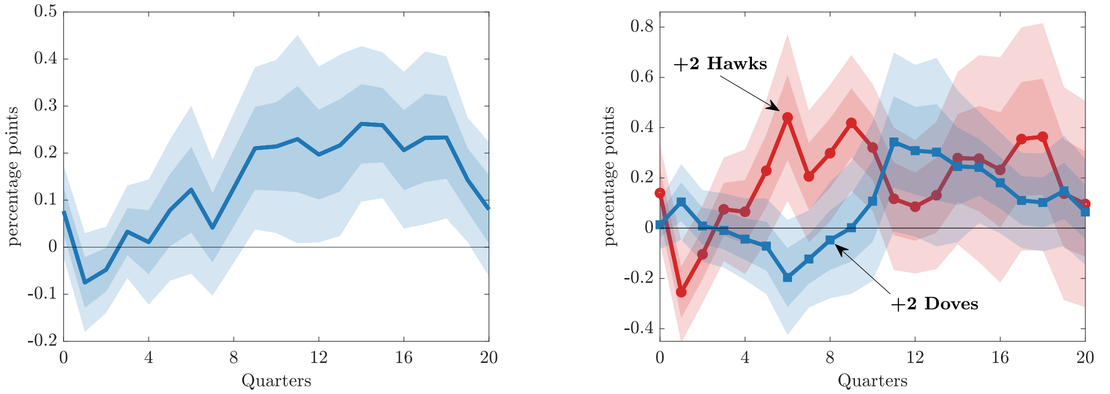
[44] In Figures H.1-H.2 and Table G.1, we only consider spending shocks that occur in the first half of the year and obtain similar results compared to the baseline. Conversely, the dependence on monetary policy becomes weaker and less significant if we only consider spending shocks in the second half of the year.

Figure 7: Responses of nominal interest rates



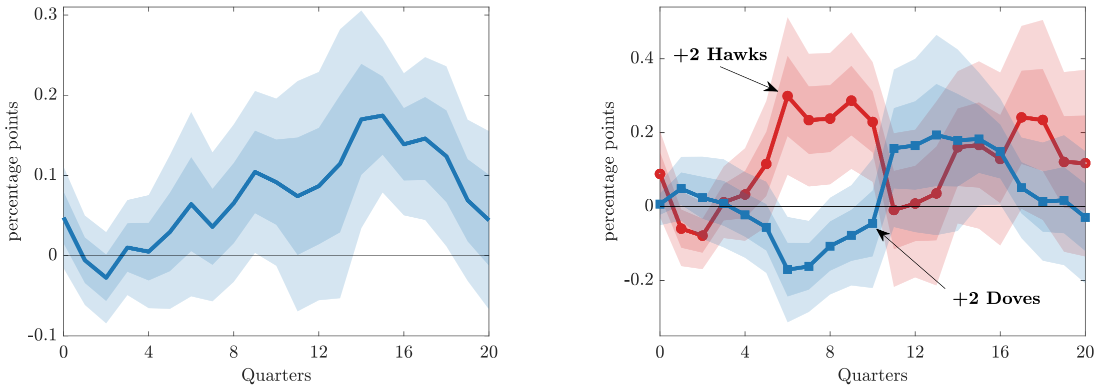
(a) Average FFR (β^h)

(b) State-dependent FFR ($\beta^h \pm \gamma^h$)



(c) Average 1-year rate (β^h)

(d) State-dependent 1-year rate ($\beta^h \pm \gamma^h$)



(e) Average 10-year rate (β^h)

(f) State-dependent 10-year rate ($\beta^h \pm \gamma^h$)

Notes: The figure shows responses of the federal funds rate (FFR), as well as the 1-year and 10-year treasury yields to an expansionary military spending shock, corresponding to one percent of potential GDP, conditional on systematic monetary policy ($Hawk_t$). All outcomes are annualized interest rates. We show IV estimates based on the local projection framework (3.5)-(3.6) as specified in Section 5.1. The β^h captures the responses when $Hawk_t$ equals its sample average. The $\beta^h \pm \gamma^h$ shows the state-dependent responses when $Hawk_t$ exceeds the sample average either by two hawks (+2 Hawks) or by two doves (+2 Doves). The shaded areas indicate 68% and 95% confidence bands using Newey-West standard errors.

tion.[45] We estimate the inflation responses using the specification of our local projection framework (3.5)-(3.6) for nominal interest rates and control for four lags of the inflation measure under consideration. The results are shown in Figure 8. Overall, the inflation responses are not precisely estimated. The average response of expected inflation tends to be positive, while the evidence is mixed for core and headline inflation. Turning to the dependence on the Hawk-Dove balance, we find that inflation expectations increase sluggishly under a dovish FOMC and peak at about three years. In contrast, inflation expectations tend to fall under a hawkish FOMC, suggesting that the FOMC is successful in containing inflation expectations. The response of core inflation follows a similar but even more sluggish pattern, suggesting that policy tightening is successful in containing inflationary pressures. Compared to the interest rate responses, the inflation response appears delayed by one to two years, broadly in line with the lags in the transmission of monetary policy. Finally, the results for headline inflation are more mixed, possibly due to larger transitory fluctuations in energy and food prices.

Real interest rates. In a large class of models, the real effects of monetary policy depend on its ability to affect real interest rates. Under a hawkish FOMC, the response of nominal rates is larger, while the response of inflation is smaller. Hence, the implied response of real interest rates is larger. In response to a government spending shock, real interest rates increase by more if the FOMC is hawkish and by less if the FOMC is dovish. We obtain similar results when directly estimating the response of the real FFR, which we compute by subtracting expected CPI inflation from the nominal rate, see Figure H.3 in Appendix H.

5.2 Discussion of results

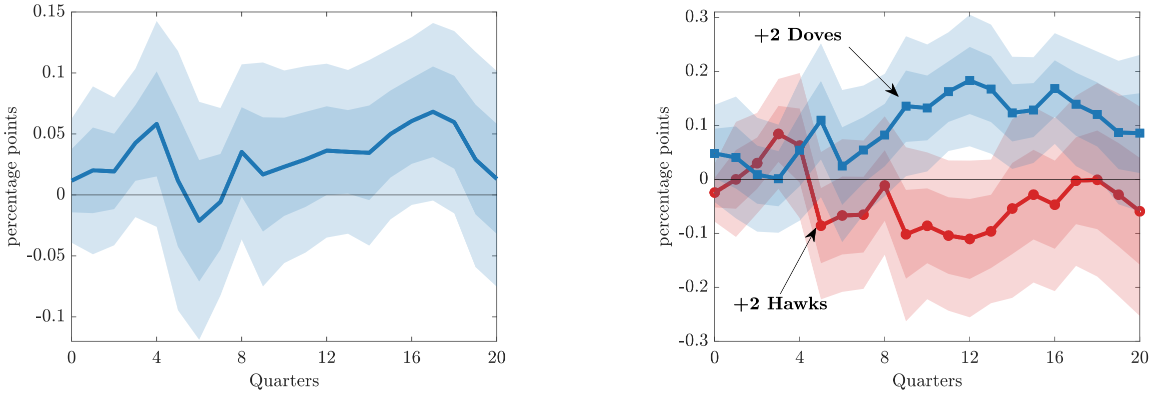
To put our empirical results into perspective, we discuss them in comparison with prior estimates for the effects of monetary policy shocks and fiscal multipliers.

Relation to monetary policy shocks. A large related empirical literature estimates the effects of monetary policy shocks. It is therefore natural to compare the effects of such shocks with our estimates. In particular, we compare the ratio of the peak output to the peak interest rate response for various monetary policy shocks with the ratio of the peak differential GDP response to the peak differential interest rate response, formally $(\min_h \gamma_y^h)/(\max_h \gamma_i^h)$ based on our estimates. For U.S. monetary policy shocks, recursively identified shocks in

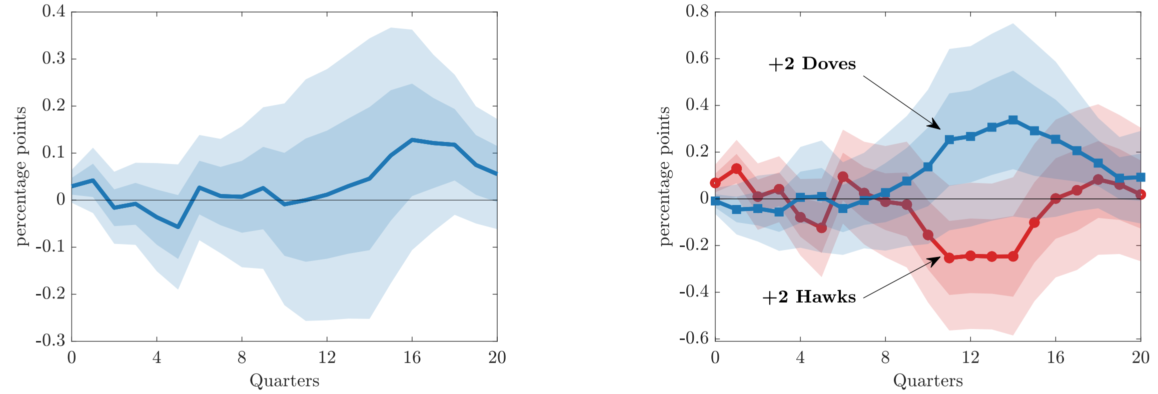
---

[45] We use one-year inflation expectations based on the CPI forecasts from the Livingston Survey of the Federal Reserve Bank of Philadelphia. It is the oldest continuous survey on the expectations of economists from industry, government, banking, and academia. For details, see Appendix B.

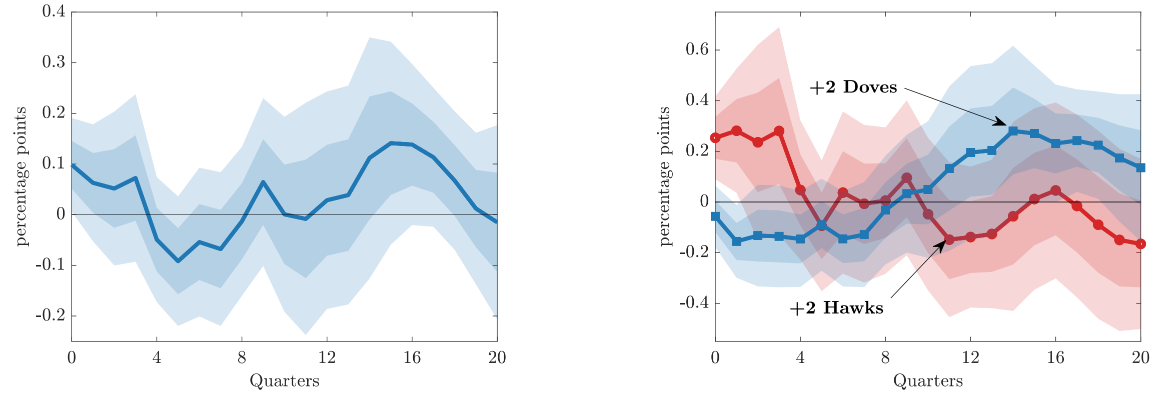
Figure 8: Responses of inflation rates


(a) Average expected inflation (β^h)

(b) State-dependent expected inflation ($\beta^H \pm \gamma^h$)


(c) Average core inflation (β^h)

(d) State-dependent core inflation ($\beta^H \pm \gamma^h$)


(e) Average headline inflation (β^h)

(f) State-dependent headline inflation ($\beta^H \pm \gamma^h$)

Notes: The figure shows responses of expected inflation, CPI core, and CPI headline inflation to an expansionary military spending shock, corresponding to one percent of potential GDP, conditional on systematic monetary policy ($Hawk_t$). All outcomes are annualized inflation rates. We show IV estimates based on the local projection framework (3.5)-(3.6) as specified in Section 5.1. The β^h captures the responses when $Hawk_t$ equals its sample average. The $\beta^h \pm \gamma^h$ shows the state-dependent responses when $Hawk_t$ exceeds the sample average either by two hawks (+2 Hawks) or by two doves (+2 Doves). The shaded areas indicate 68% and 95% confidence bands using Newey-West standard errors.

Coibion (2012) imply a ratio of −1.56, Romer and Romer (2004) shocks as estimated in Coibion (2012) a ratio of −1.99, and high-frequency identified shocks in Jarociński and Karadi (2020) a ratio of −1.34. In comparison, our estimates imply a ratio of −1.35, which is within the range implied by monetary policy shocks.

Relation to fiscal multipliers. The interest rate responses further allow us to compare our fiscal spending multiplier estimates with the findings in the related literature. Our spending multiplier is between two and three under a dovish FOMC, associated with a weakly negative response of the nominal (and real) FFR for the first two years. In theory, the multiplier may be far above one (or negative) depending on the response of interest rates (Woodford, 2011; Farhi and Werning, 2016). In an estimated medium-scale DSGE model, Christiano et al. (2011) find multipliers between two and four at the ZLB, when the short-run nominal interest rate does not respond, broadly similar to our estimates.

Our findings also relate to an empirical literature that estimates fiscal spending multipliers. For example, Nakamura and Steinsson (2014) estimates two-year regional multipliers for the U.S. of approximately 1.5. To the extent that regional multipliers correspond to the aggregate multiplier when nominal interest rates do not respond, we can compare their estimates with our two-year multipliers. In particular, we construct a spending multiplier for the case in which the nominal FFR is unresponsive by choosing the Hawk-Dove balance (χ) that minimizes the squared distance of the FFR response from zero in the first two years.[46] This monetary policy counterfactual requires a χ between 0 and "+1 Dove" in Table 2. The associated two-year spending multiplier is 1.9, similar to the estimates in Nakamura and Steinsson (2014).

We further compare our results with the estimate of the aggregate spending multiplier when monetary policy is constrained at the ZLB. Ramey and Zubairy (2018) finds a ZLB multiplier of 1.6 after two years (when excluding WWII), while Miyamoto et al. (2018) find a ZLB multiplier well above 1.5 for Japan. Notwithstanding the endogeneity of a binding ZLB, our multiplier of 1.9 under a non-responsive FFR is similar to the ZLB multipliers in the literature. Overall, our multiplier estimates and the associated interest rate path are broadly similar to previous quantitative and empirical findings.

5.3 Calibrated New Keynesian model

Finally, we rationalize our empirical findings using a calibrated New Keynesian model, which provides a structural interpretation of the mechanisms underlying our estimates. We

---

[46] Formally, we solve $\min_\chi \sum_{h=0}^{8}(\beta^h_{FFR} + \chi \cdot \gamma^h_{FFR})^2$, where χ indicates a level of the Hawk-Dove balance in deviation from the sample mean $(Hawk_t - \overline{Hawk})$.

consider a behavioral two-agent New Keynesian framework that extends the stylized model in Section 2 by introducing interest rate smoothing, hand-to-mouth agents, and cognitive discounting. While the simpler model remains nested as a special case, these features allow the model to generate persistent dynamics, fiscal multipliers above one with realistic MPCs, and a relaxation of the Taylor principle (Pfäuti and Seyrich, 2026).

Calibrated to match our empirical responses of government spending, output, inflation, and interest rates, the model reproduces the key patterns observed in the data. In particular, it explains how short-lived differences in systematic monetary policy – arising from the FOMC rotation – cause persistent differences in the responses of output, inflation, and interest rates, as illustrated by the "+2 Hawks" and "+2 Doves" scenarios. A key mechanism that transforms short-lived differences in systematic monetary policy into persistent effects is interest rate smoothing. Details on the model and the quantitative analysis are provided in Appendix I.

5.4 Evidence beyond government spending shocks

Theory predicts that the propagation of macroeconomic shocks depends on systematic monetary policy. We have so far tested this prediction using two government spending shocks: the military spending shock of Ramey and Zubairy (2018) and a spending shock in the tradition of Blanchard and Perotti (2002). We further apply our identification design to non-spending shocks and find that their propagation is meaningfully shaped by systematic monetary policy. Appendix J provides details, with the main results summarized below.

We consider two shocks: a business-cycle (demand) shock based on Angeletos et al. (2020) and a productivity shock given by the TFP residual of Fernald (2014). For both shocks, we find that a hawkish FOMC raises interest rates swiftly whereas a dovish FOMC is less responsive or even accommodative. The differential interest rate response is associated with considerable differences in real activity and inflation, although the inflation effects are less precisely estimated. The findings align well with the responses to spending shocks and thus demonstrate the broad applicability of our identification design for applied macroeconomic research and the empirical relevance of systematic monetary policy.

6 Historical FOMC records

> *Interviewer*: What would have happened, do you think, if the Fed had not raised the discount rate?
>
> *Chairman Martin*: A golden opportunity to stop inflation in its tracks would

> have been lost.
>
> *Interviewer*: It was primarily the projection of Vietnam spending; is that correct?
>
> *Chairman Martin*: Right. I kept telling him we could not have guns and butter.
>
> *Interviewer*: When you talked to Lyndon Johnson about this projection, what did he say? Did he disagree with it or did he agree with it?
>
> *Chairman Martin*: He disagreed. He thought we could have guns and butter.[47]

We complement our quantitative analysis with narrative evidence from the records of discussions and decisions at FOMC meetings. This evidence serves two purposes. First, it confirms that the FOMC members discuss changes in government defense spending, assessing the impact on economic activity and inflation as well as the FOMC's policy response. Second, it shows that the policy response depends on the composition of the FOMC.

To illustrate the FOMC discussion around military spending shocks, the FOMC composition, and the corresponding policy response, we focus on two important events during the 1960s: the acceleration of the U.S. Space Program in 1961 and the Vietnam ground war starting in 1965. The corresponding military shocks are both large while the FOMC composition is hawkish in the first part of the 1960s and dovish in the second part, see Figure 2. In this period, the Fed was headed by William McChesney Martin, a consistent hawk whose tenure as chairman from 1951 to 1970 was the longest in history. For both events, we identify three phases of the FOMC's reaction to military defense spending from the historical FOMC records. First, there is uncertainty about the extent to which the spending plans will be realized and about their impact on the economy. Second, the effects of higher spending on the economy become visible while inflation appears unresponsive, they wait until "all the evidence was in". Third, the effects on inflation become visible but the FOMC delays action. The first two phases are common to both hawkish and dovish committees; the third – delayed action despite rising inflation – is more pronounced under a dovish FOMC, broadly in line with our empirical findings.

We provide these case studies in Appendix K. The sources for our narrative evidence are the FOMC Historical Minutes until 1967 and the Memoranda of Discussion thereafter.

7 Conclusion

This paper proposes an identification design to estimate the effects of systematic monetary policy on the propagation of macroeconomic shocks. Our design combines the narrative

---

[47] Former Fed Chairman William McChesney Martin: Oral History, Interview I by Michael L. Gillette in 1987, LBJ Library Oral History Collection. The interviewer refers to the Fed's decision to raise the discount rate on December 1965. Lyndon B. Johnson was the U.S. President from 1963 to 1969.

classification of FOMC members' policy preferences from Istrefi (2019) with a novel FOMC rotation instrument for systematic monetary policy. The identification design opens up myriad research opportunities, such as revisiting the effects of various fiscal, technology, and oil shocks and their dependence on systematic monetary policy.

We use our identification design to study government spending shocks in the U.S. and find that fiscal spending multipliers depend strongly and significantly on systematic monetary policy. We inspect the mechanism behind our result and find consistent interest rate and inflation responses. In recent years, we have observed large fiscal expansions related to COVID-19 and, more recently, related to Russia's war against Ukraine. In the same period, the FOMC was rather dovish. Applied to these years, our findings suggest that the combination of fiscal and monetary policy contributed to the robust recovery of GDP.

However, a potentially misleading conclusion from our results is that the government should increase spending when the FOMC is dovish. This could be misleading because such responses of government spending to systematic monetary policy are not random shocks. This is a case of the Lucas (1976) critique. To avoid misleading conclusions, a promising avenue for future research is to use our results to discipline micro-founded models to study optimal fiscal stabilization policy.

References

Abrams, B. (2006): "How Richard Nixon Pressured Arthur Burns: Evidence from the Nixon Tapes," *Journal of Economic Perspectives*, 20, 177–188.

Anderson, T. W. and H. Rubin (1949): "Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations," *Annals of Mathematical Statistics*, 20, 46–63.

Andrews, I. (2018): "Valid Two-Step Identification-Robust Confidence Sets for GMM," *Review of Economics and Statistics*, 100, 337–348.

Andrews, I., J. H. Stock, and L. Sun (2019): "Weak Instruments in Instrumental Variables Regression: Theory and Practice," *Annual Review of Economics*, 11, 727–753.

Angeletos, G.-M., F. Collard, and H. Dellas (2020): "Business-Cycle Anatomy," *American Economic Review*, 110, 3030–3070.

Auerbach, A. and Y. Gorodnichenko (2012): "Measuring the Output Responses to Fiscal Policy," *American Economic Journal: Economic Policy*, 4, 1–27.

Barnichon, R., D. Debortoli, and C. Matthes (2022): "Understanding the Size of the Government Spending Multiplier: It's in the Sign," *Review of Economic Studies*, 89, 87–117.

Barnichon, R. and G. Mesters (2023): "A Sufficient Statistics Approach for Macro Policy," *American Economic Review*, 113, 2809–2845.

Barsky, R. and E. Sims (2011): "News Shocks and Business Cycles," *Journal of Monetary Economics*, 58, 273–289.

Barthélemy, J. and M. Marx (2017): "Solving Endogenous Regime-Switching Models," *Journal of Economic Dynamics and Control*, 77, 1–25.

Bauer, M. D., C. E. Pflueger, and A. Sunderam (2024): "Perceptions about Monetary Policy," *Quarterly Journal of Economics*, 139.

Bauer, M. D. and E. T. Swanson (2023): "An Alternative Explanation for the "Fed Information Effect"," *American Economic Review*, 113, 664–700.

Belden, S. (1989): "Policy Preferences of FOMC Members as Revealed by Dissenting Votes," *Journal of Money, Credit and Banking*, 21, 432–441.

Bernanke, B. S. and A. S. Blinder (1992): "The Federal Funds Rate and the Channels of Monetary Transmission," *American Economic Review*, 82, 901–921.

Bernanke, B. S., M. Gertler, and M. Watson (1997): "Systematic Monetary Policy and the Effects of Oil Price Shocks," *Brookings Papers on Economic Activity*, 28, 91–157.

Bianchi, F. (2013): "Regime Switches, Agents' Beliefs, and post-World War II U.S. Macroeconomic Dynamics," *Review of Economic Studies*, 80, 463–490.

Bianchi, F., R. Gómez-Cram, T. Kind, and H. Kung (2023): "Threats to Central Bank Independence: High-Frequency Identification with Twitter," *Journal of Monetary Economics*, 135, 37–54.

Bianchi, F. and C. Ilut (2017): "Monetary/Fiscal Policy Mix and Agents' Beliefs," *Review of Economic Dynamics*, 26, 113–139.

Blanchard, O. and R. Perotti (2002): "An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output," *Quarterly Journal of Economics*, 117, 1329–1368.

Blinder, A. S. (1999): *Central Banking in Theory and Practice*, vol. 1 of *MIT Press Books*, The MIT Press.

——— (2007): "Monetary Policy by Committee: Why and How?" *European Journal of Political Economy*, 23, 106–123.

——— (2022): *A Monetary and Fiscal History of the United States, 1961-2021*, Princeton University Press.

Bordo, M. D. (2016): "Some Historical Reflections on the Governance of the Federal Reserve," *Central Bank Governance and Oversight Reform*, 221–238.

Bordo, M. D. and K. Istrefi (2023): "Perceived FOMC: The Making of Hawks, Doves and Swingers," *Journal of Monetary Economics*, 136, 125–143.

Bound, J., D. A. Jaeger, and R. M. Baker (1995): "Problems with Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogenous Explanatory Variable is Weak," *Journal of the American Statistical Association*, 90, 443–450.

Broner, F., D. Clancy, A. Erce, and A. Martin (2022): "Fiscal Multipliers and Foreign Holdings of Public Debt," *Review of Economic Studies*, 89, 1155–1204.

Camous, A. and D. Matveev (2021): "Furor over the Fed: A President's Tweets and Central Bank Independence," *CESifo Economic Studies*, 67, 106–127.

Carvalho, C., F. Nechio, and T. Tristão (2021): "Taylor Rule Estimation by OLS," *Journal of Monetary Economics*, 124, 140–154.

Chang, Y., J. Maih, and F. Tan (2021): "Origins of Monetary Policy Shifts: A New Approach to Regime Switching in DSGE Models," *Journal of Economic Dynamics and Control*, 133, 104235.

Chernozhukov, V. and C. Hansen (2008): "The Reduced Form: A Simple Approach to Inference with Weak Instruments," *Economics Letters*, 100, 68–71.

Christiano, L., M. Eichenbaum, and S. Rebelo (2011): "When is the Government Spending Multiplier Large?" *Journal of Political Economy*, 119, 78 – 121.

Cieslak, A., S. Hansen, M. McMahon, and S. Xiao (2023): "Policymakers' Uncertainty," NBER Working Paper 31849, National Bureau of Economic Research.

Clarida, R., J. Gali, and M. Gertler (2000): "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory," *Quarterly Journal of Economics*, 115, 147–180.

Cloyne, J. S., Ò. Jordà, and A. M. Taylor (2023): "Decomposing the Fiscal Multiplier," NBER Working Paper 26939.

Cochrane, J. H. (2011): "Determinacy and Identification with Taylor Rules," *Journal of Political Economy*, 119, 565–615.

Cochrane, J. H. and M. Piazzesi (2002): "The Fed and Interest Rates - A High-Frequency Identification," *American Economic Review*, 92, 90–95.

Coibion, O. (2012): "Are the Effects of Monetary Policy Shocks Big or Small?" *American Economic Journal: Macroeconomics*, 4, 1–32.

Coibion, O. and Y. Gorodnichenko (2011): "Monetary Policy, Trend Inflation, and the Great Moderation: An Alternative Interpretation," *American Economic Review*, 101, 341–70.

Davig, T. and E. M. Leeper (2007): "Generalizing the Taylor Principle," *American Economic Review*, 97, 607–635.

——— (2008): "Endogenous Monetary Policy Regime Change," in *NBER International Seminar on Macroeconomics 2006*, National Bureau of Economic Research, NBER Chapters, 345–391.

DeLong, B. J. (1997): "America's Peacetime Inflation: The 1970s." in *Reducing Inflation: Motivation and Strategy*, ed. by C. Romer and D. Romer, University of Chicago Press.

Ehrmann, M., R. Tietz, and B. Visser (2022): "Voting Right Rotation, Behavior of Committee Members and Financial Market Reactions: Evidence from the U.S. Federal Open Market Committee," IMF Working Papers 2022/105, International Monetary Fund.

Farhi, E. and I. Werning (2016): "Fiscal Multipliers: Liquidity Traps and Currency Unions," in *Handbook of Macroeconomics*, Elsevier, vol. 2, 2417–2492.

Fed, S. L. (2013): "100 Years of Service, Annual Report 2013," Tech. rep.

Ferguson, N., M. Kornejew, P. Schmelzing, and M. Schularick (2023): "The Safety Net: Central Bank Balance Sheets and Financial Crisis, 1987-2020," Economics working paper 23102, Hoover Institution.

Fernald, J. G. (2014): "A Quarterly, Utilization-Adjusted Series on Total Factor Productivity," Working Paper 2012-19, Federal Reserve Bank of San Francisco.

Fernández-Villaverde, J., P. Guerrón-Quintana, and J. F. Rubio-Ramirez (2010): "Reading the Recent Monetary History of the United States, 1959-2007," *Review*, 92, 311–338.

Ferriere, A. and G. Navarro (2025): "The Heterogeneous Effects of Government Spending : It's All About Taxes," *Review of Economic Studies*, 92, 1061–1125.

Fos, V. and N. R. Xu (2025): "When Do FOMC Voting Rights Affect Monetary Policy?" Working Paper 33762, National Bureau of Economic Research.

Galí, J. (2015): *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework and Its Applications*, Princeton University Press.

Galí, J., J. López-Salido, and J. Vallés (2003): "Technology Shocks and Monetary Policy: Assessing the Fed's Performance," *Journal of Monetary Economics*, 50, 723–743.

Ghassibe, M. and F. Zanetti (2022): "State Dependence of Fiscal Multipliers: The Source of Fluctuations Matters," *Journal of Monetary Economics*, 132, 1–23.

Goodfriend, M. and R. G. King (2005): "The Incredible Volcker Disinflation," *Journal of Monetary Economics*, 52, 981–1015.

Hack, L., K. Istrefi, and M. Meier (2024): "The Systematic Origins of Monetary Policy Shocks," CEPR Discussion Paper 19063.

——— (2026): "A Local Projection Representation of State-Dependencies in DSGE Models," *A companion note to "Identification of Systematic Monetary Policy"*.

Ilzetzki, E., E. G. Mendoza, and C. A. Végh (2013): "How Big (Small?) Are Fiscal Multipliers?" *Journal of Monetary Economics*, 60, 239–254.

Istrefi, K. (2019): "In Fed Watchers' Eyes: Hawks, Doves and Monetary Policy," Banque de France Working Paper.

Jarociński, M. and P. Karadi (2020): "Deconstructing Monetary Policy Surprises — The Role of Information Shocks," *American Economic Journal: Macroeconomics*, 12, 1–43.

Jordà, Ò. and A. M. Taylor (2016): "The Time for Austerity: Estimating the Average Treatment Effect of Fiscal Policy," *The Economic Journal*, 126, 219–255.

Kilian, L. and L. T. Lewis (2011): "Does the Fed Respond to Oil Price Shocks?" *Economic Journal*, 121, 1047–1072.

Lazarus, E., D. J. Lewis, J. H. Stock, and M. W. Watson (2018): "HAR Inference: Recommendations for Practice," *Journal of Business & Economic Statistics*, 36, 541–559.

Lewis, D. J. and K. Mertens (2025): "A Robust Test for Weak Instruments for 2SLS with Multiple Endogenous Regressors," *Review of Economic Studies*, rdaf103.

Lucas, R. E. (1976): "Econometric Policy Evaluation: A Critique," *Carnegie-Rochester Conference Series on Public Policy*, 1, 19–46.

Madeira, C. and J. Madeira (2019): "The Effect of FOMC Votes on Financial Markets," *The Review of Economics and Statistics*, 101, 921–932.

McKay, A. and C. K. Wolf (2023): "What Can Time-Series Regressions Tell Us About Policy Counterfactuals?" *Econometrica*, 91, 1695–1725.

Meltzer, A. H. (2005): "Origins of the Great Inflation," *Federal Reserve Bank of St. Louis Review*, March/April, 145–176.

Miyamoto, W., T. L. Nguyen, and D. Sergeyev (2018): "Government Spending Multipliers Under the Zero Lower Bound: Evidence from Japan," *American Economic Journal: Macroeconomics*, 10, 247–77.

Montiel Olea, J. L. and C. Pflueger (2013): "A Robust Test for Weak Instruments," *Journal of Business & Economic Statistics*, 31, 358–369.

Mountford, A. and H. Uhlig (2009): "What Are the Effects of Fiscal Policy Shocks?" *Journal of Applied Econometrics*, 24, 960–992.

Nakamura, E. and J. Steinsson (2014): "Fiscal Stimulus in a Monetary Union: Evidence from US Regions," *American Economic Review*, 104, 753–92.

Orphanides, A. (2003): "The Quest for Prosperity without Inflation," *Journal of Monetary Economics*, 50, 633–663.

——— (2004): "Monetary Policy Rules, Macroeconomic Stability, and Inflation: A View from the Trenches," *Journal of Money, Credit and Banking*, 36, 151–175.

Pfäuti, O. and F. Seyrich (2026): "A Behavioral Heterogeneous Agent New Keynesian Model," Working Paper 2026-01, EMPCT Working Paper Series.

Primiceri, G. (2006): "Why Inflation Rose and Fell: Policymakers' Beliefs and US Postwar Stabilization Policy," *Quarterly Journal of Economics*, 867–901.

Primiceri, G. E. (2005): "Time Varying Structural Vector Autoregressions and Monetary Policy," *Review of Economic Studies*, 72, 821–852.

Ramey, V. (2016): "Macroeconomic Shocks and Their Propagation," *Handbook of Macroeconomics*, 2, 71 – 162.

Ramey, V. A. (2011): "Identifying Government Spending Shocks: It's all in the Timing," *Quarterly Journal of Economics*, 126, 1–50.

Ramey, V. A. and S. Zubairy (2018): "Government Spending Multipliers in Good Times and in Bad: Evidence from US Historical Data," *Journal of Political Economy*, 126, 850–901.

Riboni, A. and F. J. Ruge-Murcia (2010): "Monetary Policy by Committee: Consensus, Chairman Dominance, or Simple Majority?" *Quarterly Journal of Economics*, 125, 363–416.

——— (2022): "The power of the Federal Reserve Chair," *International Economic Review.*

Romer, C. and D. Romer (1989): "Does Monetary Policy Matter? A New Test in the Spirit of Friedman and Schwartz," in *NBER Macroeconomics Annual 1989, Volume 4*, NBER, Inc, 121–184.

Romer, C. D. and D. H. Romer (2004): "A New Measure of Monetary Shocks: Derivation and Implications," *American Economic Review*, 94, 1055–1084.

——— (2010): "The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks," *American Economic Review*, 100, 763–801.

Sargent, T. J. (1979): "Estimating Vector Autoregressions Using Methods not Based on Explicit Economic Theories," *Quarterly Review*, 3.

Taylor, J. B. (1999): "A Historical Analysis of Monetary Policy Rules," in *Monetary Policy Rules*, National Bureau of Economic Research, Inc, NBER Chapters, 319–348.

Wolf, C. K. (2023): "Fiscal Stimulus and the Systematic Response of Monetary Policy," *AEA Papers and Proceedings*, 113, 388–393.

Woodford, M. (2011): "Simple Analytics of the Government Expenditure Multiplier," *American Economic Journal: Macroeconomics*, 3, 1–35.

Identification of Systematic Monetary Policy

Online Appendix

Lukas Hack   Klodiana Istrefi   Matthias Meier

March, 2026

Appendix A New Keynesian model

A.1 Equilibrium dynamics

In the following, we derive equation (2.3). Denoting by lower case letters (log) deviations from steady state, we obtain three equilibrium conditions for the model described in Section 2:

$$\pi_t = \beta \mathbb{E}_t\left[\pi_{t+1}\right] + \lambda\left(\varphi + \frac{1}{1-\gamma}\right) y_t - \frac{\lambda\gamma}{1-\gamma} x_t^s - \lambda(1+\varphi) x_t^a, \tag{A.1}$$

$$y_t = \mathbb{E}_t\left[y_{t+1}\right] - (1-\gamma)(i_t - \mathbb{E}_t\left[\pi_{t+1}\right]) + \gamma\left(1-\rho_s\right) x_t^s, \tag{A.2}$$

$$i_t = \widetilde{\phi}_t \pi_t, \tag{A.3}$$

where $\lambda = (1-\theta)(1-\beta\theta)/\theta$ and where $\widetilde{\phi}_t = \phi + \phi_t$ follows

$$\phi_t = \rho_\phi \phi_{t-1} + \zeta^s \varepsilon_t^s + \zeta^a \varepsilon_t^a + \eta_t, \quad |\rho_\phi| < 1.$$

We assume the macroeconomic shocks $(\varepsilon_t^a, \varepsilon_t^s)$ and the exogenous shifter η_t are mutually independent and identically distributed over time. We combine the equations to obtain

$$y_t = \frac{1-\gamma}{1+\lambda\left(\varphi(1-\gamma)+1\right)\tilde{\phi}_t}\Bigg[\frac{\mathbb{E}_t\left[y_{t+1}\right]}{1-\gamma} + (1-\beta\tilde{\phi}_t)\mathbb{E}_t\left[\pi_{t+1}\right] + \frac{\gamma}{1-\gamma}\left(\tilde{\phi}_t\lambda + (1-\rho_s)\right) x_t^s + \tilde{\phi}_t\lambda\left(\varphi+1\right) x_t^a\Bigg]. \tag{A.4}$$

Combining (A.1) and (A.4), the model dynamics follow $\mathcal{Y}_t = A(\phi_t)\ \mathbb{E}_t\left[\mathcal{Y}_{t+1}\right] + B(\phi_t)\ \mathcal{X}_t$, with $\mathcal{Y}_t = (y_t, \pi_t)'$, $\mathcal{X}_t = (x_t^s, x_t^a)'$ and $A(\phi_t), B(\phi_t)$ depending only on model parameters. A first-order approximation around $\phi_t = 0$ yields

$$\mathcal{Y}_t = A\mathbb{E}_t\left[\mathcal{Y}_{t+1}\right] + B\mathcal{X}_t + \Big(\partial_{\phi_t} A\mathbb{E}_t\left[\mathcal{Y}_{t+1}\right] + A\mathbb{E}_t\left[\partial_{\phi_t}\mathcal{Y}_{t+1}\right] + \partial_{\phi_t} B\mathcal{X}_t\Big)\phi_t, \tag{A.5}$$

where $A \equiv A(0), B \equiv B(0)$, $\partial_{\phi_t}(\cdot)$ denotes a derivative with respect to ϕ_t that is evaluated at $\phi_t = 0$. We next guess the solution to (A.5) satisfies

$$\mathcal{Y}_t = \mathcal{A} + \mathcal{B}\mathcal{X}_t + \mathcal{C}\mathcal{X}_t\phi_t + \mathcal{D}\phi_t, \tag{A.6}$$

which is straightforward to verify. The coefficients of the guess depend on the deep structural parameters of the model and can be determined via the method of undetermined coefficients. This fully describes the approximate state-dependent model dynamics with respect to systematic monetary policy ϕ_t and provides equation (2.3) in the main text, where $a = \mathcal{A}_1$, $b_s = \mathcal{B}_{11}$, $b_a = \mathcal{B}_{12}$, and analogously for $\mathcal{C}$ and $\mathcal{D}$. In the special case $\rho_s = \rho_a = \rho_\phi = 0$, the coefficients in (2.3) are given by (2.6).

A.2 Identification

We next describe the identification results in Section 2 in more detail. Using (2.2), (2.3), and the laws of motion for x_t^s and x_t^a, we obtain

$$v_{t+h}^h = F^h \cdot z_{t+h}^h,$$

where F^h is a coefficient vector and z_{t+h}^h is the following vector of variables:

$$z_{t+h}^h = \Big[x_{t-1}^s, \{\varepsilon_{t+i}^s\}_{i=1}^h, x_{t-1}^s\phi_{t+h}, \varepsilon_t^s\{\eta_{t+i}\}_{i=1}^h, \varepsilon_t^s\{\varepsilon_{t+i}^s\}_{i=1}^h, \varepsilon_t^s\{\varepsilon_{t+i}^a\}_{i=1}^h,$$
$$\{\varepsilon_{t+i}^s\phi_{t+h}\}_{i=1}^h, \{\eta_{t+i}\}_{i=1}^h, \{\varepsilon_{t+i}^a\}_{i=1}^h, x_{t+h}^a, x_{t+h}^a\phi_{t+h}\Big]',$$

where $\{\varepsilon_{t+i}^s\}_{i=1}^h$ denotes the vector of all ε_{t+i}^s for $i = 1$ through $i = h$, and analogously for all terms in braces. Defining the vector of regressors (excluding the intercept) in (2.4) by $X_t = [\varepsilon_t^s,\ \varepsilon_t^s\phi_t,\ \phi_t]'$, consistency of the OLS estimates of $(\beta^h, \gamma^h, \delta^h)$ requires

$$E[X_t(z_{t+h}^h)'] = \mathbf{0},$$

where $\mathbf{0}$ denotes a zero matrix with conforming dimension. This orthogonality condition is satisfied if $\zeta^s = \zeta^a = 0$. We next turn to the IV estimator of $(\beta^h, \gamma^h, \delta^h)$. Consider an instrument ϕ_t^{IV} with the following properties:

$$E[\phi_t^{IV}\varepsilon_{t+i}^s] = E[\phi_t^{IV}\varepsilon_{t+i}^a] = 0 \quad \forall i, \qquad E[\phi_t^{IV}\eta_t] \neq 0, \qquad E[\phi_t^{IV}\eta_{t+i}] = 0 \quad \forall i \neq 0.$$

Defining as instrument vector $Q_t = \left[\varepsilon_t^s, \varepsilon_t^s \phi_t^{IV}, \phi_t^{IV}\right]'$, consistency of the IV estimator requires

$$E[Q_t(z_{t+h}^h)'] = \mathbf{0}.$$

This condition is satisfied given the properties of the instrument.[48] Hence, the IV estimator consistently estimates $(\beta^h, \gamma^h, \delta^h)$ even absent strong exogeneity assumptions for ϕ_t.

Appendix B Data

B.1 Narrative data

We use the classification of FOMC members' policy preferences at each FOMC meeting for 1960–2023 from Istrefi (2019). The news coverage of FOMC members is relatively sparse during the first six years in our sample, leaving us with relatively many unclassified FOMC members in this period. In those years, we impute missing data by assuming that the unobserved preferences coincide with the first observed preference of the respective FOMC member. Occasionally, voting FOMC members do not attend the meetings personally, but are replaced by a substitute. We assume short-term substitutes act in the best interest of the substituted person – the substitutes are often subordinates of the original voting member. Specifically, we impose the original member's preference on the short-term substitute if three criteria hold: (i) the substitution period is no longer than six months when the substitute is from the same FRB, (ii) the substitution period is no longer than three months if the substitute is not from the same FRB, (iii) the substitution does not take place at the beginning or the end of a rotation cycle within a rotation group.[49] However, it frequently holds that the preferences of the substitute and the original voter coincide. In total, we change fewer than 1% of preferences in case of substitutions and our results are insensitive to these changes.

B.2 Macroeconomic data

We take the series for potential output (*rgdp_pott6*), real GDP (*rgdp*), nominal government spending (*ngov*), the GDP deflator (*pgdp*) and the military spending news shock (*news*) from the replication package of Ramey and Zubairy (2018). We follow their data preparation

---

[48] Note that $E[Q_t(z_{t+h}^h)'] = 0$ requires not only $E[\phi_t^{IV}\varepsilon_{t+i}] = 0$ but also $E[\phi_t^{IV}(\varepsilon_{t+i})^2] = 0$. However, given the assumption that ε_t and η_t are mutually independently distributed, and given the law of motion for ϕ_t, the second condition is satisfied if the first condition is satisfied.

[49] For example, suppose the Chicago president had the voting right until meeting τ and the Cleveland president thereafter. If Chicago exercises the voting right in $\tau + 1$ on behalf of Cleveland, we would use the preference of the Chicago president in $\tau + 1$.

steps to create the aggregate series as in their paper.[50] From FRED, we use headline CPI (*CPIAUCSL*) and CPI core (*CPILFESL*) inflation defined as the year-over-year growth rate of the respective price index, and the effective federal funds rate (*DFF*). The 10-year treasury market yield (*DGS10*) starts only in 1962q1 and is therefore combined with the same variable from Romer and Romer (2010) to obtain a series that starts in 1960q1. Similarly, we use the 1-year market yield from Liu and Wu (2021) and use for the first four observations (1960q1 to 1960q4) the 1-year treasury market yield from FRED (*DTB1YR*). Federal government defense expenditures (*FDEFX*) is divided by the GDP deflator and by real potential GDP, as described above. Variables are averaged to quarterly frequency, if applicable. We use inflation expectations from the Livingston survey. Our measure of inflation expectation is the annualized expected growth rate of CPI forecasts from 6 to 12 months ahead. Because the survey is biannual, we assume that inflation expectations remain constant in quarters in which no new data is available.

The validation exercise in Section 3 is based on the Fed's Greenbook forecasts. We use three-quarter ahead forecasts for the output gap and for inflation, both as quarterly averages. We construct the Blanchard and Perotti (2002) shock controlling for anticipation in government spending by including the one-quarter projected growth rate of government spending from Ramey (2011). We further consider the primary surplus (svt_q) from Cochrane (2022), seasonally adjusted via X-13 ARIMA-SEATS procedure from the U.S. Census Bureau.

Appendix C Hawk-Dove decompositions

We derive a decomposition of the aggregate Hawk-Dove balance. We first rewrite the aggregate Hawk-Dove balance in equation (3.2) as $Hawk_t = \sum_{i\in\mathcal{M}_t} s_t Hawk_{it}$ with $s_t = 1/|\mathcal{M}_t|$. We define a decomposition over p-period changes in $Hawk_t$:

$$\Delta^p Hawk_t = Hawk_t - Hawk_{t-p} = \sum_{i\in\mathcal{M}_t} s_t Hawk_{it} - \sum_{i\in\mathcal{M}_{t-p}} s_{t-p} Hawk_{it-p} \tag{C.1}$$

We next partition the set $\mathcal{M}_t$ into the set of "surviving" FOMC members S_t present in $t-p$ and t, the set of entering FOMC members E_t present in t but not in $t-p$, and the set of exiting FOMC members X_t present in $t-p$ but not in t. We further distinguish between the rotating and non-rotating FOMC members in the set of entering and exiting FOMC

---

[50] The fiscal shock is computed as $news_t/(pgdp_{t-1} \times rgdp\_pott6_{t-1}) \times 100$. Detrended real GDP is $rgdp_t/rgdp\_pott6_t \times 100$ and detrended real government spending is $ngov_t/(pgdp_t \times rgdp\_pott6_t) \times 100$.

members, denoted E_t^R, E_t^N, X_t^R and X_t^N to obtain the following decomposition:

$$\begin{aligned}\Delta^p Hawk_t = &\sum_{i\in S_t} s_{t-p}(Hawk_{it} - Hawk_{it-p}) + \sum_{i\in S_t}(s_t - s_{t-p})Hawk_{it} + \sum_{i\in E_t^N} s_t Hawk_{it} \\ &- \sum_{i\in X_t^N} s_{t-p}Hawk_{it-p} + \sum_{i\in E_t^R} s_t Hawk_{it} - \sum_{i\in X_t^R} s_{t-p}Hawk_{it-p} \qquad \text{(C.2)}\end{aligned}$$

The first two terms capture changes in the preferences and the number of surviving FOMC members. The third and fourth terms capture changes in the aggregate Hawk-Dove balance due to the entry and exit of rotating FOMC members, while the fifth and sixth terms capture the contribution of entry and exit of non-rotating FOMC members.

The variance in yearly changes of the aggregate Hawk-Dove balance ($p = 4$) is 0.084. The variance of the first term of (C.2) corresponds to 11% of the total variance and the second term is negligible in size. The variance of the third and fourth term, capturing extensive margin changes of non-rotating FOMC members, corresponds to 29% of the total variance. The variance of the fifth and sixth term, capturing extensive margin changes of rotating FOMC members, corresponds to 61% of the total variance. Finally, the covariances between these terms account for the remainder of the total variance. The results differ little for quarterly changes ($p = 1$). Notably, extensive margin changes of rotating FOMC members still account for 59% of the total variance.

Analogously, we propose a decomposition for the FOMC rotation instrument

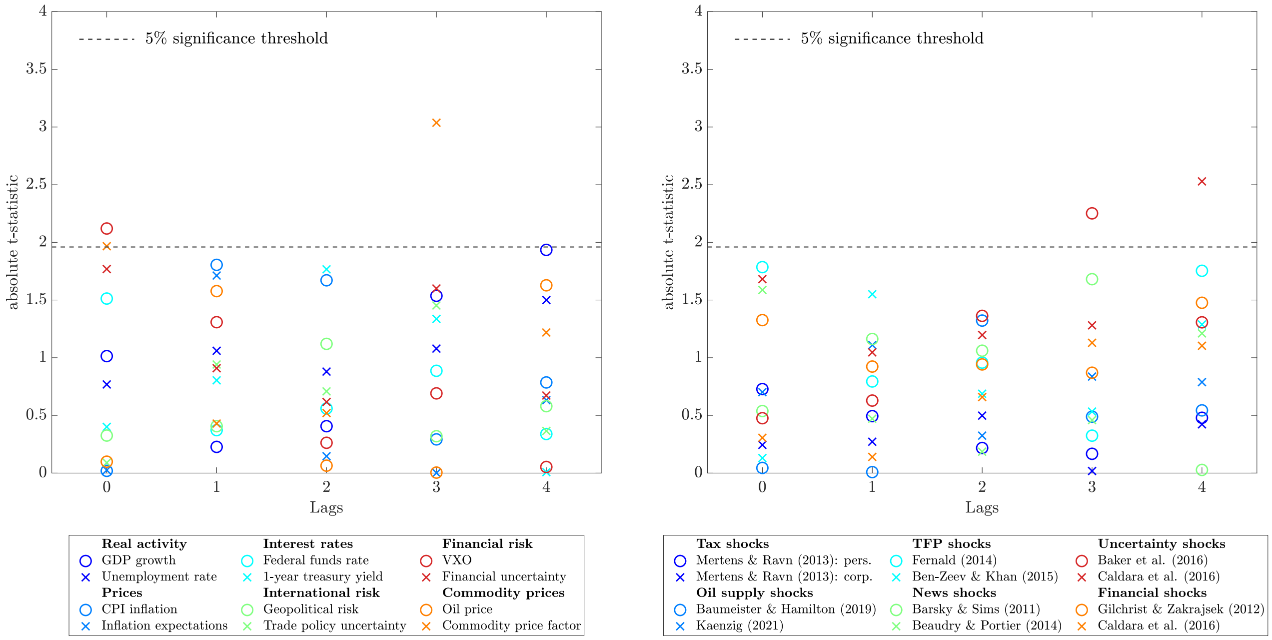
$$\begin{aligned}\Delta^p Hawk_t^{\mathcal{R}} = &\sum_{i\in S_t^R} s_{t-p}^R(Hawk_{it} - Hawk_{it-p}) + \sum_{i\in S_t^R}(s_t^R - s_{t-p}^R)Hawk_{it} \\ &+ \left(\sum_{i\in E_t^{RA}} s_t^R Hawk_{it} - \sum_{i\in X_t^{RA}} s_{t-p}^R Hawk_{it-p}\right) \\ &+ \left(\sum_{i\in E_t^{RI}} s_t^R Hawk_{it} - \sum_{i\in X_t^{RI}} s_{t-p}^R Hawk_{it-p}\right), \qquad \text{(C.3)}\end{aligned}$$

with the weights given by $s_t^R = 1/|\mathcal{R}_t|$, S_t^R the set of surviving rotating FOMC members, and distinguishing between the sets of entering rotating FOMC members whose appointments start or end in t (A), and incumbent (I) regional FRB presidents.

For yearly changes in the rotation instrument, we find that 93% of the variance is due to the rotation of incumbent members, while 7% is due to appointments starting or ending. All other variances and covariances are negligible in size. We also study quarterly changes ($p = 1$). Intensive margin changes now explain 4% of the variance, appointments account for 26%, and rotations of incumbent members account for 69%. Appointments become relatively more important for $p = 1$ because only every fourth quarter of $\Delta^1 Hawk_t^{\mathcal{R}}$ features a rotation. Compared to $\Delta^4 Hawk_t^{\mathcal{R}}$ for which the rotation affects all quarters, we mechanically lower the importance of rotations and the overall variance for $p = 1$.

Appendix D Further validation

Figure D.1: Predictive regression of $Hawk_t^{\mathcal{R}}$ on macroeconomic variables

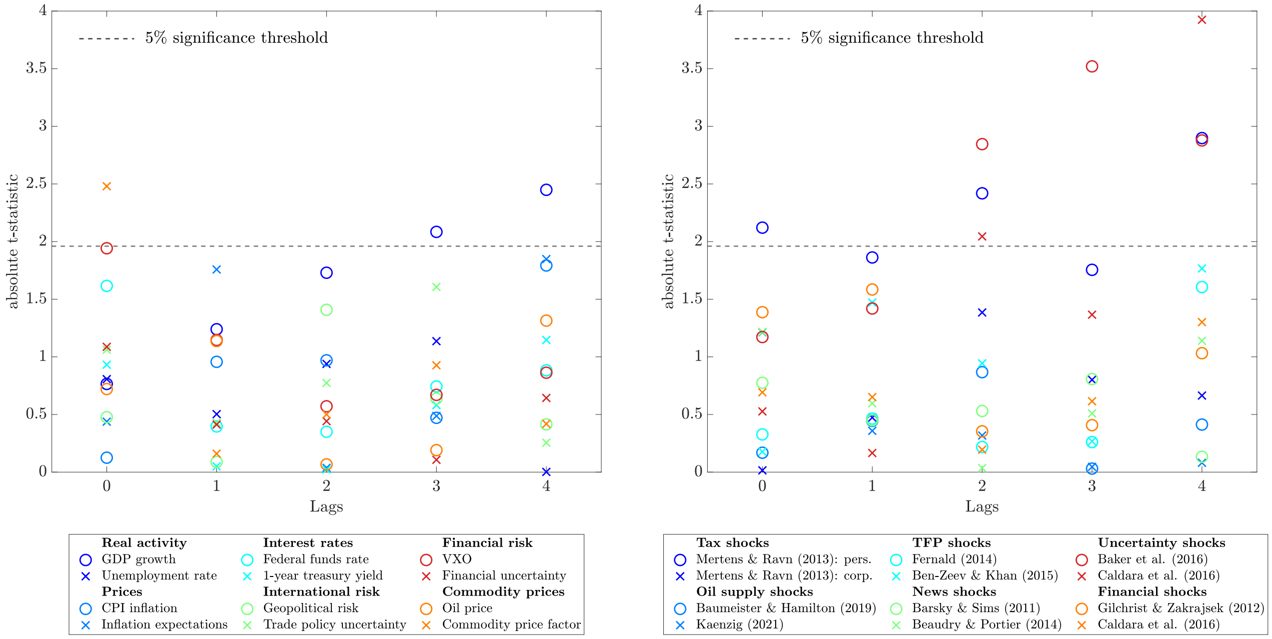


(a) Macroeconomic indicators

(b) Macroeconomic shocks

Notes: This figure shows absolute t-statistics from OLS regressions of $Hawk_t^{\mathcal{R}}$ on the contemporaneous value and four quarterly lags of a given predictor. The predictor is either a macroeconomic indicator in Panel (a) or a macroeconomic shock from the literature in Panel (b). Macroeconomic indicators: Real GDP growth is the quarter-on-quarter growth rate. The unemployment rate is from the CPS survey. The inflation rate is based on the CPI. Inflation expectations are from the Livingston survey. The indicators of geopolitical risk and trade policy uncertainty are from Caldara and Iacoviello (2022) and Caldara et al. (2020), respectively. The VXO is the S&P 100 Volatility Index from the Chicago Board Options Exchange. Financial uncertainty is taken from Jurado et al. (2015). The oil price is the natural logarithm of the Western Texas Intermediate price. The commodity price factor is taken from Baumeister and Guérin (2021). Macroeconomic shocks: The personal and corporate income tax shocks are from Mertens and Ravn (2013). The oil supply shocks are from Baumeister and Hamilton (2019) and Känzig (2021). The TFP shocks are from Fernald (2014) and Ben-Zeev and Khan (2015). The news shocks are from Barsky and Sims (2011) and Beaudry and Portier (2014). The uncertainty shocks are from Baker et al. (2016) and Caldara et al. (2016). The financial shocks are from Gilchrist and Zakrajšek (2012) and Caldara et al. (2016). The horizontal dashed line indicates an absolute t-statistic of 1.96 corresponding to testing the null of a zero coefficient at the 5% level.

Figure D.2: Predictive regression of $Hawk_t$ on macroeconomic variables



(a) Macroeconomic indicators

(b) Macroeconomic shocks

Notes: This figure shows absolute t-statistics from OLS regressions of $Hawk_t$ on the contemporaneous value and four quarterly lags of a given predictor, see Figure D.1 for details. The horizontal dashed line indicates an absolute t-statistic of 1.96 corresponding to testing the null of a zero coefficient at the 5% level.

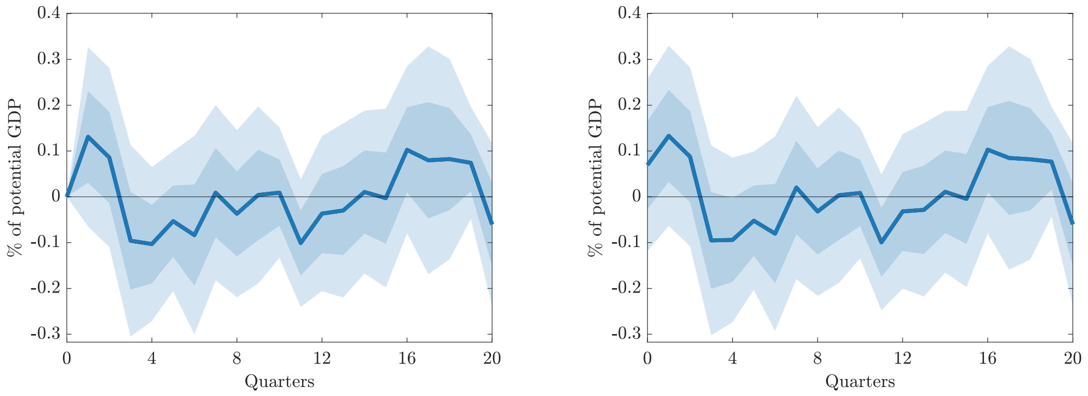
Table D.1: Communication of FRB presidents with and without voting right

	Number of interventions			Intervention length in words		
	Full meeting	Policy round	Remaining meeting	Full meeting	Policy round	Remaining meeting
	(1)	(2)	(3)	(4)	(5)	(6)
Voting indicator	1.664^{***}	0.998^{***}	0.578^{***}	57.805^{***}	38.391^{***}	15.792
	(0.321)	(0.198)	(0.211)	(17.944)	(9.439)	(12.682)
Member fixed effects	✓	✓	✓	✓	✓	✓
Meeting fixed effects	✓	✓	✓	✓	✓	✓
Sample average	9.22	3.34	5.94	1425.73	398.26	1041.26
R^2	0.60	0.52	0.59	0.84	0.74	0.81
Observations	4006	3931	3931	4006	3931	3931

Notes: This table shows results on the communication behavior of FRB presidents, based on FOMC Transcripts (1976 - 2019). We show OLS estimates of β based on $y_{i,\tau} = \alpha_i + \alpha_\tau + \beta \text{VotingIndicator}_{i,\tau} + u_{i,\tau}$, where $y_{i,\tau}$ is either the *Number of interventions* or the *Length of interventions* (number of words) by member i in meeting τ, at different segments of the FOMC meeting (Full, Policy Round, Remaining), and $\text{VotingIndicator}_{i,\tau} \in \{0,1\}$ equals unity when i has voting right at meeting τ. Additionally, α_i and α_τ denote member and meeting fixed effects, and $u_{i,\tau}$ is an error term. Standard errors clustered at the member and meeting level are reported in parentheses, and *,**,*** indicates significance at the 0.1, 0.05, 0.01 levels, respectively.

Appendix E Additional results for Section 4

Figure E.1: Responses of military spending shocks to systematic monetary policy



(a) Baseline model (δ^h)

(b) Linear model (δ^h)

Notes: The figure shows responses of the military spending shock to systematic monetary policy ($Hawk_t$). We show IV estimates based on the local projection framework (3.5)-(3.6) as specified in Section 4.1. The δ^h captures the response when $Hawk_t$ exceeds the sample average by two hawks. Panel (a) shows the results for our baseline model whereas Panel (b) shows the results when we restrict $\beta^h = \gamma^h = 0$ in the local projection (3.5). The shaded areas indicate 68% and 95% confidence bands using Newey-West standard errors.

Table E.1: Testing for differences across regimes, p-values

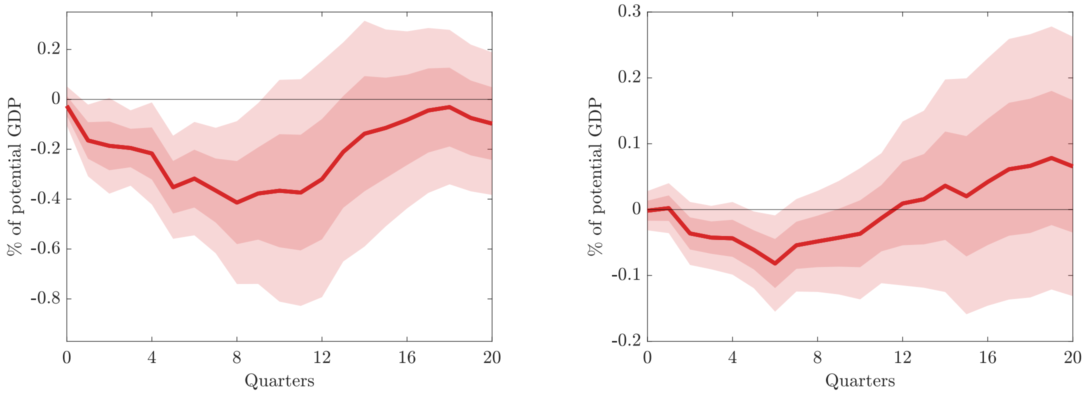
Outcome	+2 Hawk vs. Average	+1 Hawks vs. Average	Average vs. +1 Dove	Average vs. +2 Doves
	Two-year horizon			
Multiplier	0.211	0.114	0.100	0.103
GDP (*cum*)	0.000			
G (*cum*)	0.080			
	Four-year horizon			
Multiplier	0.233	0.114	0.039	0.039
GDP (*cum*)	0.007			
G (*cum*)	0.437			

Notes: The table shows p-values corresponding to statistical tests for whether the fiscal multiplier or its components are significantly different across monetary regimes ($Hawk_t$). The tests are based on the multiplier estimates reported in Table 2 in Section 4.3, using Driscoll-Kraay standard errors, see Appendix F for details.

Appendix F Weak instruments and robust inference

F.1 Weak instrument tests

Figure F.1: Differential responses of GDP and government spending, reduced-form



(a) Differential GDP (γ^h)

(b) Differential G (γ^h)

Notes: The figure shows differential responses of real GDP and real government spending (G) to an expansionary military spending shock, corresponding to one percent of GDP, conditional on the FOMC rotation instrument ($Hawk_t^{\mathcal{R}}$). We show reduced-form estimates based on the local projection framework (3.5)-(3.6) as specified in Section 4.1. The γ^h captures the differential responses when $Hawk_t$ exceeds the sample average by two hawks. Moreover, testing whether γ^h is statistically significant from zero is equivalent to testing for zero relevance of the instrument, as explained in the main text. The shaded areas indicate 68% and 95% confidence bands using Newey-West standard errors.

F.2 Robust inference

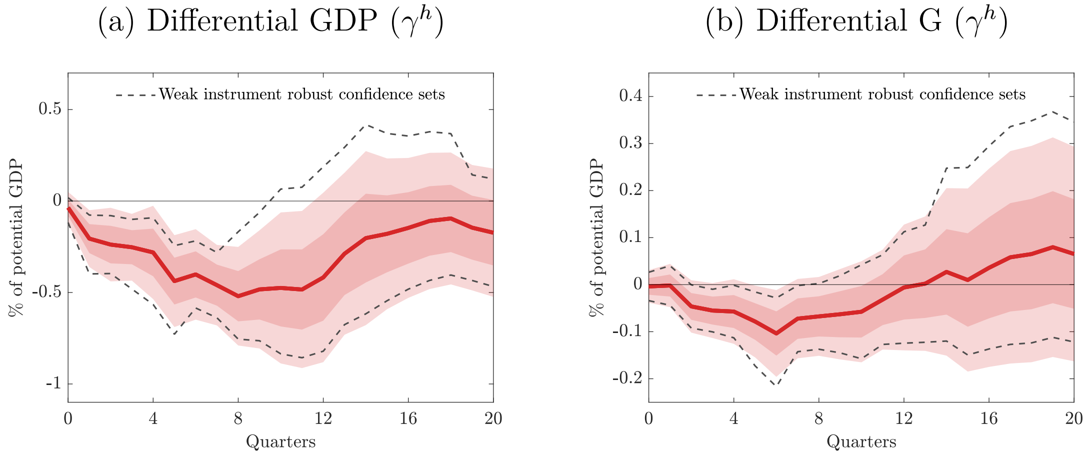
Differential responses. We compute robust inference for the differential GDP and government spending effects based on Andrews (2018).

Baseline multiplier inference. To provide (baseline) inference for the multiplier estimates, we estimate the responses of cumulative GDP and cumulative G through

$$\tilde{x}_t = \tilde{\alpha}_x + \tilde{\beta}_x \varepsilon_t^s + \tilde{\gamma}_x \varepsilon_t^s (Hawk_t - \overline{Hawk}) + \tilde{\delta}_x (Hawk_t - \overline{Hawk}) + \tilde{\zeta}_x Z_{t-1} + \tilde{v}_t, \qquad \text{(F.1)}$$

where $\tilde{x}_t$ is cumulative GDP ($\tilde{x}_t = \sum_{h=0}^{H} GDP_{t+h}$) or cumulative G ($\tilde{x}_t = \sum_{h=0}^{H} G_{t+h}$). This yields estimates $\tilde{\beta}_{\text{GDP}} = \sum_{h=0}^{H} \beta_{\text{GDP}}^h$, and $\tilde{\beta}_{\text{G}} = \sum_{h=0}^{H} \beta_{\text{G}}^h$. The coefficients $\tilde{\alpha}_x$, $\tilde{\gamma}_x$, $\tilde{\delta}_x$, $\tilde{\zeta}_x$ are analogously related to (3.5). To obtain a covariance matrix for the IV estimates $\hat{\vartheta} = (\tilde{\beta}_{\text{GDP}}, \tilde{\beta}_{\text{G}}, \tilde{\gamma}_{\text{GDP}}, \tilde{\gamma}_{\text{G}})'$, we estimate the stacked regression (i.e., for GDP and G). We use the Driscoll and Kraay (1998) covariance estimator, allowing for serial correlation and cross-correlation between GDP and G. We compute standard errors for the fiscal multiplier by applying the Delta method to the fiscal multiplier in (4.1).

Figure F.2: Responses of GDP and government spending, robust inference


(a) Differential GDP (γ^h)

(b) Differential G (γ^h)

Notes: The figure shows differential responses of real GDP and real government spending (G) to an expansionary military spending shock, corresponding to one percent of GDP, conditional on systematic monetary policy ($Hawk_t$). We show IV estimates based on the local projection framework (3.5)-(3.6) as specified in Section 4.1. The γ^h captures the differential responses when $Hawk_t$ exceeds the sample average by two hawks. The shaded areas indicate 68% and 95% confidence bands using Newey-West standard errors. The dashed bands provide 95% confidence sets, robust to weak identification based on Andrews (2018), constructed via the refined projection method from Chaudhuri and Zivot (2011).

Anderson-Rubin multiplier inference. We construct robust confidence sets for the fiscal multiplier by inverting an Anderson and Rubin (1949) test (AR henceforth) following Andrews et al. (2019) via two regressions. First, consider the reduced-form regression

$$\tilde{x}_t = \tilde{\alpha}_x^{rf} + \tilde{\beta}_x^{rf}\varepsilon_t^s + \tilde{\gamma}_x^{rf}\varepsilon_t^s(Hawk_t^{\mathcal{R}} - \overline{Hawk}^{\mathcal{R}}) + \tilde{\delta}_x^{rf}(Hawk_t^{\mathcal{R}} - \overline{Hawk}^{\mathcal{R}}) + \tilde{\zeta}_x^{rf} Z_{t-1} + \tilde{v}_t^{rf}, \quad \text{(F.2)}$$

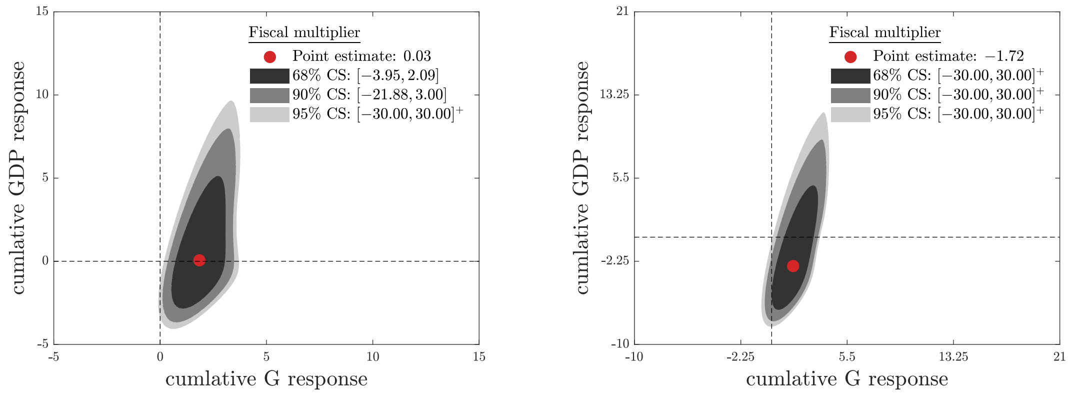
and ϱ denotes the OLS estimator of parameters $(\tilde{\beta}_{\text{GDP}}^{rf},\ \tilde{\gamma}_{\text{GDP}}^{rf},\ \tilde{\beta}_{\text{G}}^{rf},\ \tilde{\gamma}_{\text{G}}^{rf})'$. Second, consider the first-stage regressions

$$\begin{aligned} \varepsilon_t^s &= \tilde{\alpha}^{fs1} + \tilde{\beta}^{fs1}\varepsilon_t^s + \tilde{\gamma}^{fs1}\varepsilon_t^s(Hawk_t^{\mathcal{R}} - \overline{Hawk}^{\mathcal{R}}) \\ &\quad + \tilde{\delta}^{fs1}(Hawk_t^{\mathcal{R}} - \overline{Hawk}^{\mathcal{R}}) + \tilde{\zeta}_x^{fs1} Z_{t-1} + \tilde{v}_t^{fs1}, && \text{(F.3)} \\ \varepsilon_t^s(Hawk_t - \overline{Hawk}) &= \tilde{\alpha}^{fs2} + \tilde{\beta}^{fs2}\varepsilon_t^s + \tilde{\gamma}^{fs2}\varepsilon_t^s(Hawk_t^{\mathcal{R}} - \overline{Hawk}^{\mathcal{R}}) \\ &\quad + \tilde{\delta}^{fs2}(Hawk_t^{\mathcal{R}} - \overline{Hawk}^{\mathcal{R}}) + \tilde{\zeta}_x^{fs2} Z_{t-1} + \tilde{v}_t^{fs2}, && \text{(F.4)} \end{aligned}$$

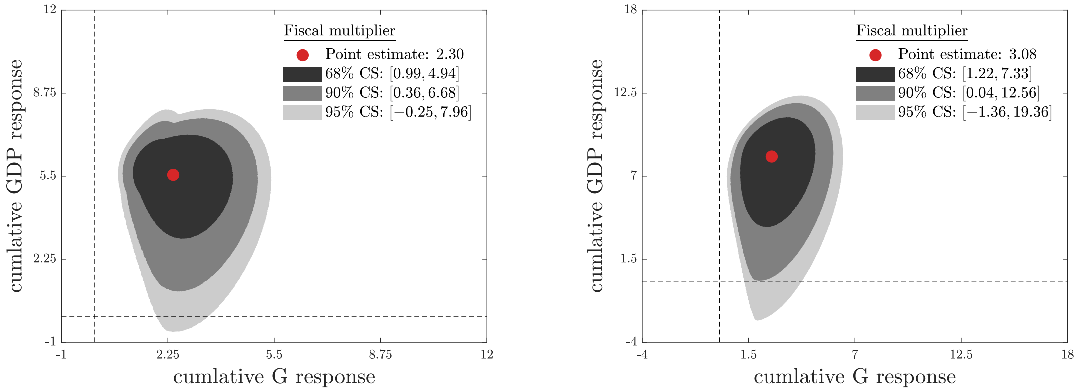
and π denotes the OLS estimator of the 2×2 parameter matrix $((\tilde{\beta}^{fs1}, \tilde{\gamma}^{fs1})', (\tilde{\beta}^{fs2}, \tilde{\gamma}^{fs2})')$. We further define $\Pi = I_2 \otimes \pi$ with I_2 a 2×2 identity matrix, which corresponds to the OLS estimators of the stacked first stage regressions for GDP and G. The AR statistic builds on the identity $\varrho = \Pi\vartheta$ where ϑ is the IV estimator of the coefficients of interest, see Andrews et al. (2019).

Figure F.3: Anderson-Rubin confidence sets for four-year fiscal multipliers

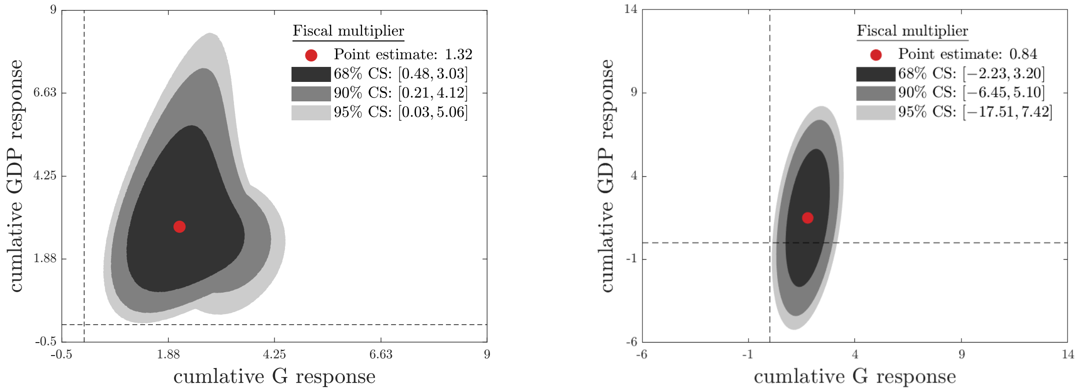
(a) +1 Hawk (b) +2 Hawks



(c) +1 Dove (d) +2 Doves



(e) Average (f) Linear model



Notes: This figure shows Anderson-Rubin type confidence sets for the cumulative four-year fiscal multiplier. We depict the numerator and denominator of the multiplier on the vertical and horizontal axis, respectively. The shaded areas depict the confidence sets and various levels of significance. The red circle is the baseline point estimate from Table 2. The dashed lines indicate the zero values on each axis, respectively. The confidence sets reported in the legend are defined by the minimum and maximum fiscal multiplier that is contained in the respective confidence set, capped at ±30 for readability. Panels (a)-(d) correspond to the fiscal multipliers when $Hawk_t$ exceeds the sample average by either one or two hawks or doves. Panel (e) corresponds to the fiscal multiplier when $Hawk_t$ equals its sample average. Panel (f) corresponds to the fiscal multiplier estimate when we restrict $\tilde{\gamma}_{\text{GDP}} = \tilde{\delta}_{\text{GDP}} = \tilde{\gamma}_{\text{G}} = \tilde{\delta}_{\text{G}} = 0$.

The test statistic for H_0: $\vartheta = \vartheta_0$ is given by

$$AR(\vartheta_0) = \hat{g}(\vartheta_0)' \, \hat{\Omega}(\vartheta_0)^{-1} \, \hat{g}(\vartheta_0), \tag{F.5}$$

with $\hat{g}(\vartheta_0) = \hat{\varrho} - \hat{\Pi}\vartheta_0$, and $\hat{\Omega}(\vartheta_0) = \hat{\mathbb{E}}\left[\varrho \; \varrho'\right] - \hat{\mathbb{E}}\left[\varrho \; \vartheta_0' \; \Pi'\right] - \hat{\mathbb{E}}\left[\Pi \; \vartheta_0 \; \varrho'\right] + \hat{\mathbb{E}}\left[\Pi \; \vartheta_0 \; \vartheta_0' \; \Pi'\right]$, and hats denote estimates. We estimate all covariance terms in $\hat{\Omega}(\vartheta_0)$ accounting for cross-correlations between estimators as well as for serial correlation using the Driscoll-Kraay covariance estimator. Under weak assumptions, it holds that $AR(\vartheta_0) \xrightarrow{d} \chi^2(4)$, see Andrews et al. (2019). This holds regardless of the strength of the instrument and regardless of whether the denominator of the fiscal multiplier is zero.

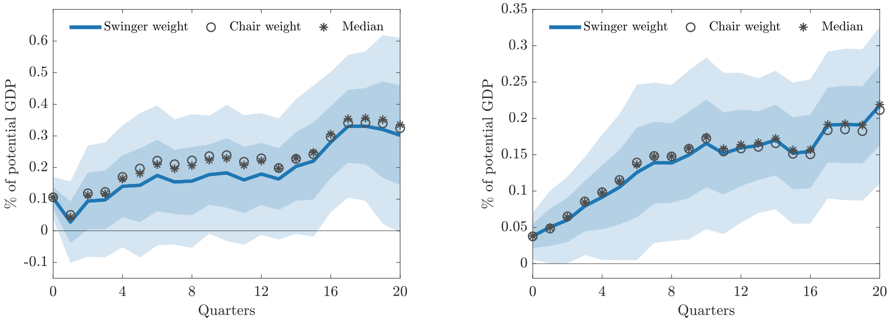
We compute the AR confidence set $CS^{FM}(\chi)$ for the fiscal multiplier $FM^H(\chi)$ from equation (4.1) by inverting the AR test. This requires four steps: First, define the set Θ that contains the confidence region of $FM^H(\chi)$. Second, define the discrete set $\Theta_N \subset \Theta$ that contains N vectors of ϑ. Third, construct the set $CS^{\vartheta} = \left\{\vartheta \in \Theta_N \mid AR(\vartheta) \leq c_{1-\alpha,\chi^2(4)}\right\}$. Fourth, compute the confidence set for the fiscal multiplier as

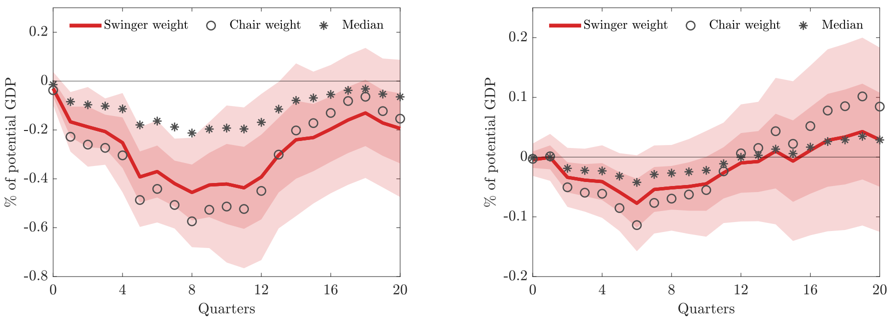
$$CS^{FM}(\chi) = \left\{ FM \;\middle|\; FM = \frac{\tilde{\beta}_{\mathrm{GDP}} + \chi \; \tilde{\gamma}_{\mathrm{GDP}}}{\tilde{\beta}_{\mathrm{G}} + \chi \; \tilde{\gamma}_{\mathrm{G}}}, \; \forall \; (\tilde{\beta}_{\mathrm{GDP}}, \; \tilde{\gamma}_{\mathrm{GDP}}, \; \tilde{\beta}_{\mathrm{G}}, \; \tilde{\gamma}_{\mathrm{G}})' = \vartheta \in CS^{\vartheta} \right\}.$$

Note that $c_{1-\alpha,\chi^2(4)}$ is the $1-\alpha$ quantile of a χ^2 distribution with four degrees of freedom. We implement step 1 by choosing a closed interval for each entry of the vector ϑ. Θ is the Cartesian product of the four closed intervals. Specifically for entry i of ϑ, which we denote by ϑ_i, we use the interval $[-1.5 \; \hat{\vartheta}_i, \; 3.5 \; \hat{\vartheta}_i]$, when $\hat{\vartheta}_i > 0$, and $[3.5 \; \hat{\vartheta}_i, \; -1.5 \; \hat{\vartheta}_i]$ when $\hat{\vartheta}_i < 0$, where $\hat{\vartheta}_i$ denotes the IV estimate, based on (F.1). We verify that the chosen intervals are not binding in the sense that the upper or lower bound of CS^{θ} is not the boundary of Θ. For the linear model, we require a larger set Θ with $[-4 \; \hat{\vartheta}_i, \; 10 \; \hat{\vartheta}_i]$ if $\hat{\vartheta}_i > 0$ and analogously if $\hat{\vartheta}_i < 0$. For step 2, we define Θ_N based on a Sobol sequence of length $N = 2{,}000{,}000{,}000$. Finally, we have verified that increasing or decreasing N by 5% does not affect our results.

Appendix G Sensitivity analysis

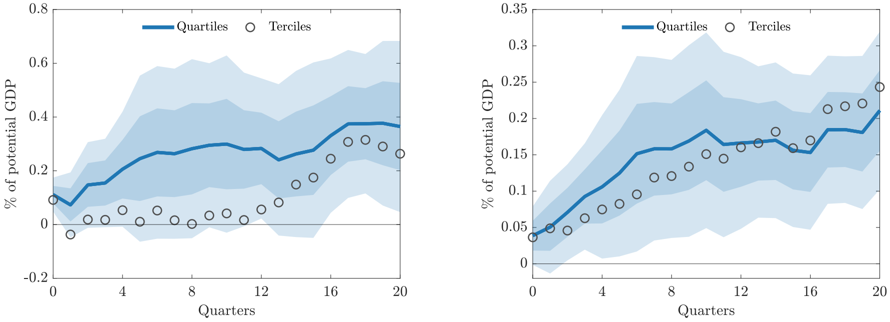
Figure G.1: Responses of GDP and government spending, aggregation schemes

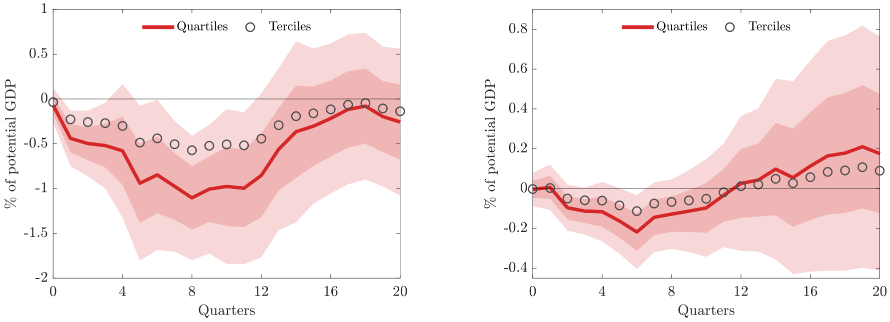

(a) Average GDP (β^h) (b) Average G (β^h)


(c) Differential GDP (γ^h) (d) Differential G (γ^h)

Notes: We use three variants of $Hawk_t$. Swinger weight: We do not discriminate between swingers and consistent members. Chair weight: We assign the preferences of the Fed Chair twice the weight of an ordinary member when aggregating to $Hawk_t$. Median: We aggregate the cross-section of FOMC members by the median, instead of the arithmetic average.

Figure G.2: Responses of GDP and government spending, discrete Hawk-Dove balance

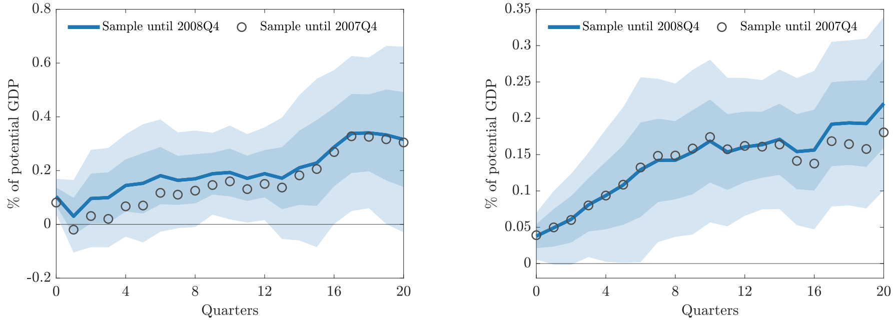

(a) Average GDP (β^h) (b) Average G (β^h)


(c) Differential GDP (γ^h) (d) Differential G (γ^h)

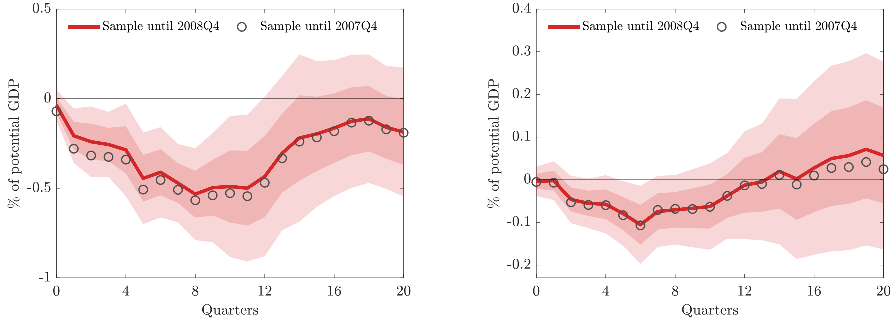
Notes: We use two discrete variants of $Hawk_t$. We define that the discrete $Hawk_t$ equals -1 if $Hawk_t$ falls below the first quartile or tertile of the distribution of $Hawk_t$ over time, +1 if above the highest quartile or tertile, and zero else.

Figure G.3: Responses of GDP and government spending, accounting for the ZLB

(a) Average GDP (β^h) (b) Average G (β^h)



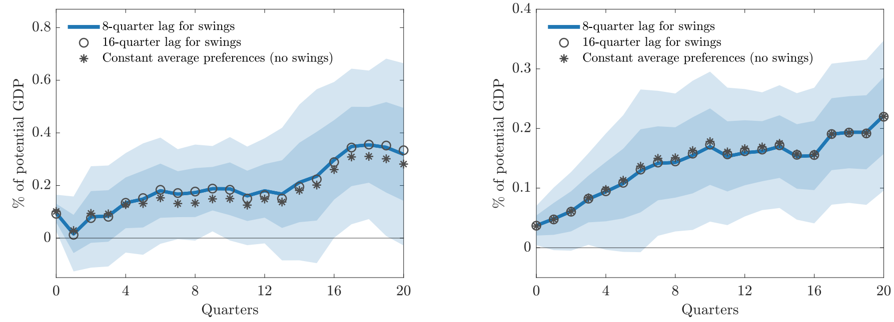
(c) Differential GDP (γ^h) (d) Differential G (γ^h)



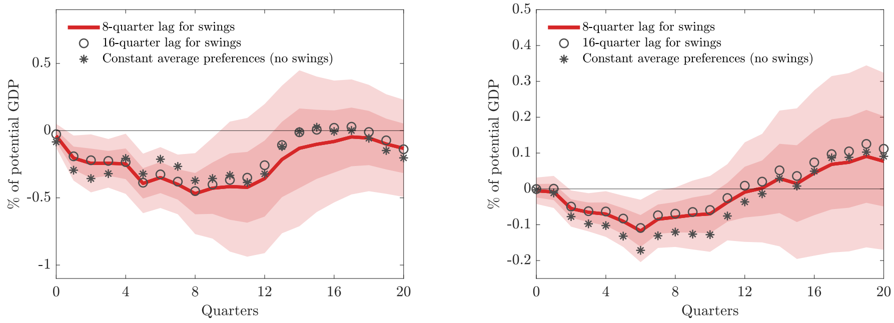
Notes: We use a sub-sample with shocks before either 2008Q4 or 2007Q4 to exclude the ZLB, or both the ZLB and the Great Recession.

Figure G.4: Responses of GDP and government spending, alternative IVs

(a) Average GDP (β^h) (b) Average G (β^h)



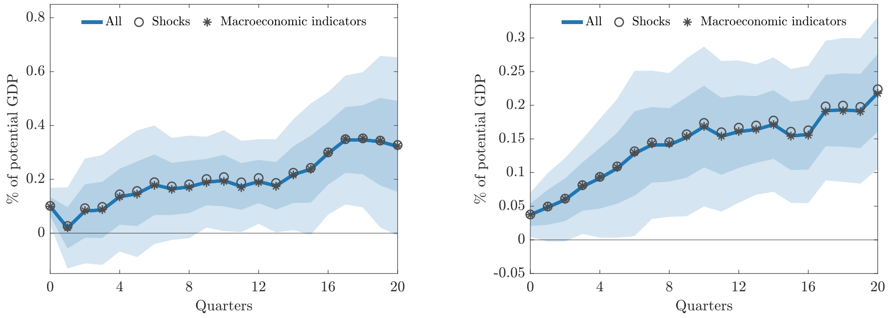
(c) Differential GDP (γ^h) (d) Differential G (γ^h)



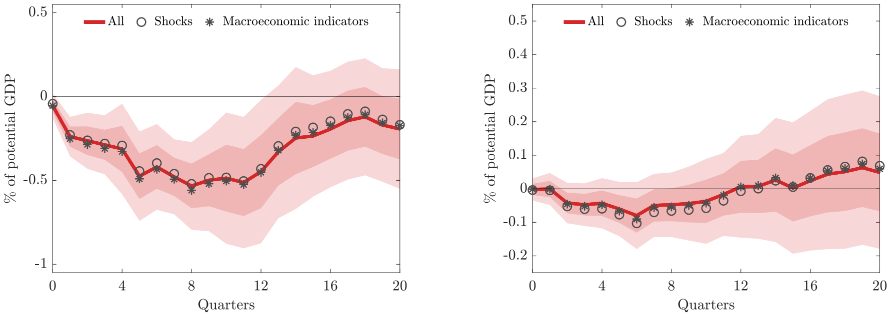
Notes: We use an alternative definition of the instrumental variable $Hawk_t^{IV}$ where swings affect the individual preference only 8 or 16 quarters after the date of the swing, or where no swing occurs because we set the individual preference to the average, rendering them time-invariant.

Figure G.5: Responses of GDP and government spending, residualized IVs

(a) Average GDP (β^h) (b) Average G (β^h)



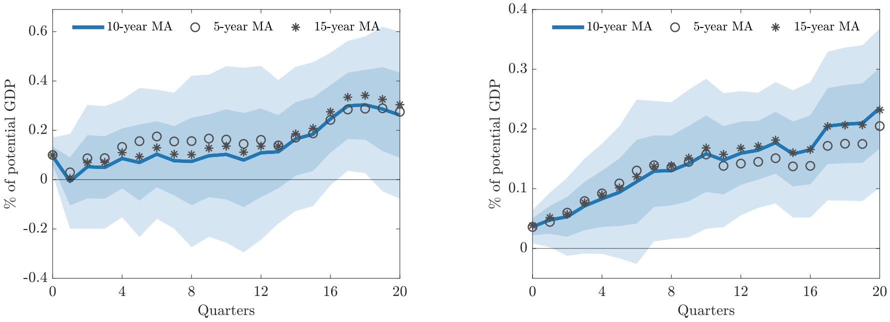
(c) Differential GDP (γ^h) (d) Differential G (γ^h)



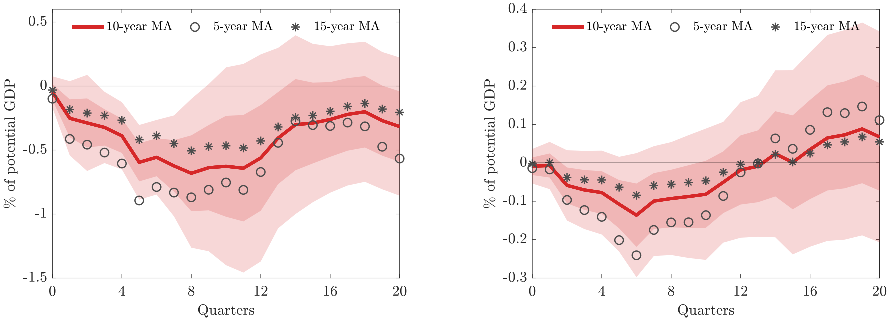
Notes: The different specifications replace $Hawk_t^{\mathcal{R}}$ by a purged version, where we first orthogonalize $Hawk_t^{\mathcal{R}}$ with respect to the contemporaneous values and the first four lags of both uncertainty shocks (Shocks), or with respect to the commodity price factor and the VXO (Macroeconomic indicators). Finally, we jointly orthogonalize $Hawk_t^{\mathcal{R}}$ with respect to all four variables jointly (All). These variables are the only ones with significant predictive power for $Hawk_t^{\mathcal{R}}$ as shown in Figure D.1.

Figure G.6: Responses of GDP and government spending, accounting for trends

(a) Average GDP (β^h) (b) Average G (β^h)

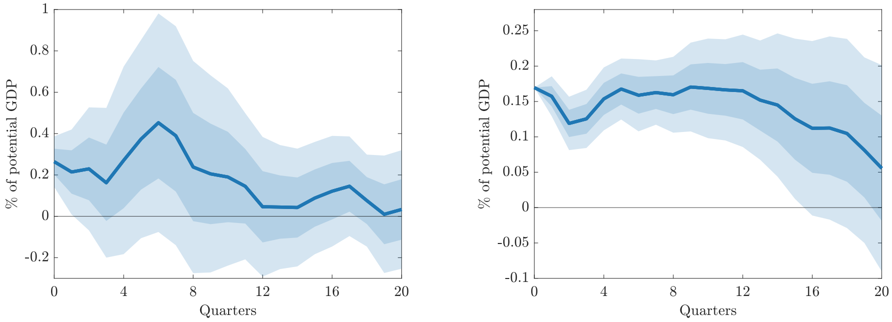


(c) Differential GDP (γ^h) (d) Differential G (γ^h)

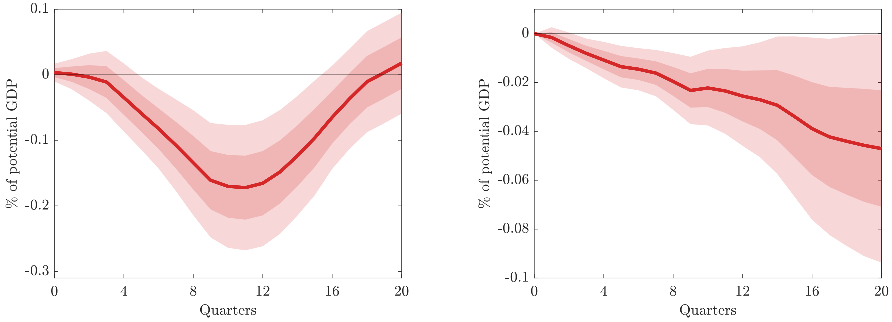


Notes: We use three variants of $Hawk_t$ where we subtract the backward-looking 5, 10, or 15-year moving average from $Hawk_t$ prior estimation.

Figure G.7: Responses of GDP and government spending, Blanchard-Perotti shock


(a) Average GDP (β^h)

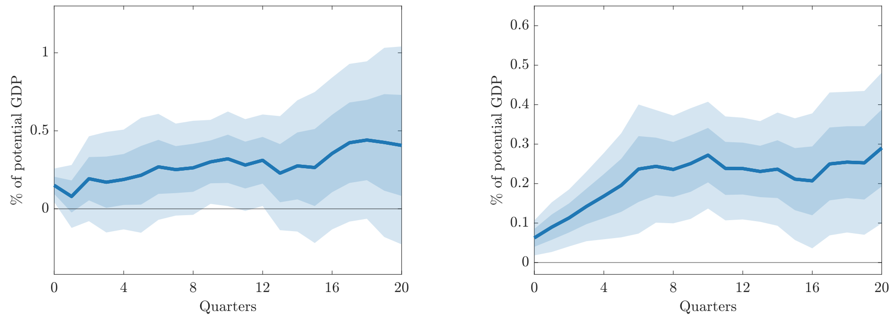
(b) Average G (β^h)


(c) Differential GDP (γ^h)

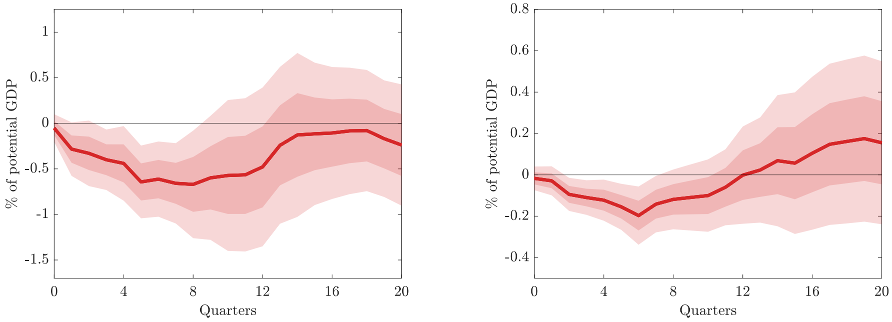
(d) Differential G (γ^h)

Notes: The shock is contemporaneous G, conditional on controls that include four lags of real GDP and real government spending, as well as the projected growth rate of real government spending. The projected growth rate is taken from the Survey of Professional Forecasters and is available from 1969 onward, which is the start of our sample, see Appendix B.

Figure G.8: Responses of GDP and government spending, winsorize large shocks


(a) Average GDP (β^h)

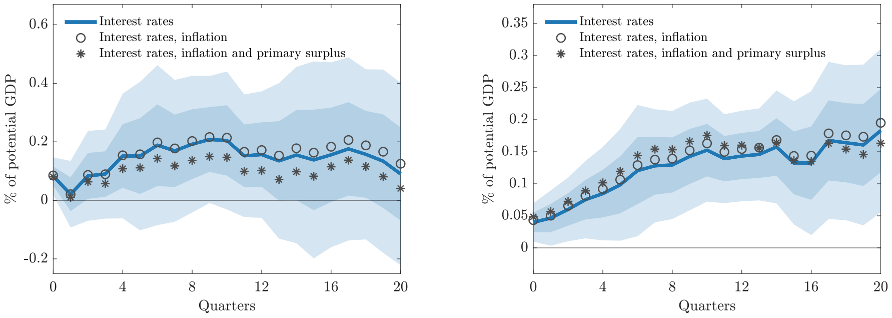
(b) Average G (β^h)


(c) Differential GDP (γ^h)

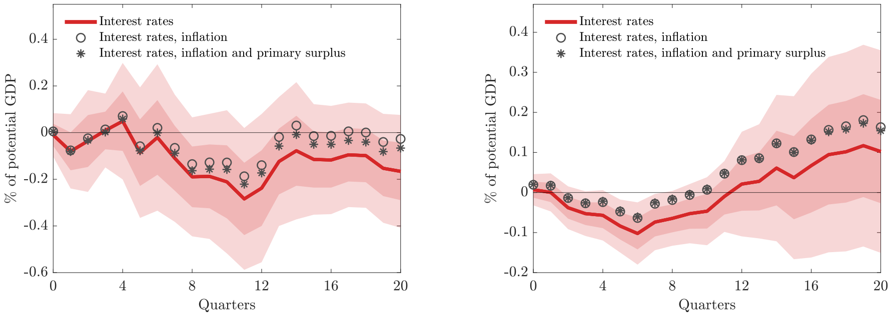
(d) Differential G (γ^h)

Notes: We use a winsorized version of the military spending news shocks from Ramey and Zubairy (2018), where we winsorize all realizations of the shock series that are above a threshold in absolute value to minimize the excess kurtosis.

Figure G.9: Responses of GDP and government spending, additional controls


(a) Average GDP (β^h)

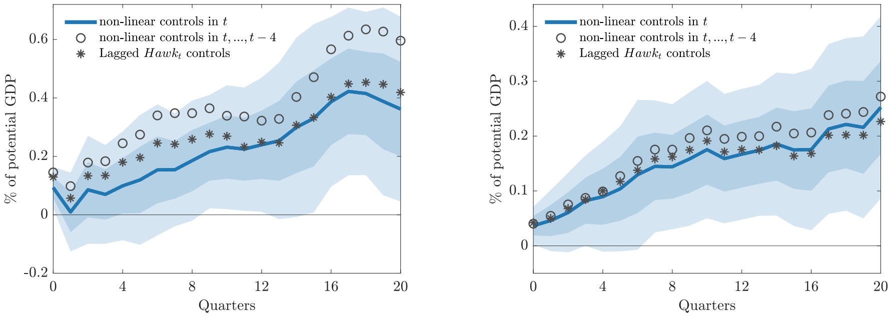
(b) Average G (β^h)


(c) Differential GDP (γ^h)

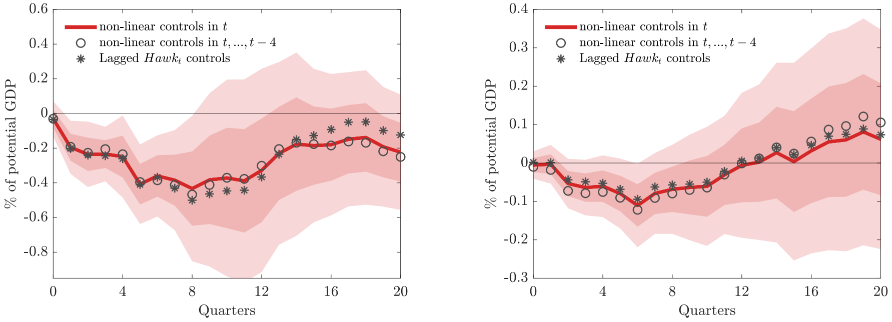
(d) Differential G (γ^h)

Notes: The different specifications augment the control vector Z_{t-1} gradually by four lags of treasury yields with 1-year and 10-year maturity, the fed funds rate (interest rates), CPI inflation, and the primary surplus from Cochrane (2022).

Figure G.10: Responses of GDP and government spending, non-linear controls


(a) Average GDP (β^h)

(b) Average G (β^h)


(c) Differential GDP (γ^h)

(d) Differential G (γ^h)

Notes: Non-linear controls in t: Controls Z_{t-1} include four lags of ε_t^s, real GDP and real government spending, both divided by potential GDP in all specifications. All controls are in levels, as well as interacted with $Hawk_t$, and instrumented accordingly. Non-linear controls in $t, \ldots, t-4$: Augments the control vector by also including and instrumenting lagged interaction terms, i.e. $Hawk_{t-i} \times C_{t-i}$ with $i = 1, \ldots, 4$ and C_t referring to G, GDP, and ε_t^s. Lagged $Hawk_t$ controls: Baseline controls augmented by four lags of $Hawk_t$ in levels, and instrumented accordingly.

Table G.1: Cumulative 4-year government spending multipliers, robustness

	Multipliers across regimes			p-values for differences across regimes		
Specification	+2 Hawks	Average	+2 Doves	+2 Hawks vs. +2 Doves	+2 Hawks vs. Average	Average vs. +2 Doves
RZ shock (baseline)	-1.723	1.316	3.077	0.114	0.233	0.039
	(2.515)	(0.466)	(1.135)			
RZ shock (winsorized)	-1.166	1.221	2.514	0.217	0.333	0.081
	(2.481)	(0.472)	(0.884)			
RZ shock (only Q1/Q2)	-4.574	0.499	2.306	0.275	0.392	0.007
	(6.082)	(0.570)	(0.600)			
BP shock	0.861	1.349	1.731	0.084	0.098	0.070
	(1.077)	(0.842)	(0.697)			
Aggregation schemes						
Median	0.495	1.462	2.257	0.039	0.056	0.035
	(0.538)	(0.543)	(0.808)			
Chair weight	-1.594	1.532	3.445	0.067	0.167	0.075
	(2.126)	(0.649)	(1.487)			
Swinger weight	-1.521	1.272	3.021	0.085	0.169	0.042
	(2.068)	(0.543)	(1.114)			
Accounting for trends						
5-year MA	-12.330	1.239	3.661	0.816	0.841	0.217
	(68.128)	(1.176)	(2.022)			
10-year MA	-4.888	0.836	3.188	0.454	0.562	0.086
	(10.539)	(0.876)	(1.193)			
15-year MA	-2.058	0.982	2.929	0.133	0.271	0.033
	(2.967)	(0.423)	(0.849)			
Accounting for swings in the IV						
8-quarter lag	-1.396	1.276	2.661	0.207	0.319	0.065
	(2.711)	(0.528)	(1.058)			
16-quarter lag	-0.759	1.222	2.496	0.341	0.446	0.243
	(2.677)	(0.592)	(1.415)			
Average preferences	-1.481	1.065	1.878	0.548	0.580	0.491
	(4.813)	(0.650)	(1.216)			

(Table continues on the next page)

Table G.1 (continued): Cumulative 4-year government spending multipliers, Robustness

	Multipliers across regimes			p-values for differences across regimes		
Specification	+2 Hawks	Average	+2 Doves	+2 Hawks vs. +2 Doves	+2 Hawks vs. Average	Average vs. +2 Doves
Residualized IV						
Shocks	-1.729 (2.498)	1.337 (0.463)	3.084 (1.116)	0.110	0.232	0.038
Macro indicators	-2.004 (2.673)	1.290 (0.459)	3.463 (1.261)	0.080	0.237	0.043
All	-1.896 (2.487)	1.306 (0.464)	3.478 (1.255)	0.070	0.216	0.039
Accounting for the ZLB						
End sample ’08	-1.941 (2.559)	1.308 (0.503)	3.070 (1.105)	0.100	0.213	0.030
End sample ’07	-3.117 (4.115)	0.937 (0.524)	2.984 (1.105)	0.188	0.326	0.031
Additional controls						
Interest rates	0.389 (1.244)	1.263 (0.682)	1.863 (0.736)	0.289	0.304	0.345
Interest rates, inflation	0.719 (0.877)	1.260 (0.641)	2.047 (0.989)	0.307	0.288	0.360
Interest rates, inflation, surplus	0.170 (0.815)	0.772 (0.557)	1.645 (0.804)	0.191	0.244	0.220
Non-linear controls						
in t	-1.017 (2.641)	1.428 (0.555)	2.801 (1.118)	0.210	0.358	0.068
in $t, ..., t-4$	0.285 (2.382)	2.021 (0.636)	3.051 (1.184)	0.345	0.483	0.149
Lagged $Hawk_t$	-0.587 (1.718)	1.677 (0.536)	3.241 (1.210)	0.112	0.220	0.068

Notes: The table shows IV estimates of the cumulative fiscal spending multipliers $FM^H(\chi)$ in equation (4.1) for $H = 16$ quarters. The last three columns show p-values corresponding to statistical tests for whether the fiscal multiplier is significantly different across monetary regimes ($Hawk_t$). The baseline coefficients are estimated using a cumulative version of the local projection framework (3.5)-(3.6) as specified in Section 4.1. The columns present different states of the Hawk-Dove balance between “+2 Hawks” ($\chi = +2/12$), “Average” ($\chi = 0$), and “+2 Doves” ($\chi = -2/12$). Driscoll-Kraay standard errors are in parenthesis, see Appendix **F** for details. The various exercises correspond to the impulse responses presented in Figures **G.1**-**G.10**, see the respective figure notes for details.

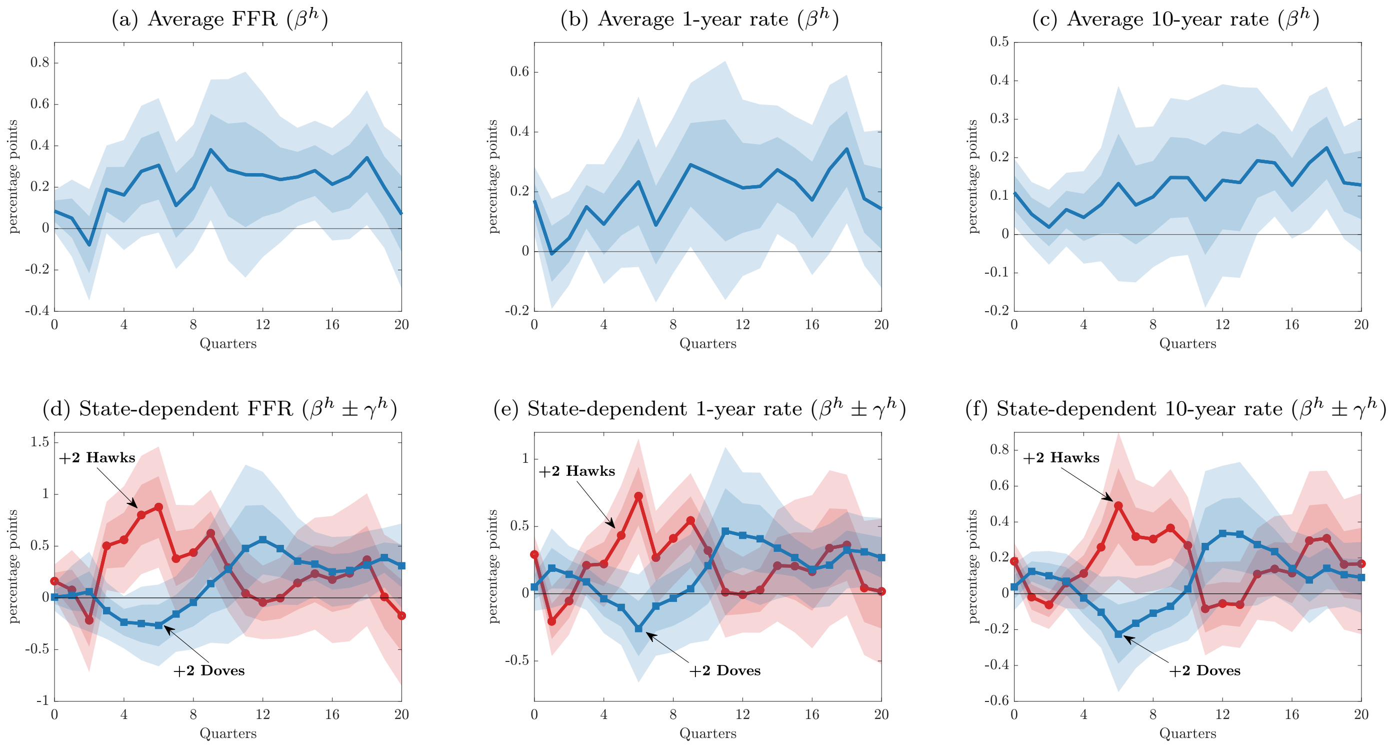
Table G.2: Cumulative 4-year government spending multipliers, Discrete Hawk-Dove balance

	Multipliers across regimes			p-values for differences across regimes		
Specification	Hawkish	Average	Dovish	Hawkish vs. Dovish	Hawkish vs. Average	Average vs. Dovish
Quartiles	-5.193 (8.557)	1.762 (0.762)	4.817 (2.787)	0.214	0.423	0.209
Tertiles	-3.262 (5.713)	0.524 (0.759)	2.834 (1.068)	0.314	0.471	0.043

Notes: The table shows IV estimates of the cumulative fiscal spending multipliers $FM^H(\chi)$ in equation (4.1) for $H = 16$ quarters. The last three columns show p-values to statistical tests for whether the fiscal multiplier is significantly different across monetary regimes ($Hawk_t$). The coefficients are estimated using a cumulative version of the local projection framework (3.5)-(3.6) as specified in Section 4.1. We use two discrete variants of $Hawk_t$. We define that the discrete $Hawk_t$ equals -1 if $Hawk_t$ falls below the first quartile or tertile of the distribution of $Hawk_t$ over time, +1 if above the highest quartile or tertile, and zero else. The columns present different states of the Hawk-Dove balance between "Hawkish" (χ within the last quartile or tertile), "Average" (χ between the first and last quartile or tertile) "Dovish" (χ within the first quartile or tertile).

Appendix H Additional results for Section 5

Figure H.1: Responses of nominal interest rates, omit shocks at end of rotation cycle

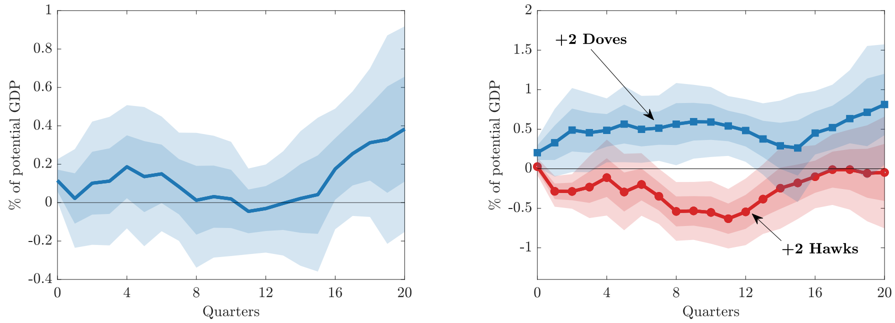


Notes: We set the military spending shocks occurring in Q3 or Q4 to zero.

Figure H.2: Responses of GDP and government spending, omit shocks at end of rotation cycle

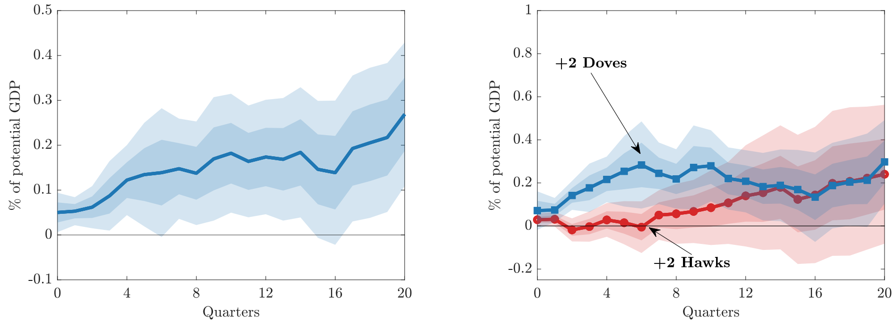
(a) Average GDP (β^h)

(b) State-dependent GDP ($\beta^h \pm \gamma^h$)



(c) Average G (β^h)

(d) State-dependent G ($\beta^h \pm \gamma^h$)

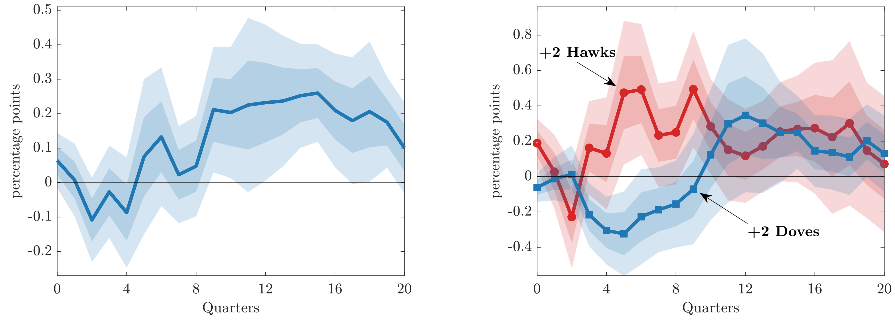


Notes: We set the military spending shocks occurring in Q3 or Q4 to zero.

Figure H.3: Responses of real interest rates

(a) Average real FFR (β^h)

(b) State-dependent real FFR ($\beta^h \pm \gamma^h$)



Notes: We compute the real federal funds rate by as nominal rate minus one-year ahead inflation expectations according to the Livingston Survey.

Appendix I Details on Calibrated NK Model

Setup and equilibrium. We consider a behavioral New Keynesian model populated by two types of households. A fraction λ^H of households are hand-to-mouth (indexed by H), and the remainder are Ricardian (indexed by R). Ricardian households choose consumption C_t^R, labor N_t^R, and savings in a risk-free, zero-coupon bond B_t^R to maximize $\tilde{E}_0 \sum_{t=0}^{\infty} \beta^t \Big(\frac{(C_t^R)^{1-\sigma}-1}{1-\sigma} - \frac{(N_t^R)^{1+\varphi}}{1+\varphi} \Big)$, subject to a sequence of budget constraints $P_t C_t^R + Q_t B_t^R =$

$B_{t-1}^R + W_t N_t^R + D_t^R + T_t^R$, where P_t is the aggregate price level, Q_t the bond price, W_t the nominal wage, D_t dividends, and T_t^R a lump-sum tax. The log-linearized first-order conditions are $c_t^R = \tilde{E}_t c_{t+1}^R - \frac{1}{\sigma}(i_t - \tilde{E}_t \pi_{t+1})$, and $\varphi n_t^R + \sigma c_t^R = w_t - p_t$, where lower case letters denote log deviations from steady state. The expectation operator $\tilde{E}_t$ entails cognitive discounting. Given some variable x_t and its steady state value x, we define $\tilde{E}_t[x_{t+1}] = x + \bar{m} E_t[x_{t+1} - x]$, where $\bar{m} \in [0,1]$ is the cognitive discounting parameter and E_t the rational expectations operator (Gabaix, 2020). The hand-to-mouth agents do not have access to bonds or shares and consume their income every period $c_t^H = y_t^H = \chi^H y_t$, with χ^H depending on transfers between household types (Bilbiie, 2020). These agents choose labor to maximize the same lifetime utility function as Ricardians, yielding an analogous intratemporal optimality condition $\varphi n_t^H + \sigma c_t^H = w_t - p_t$. Given a full-insurance steady state, aggregate consumption and labor supply are given by $c_t = (1 - \lambda^H) c_t^R + \lambda^H c_t^H$ and $n_t = (1 - \lambda^H) n_t^R + \lambda^H n_t^H$. The firm side of the economy is identical to the stylized model in Section 2 except that we allow for decreasing returns to scale. Intermediate good firms produce variety goods using the Cobb-Douglas technology $Y_{it} = N_{it}^{\alpha}$ with $\alpha \in (0,1]$ and set prices subject to a Calvo friction with reset probability $1 - \theta$. The variety goods are aggregated into final goods Y_t using a CES aggregator with elasticity ϵ.

We model fiscal policy analogous to the stylized model in Section 2 but allowing government spending to follow an $AR(p)$ process. Goods market clearing implies $y_t = (1 - \gamma) c_t + \gamma g_t$. The equilibrium dynamics are described by a dynamic IS equation $y_t = \tilde{E}_t[y_{t+1}] - \psi_r\left(i_t - \tilde{E}_t[\pi_{t+1}]\right) + \psi_g\left(g_t - \tilde{E}[g_{t+1}]\right)$, with $\psi_r = \frac{(1-\lambda^H)(1-\gamma)}{(1-\lambda^H \chi^H (1-\gamma))\sigma}$ and $\psi_g = \frac{\gamma}{1-\lambda^H \chi^H (1-\gamma)}$, and a New Keynesian Phillips Curve (NKPC) $\pi_t = \beta \tilde{E}_t[\pi_{t+1}] + \omega_y y_t - \omega_g g_t$, with $\omega_y = \frac{(1-\theta)(1-\beta\theta)}{\theta} \frac{1}{1+\varepsilon\frac{1-\alpha}{\alpha}} \left(\varphi + \frac{\sigma}{1-\gamma}\right)$ and $\omega_g = \frac{(1-\theta)(1-\beta\theta)}{\theta} \frac{1}{1+\varepsilon\frac{1-\alpha}{\alpha}} \frac{\gamma\sigma}{1-\gamma}$.

Finally, the monetary authority sets the nominal interest rate according to $i_t = \rho_i(L) i_{t-1} + (1 - \rho_i(1))[(\phi_\pi + \tilde{\phi}_{\pi,t})\pi_t + \phi_y y_t]$, with the lag polynomial $\rho_i(L) = \sum_{k=1}^{\infty} \rho_{ik} L^{k-1}$ capturing interest rate smoothing, and the inflation response (systematic monetary policy) varying over time according to $\tilde{\phi}_{\pi,t} = \rho_\phi \tilde{\phi}_{\pi,t-1} + v_t$. Monetary policy responds to both output and inflation, but only the inflation response varies over time. This specification is parsimonious and supported by the empirical evidence (Figure 3).[51]

Calibration and quantitative findings. A period in the model is a quarter. We calibrate the autoregressive process for government spending to match the estimated response of government spending to the spending shock given an average hawkishness of the FOMC,

---

[51] For simplicity, we have omitted endogenous fluctuations in systematic monetary policy. Comparing the model with our causal evidence makes this omission inconsequential.

as shown in Figure 4 (b).[52] The minimal AR(p) process required to match the government spending response has $p = 4$ lags. Figure I.1 (b) below shows that an AR(4) process closely matches the (average) government spending response. We next calibrate the monetary policy rule. Interest rate smoothing with two lags (e.g., as in Coibion and Gorodnichenko, 2011) allows the model to match the empirical interest rate response (Figure 7) substantially better compared to allowing only for one (non-zero) lag coefficient in $\rho_i(L)$. The calibration of ρ_{i1} and ρ_{i2} is discussed below. To calibrate the persistence of shocks to systematic monetary policy, we target the empirical autocorrelation $\text{corr}(Hawk_t^{\mathcal{R}}, Hawk_{t-p}^{\mathcal{R}})$ for $p = 1, \dots, 4$. An AR(1) coefficient of $\rho_\phi = 0.75$ matches these four moments relatively well.[53]

We adopt standard values for $\beta = 0.99$, $\alpha = 0.65$, and $\theta = 0.75$. We set $\gamma = 0.20$ to match the ratio of government spending over GDP in our sample. We follow Pfäuti and Seyrich (2026) and set $\lambda^H = 0.33$ and $\chi^H = 1.35$. The remaining model parameters are: $\{\bar{m}, \varphi, \sigma, \rho_{i1}, \rho_{i2}, \phi_\pi, \phi_y, \epsilon\}$. We calibrate these parameters to match the responses of real GDP, the federal funds rate, and the inflation rate in the "+2 Hawks" and the "+2 Doves" scenarios in Figures 4 (e), 7 (b), and 8 (d). In the model, we replicate the two scenarios via a shock to government spending (g_t) that coincides with a shock v_t (to the inflation coefficient). The latter shock is positive for the "+2 Hawks" scenario and negative (but of equal magnitude) for the "+2 Doves" scenario. The size of the shock is also calibrated. We solve the model using perfect foresight and replicate the empirical moments in the model by imposing the structure of the local projections on the simulated model output. We obtain the following parameters: $\bar{m} = 0.75$, $\varphi = 3.38$, $\sigma = 1.77$, $\rho_{i1} = 1.63$, $\rho_{i2} = -0.65$, $\phi_\pi = 2.49$, $\phi_y = 0.11$, $\epsilon = 17.48$. The first calibrated parameter is a corner solution. We impose a lower bound of cognitive discounting of 0.75, which is conservative in the sense that multiple empirical estimates fall below 0.75 (Pfäuti and Seyrich, 2026). The calibrated parameters for φ, ϕ_π, ϕ_y, and ϵ are all well within the range of common parameterizations. Finally, the size of the shock v_t is calibrated to 18.60. Interest rate smoothing implies that the effect of this shock on interest rates is substantially dampened.

Figure I.1 shows that the calibrated model replicates well the empirical responses under the hawkish and dovish FOMC scenarios. The GDP responses are closely matched at all horizons. The model captures an initially lower federal funds rate in the dovish scenario and a rise in the hawkish case. It exhibits overshooting in the dovish scenario, albeit with a delay of about 20 quarters. The model broadly replicates the pattern of the inflation responses, though it does not explain the delayed inflation effect (within the first 20 quarters) Overall,

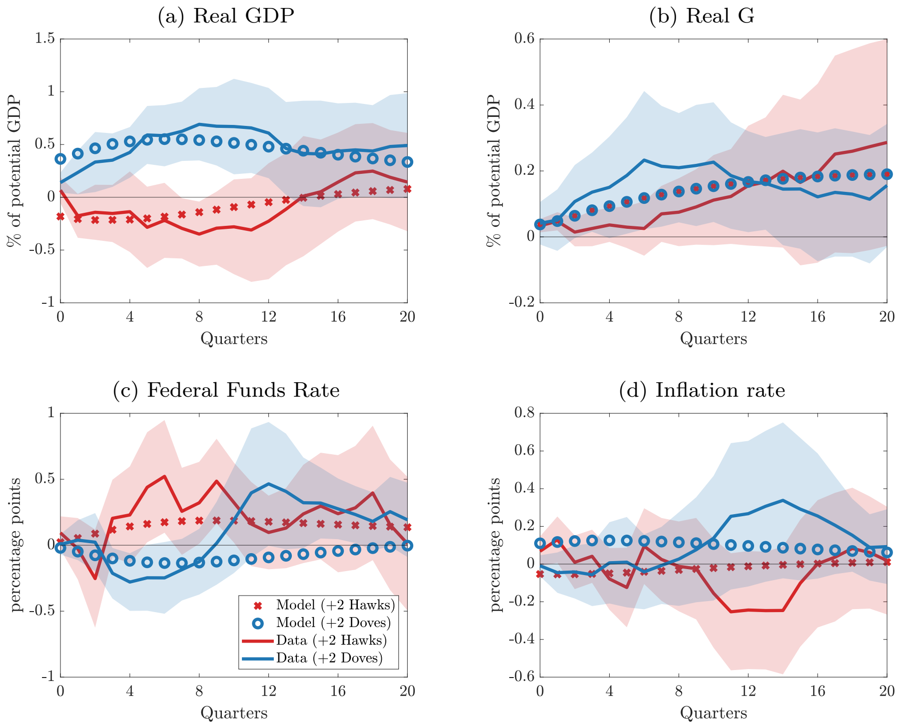
---

[52] We abstract from different government spending responses as a function of systematic monetary policy. The assumption is supported empirically and renders the quantitative analysis more transparent.

[53] The four empirical autocorrelations are given by 0.78, 0.58, 0.40, and 0.22.

our quantitative analysis shows that short-lived changes in systematic monetary policy (with a half-life of only 2.5 quarters) explain persistent differences in the responses of interest rates, GDP, and inflation to a government spending shock.

Figure I.1: State-dependent responses in a calibrated New-Keynesian model



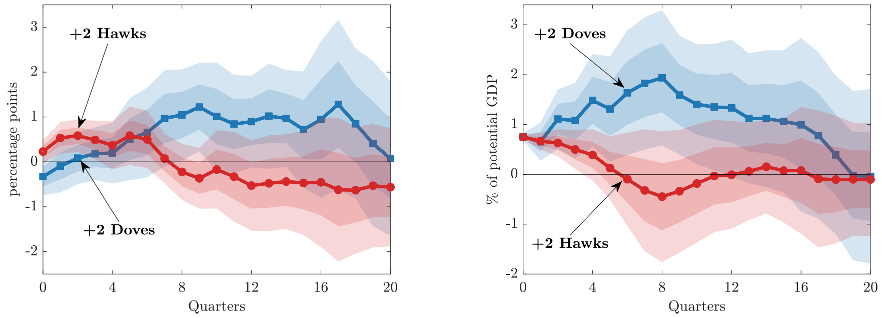
Notes: The solid lines and shaded areas replicate the point estimates and 95% confidence bands in Figures 4 (e) and (f), 7 (b), and 8 (d). Crosses and circles indicate the corresponding responses in the calibrated model.

Appendix J Beyond government spending shocks

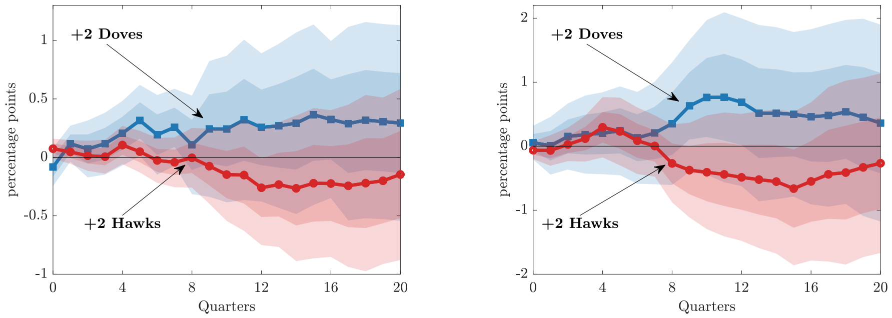
We apply our identification design to two macroeconomic shocks that are not government spending shocks. The first shock is the main business cycle (MBC) shock based on Angeletos et al. (2020), constructed as the shock that maximizes the variance of fluctuations in real GDP per capita at frequencies between 6 and 32 quarters. This shock appears as a demand shock; Angeletos et al. (2020) show that the shock resembles the confidence shock in Angeletos et al. (2018). We also consider the TFP residual from Fernald (2014) as a productivity shock. We estimate responses to these two shocks, as a function of systematic monetary policy, by applying the identification design from Section 3. We consider the responses of the federal funds rate, real GDP relative to its potential, headline inflation based on the CPI, and inflation expectations from the Livingston survey. The control vector includes 2 lags of these variables, the unemployment, the 1-year treasury rate, and the shock under consideration.

Additionally, we interact the control vector with $Hawk_t$ and instrument these regressors accordingly using $Hawk_t^{\mathcal{R}}$.[54]

Figure J.1: Responses to a main business cycle shock conditional on monetary policy


(a) Federal funds rate

(b) Real GDP


(c) Expected inflation

(d) Headline inflation

Notes: The figure shows responses to a main business cycle shock, conditional on systematic monetary policy ($Hawk_t$). We show IV estimates based on the local projection framework (3.5)-(3.6) as specified in Appendix J. The responses show $\beta^h \pm \gamma^h$ from (3.5) when $Hawk_t$ exceeds the sample average either by two hawks (+2 Hawks) or by two doves (+2 Doves). The shaded areas indicate 68% and 95% confidence bands using Newey-West standard errors.

Figure J.1 provides the results for the MBC shock. We show responses for the "+2 Hawks" and "+2 Doves" scenarios. In the hawkish scenario, the FOMC responds with an immediate interest rate increase to an expansionary shock, which is highly statistically significant for multiple quarters, see panel (a). In contrast, in the dovish scenario, the interest rate response appears delayed and becomes significant only after around two years. This finding is broadly in line with the interest rate responses to government spending shocks shown in Figure 7: in response to both shocks, a dovish FOMC tends to delay the interest rate hike before catching up with higher rates, potentially to reign in the larger economic expansion due to the initially delayed response of monetary policy. Consistent with the differential interest rate responses, we find strong state dependence in the responses of real GDP, with a hawkish FOMC strongly dampening the expansionary shock, see panel (b). Finally, a hawkish FOMC

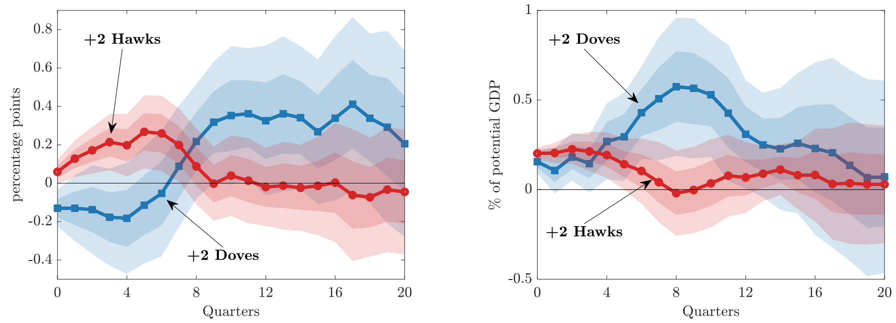
---

[54] We control for the interaction because both shocks are originally computed from linear models, leading to possible misidentification due to time variation in systematic monetary policy. We discuss this problem in more detail in Hack et al. (2024) in the context of monetary policy shock identification.

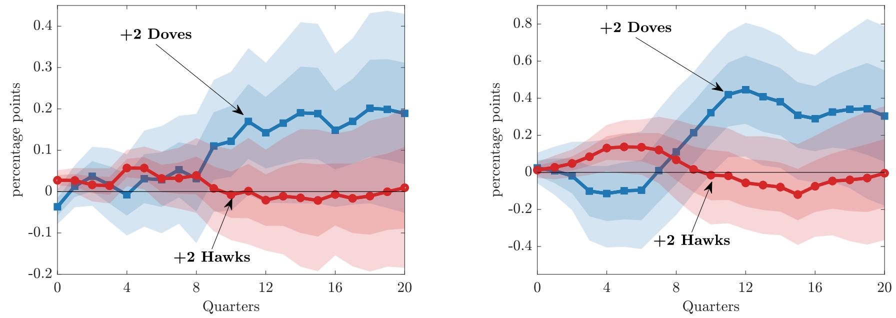
appears to be more successful in containing inflationary pressure, as shown in panels (c) and (d), although these effects are imprecisely measured.

Figure J.2: Responses to a TFP shock conditional on monetary policy

(a) Federal funds rate (b) Real GDP



(c) Expected inflation (d) Headline inflation



Notes: The figure shows responses to a TFP shock, conditional on systematic monetary policy ($Hawk_t$). We show IV estimates based on the local projection framework (3.5)-(3.6) as specified in Appendix J. The responses show $\beta^h \pm \gamma^h$ from (3.5) when $Hawk_t$ exceeds the sample average either by two hawks (+2 Hawks) or by two doves (+2 Doves). The shaded areas indicate 68% and 95% confidence bands using Newey-West standard errors.

The results for the TFP shock are presented in Figure J.2. The shock raises productivity leading to an expansion in real activity, see panel (b). However, the expansion is considerably dampened under a hawkish FOMC, which raises interest rates swiftly, see panel (a). In contrast, a dovish FOMC again delays the interest rate hike for about two years. Lastly, in panels (c) and (d), we find that a hawkish FOMC is more successful in stabilizing (expected) inflation, though the confidence bands are relatively wide.[55]
Overall, our results align well with our (main) evidence for government spending shocks.

Appendix K Two case studies from FOMC records

The U.S. Space Program. In the first half of 1961, Ramey and Zubairy (2018) identify two expansionary shocks related to President Kennedy's defense spending plans, including

---

[55] Note that the expansionary TFP shock tends to be inflationary, suggesting it contains a demand component. That may explain the initial FFR increase under the hawkish FOMC.

the Space Program to "go to the Moon". In the FOMC meeting of August 1, 1961, the staff presents the following assessment:

> *On top of substantial increases in expenditures to finance space exploration and longer-run defense measures [...] the President has found it necessary to recommend an increase of $3-1/2 billion in current defense expenditures [...]. More important, the President accompanied his recommendations with a very firm statement regarding his intentions with respect to the 1963 budget. These factors have certainly tended to minimize the immediate inflationary expectations and the urgency of the need for counter-measures. As of this moment in time, actual developments do not seem to call for any change in monetary policy.* (p.8)

The majority of the FOMC members argued similarly for no change in policy because the effects could not yet be evaluated. Hawkish FOMC members suggested the need for alertness to avoid getting into an inflationary situation while agreeing to no policy change in this meeting. In this regard, New York Fed first-vice president, William Treiber noted: *If expenditures and related private spending result in an upsurge of activity with inflationary aspects, we may have to modify our policy of basic monetary ease sooner than we would otherwise have done. In the coming period undue ease should be avoided.* (p.22-23)
FOMC members started to acknowledge the expansionary impact on employment and business sentiment in defense-related industries by the end of 1961 and later in 1963 on prices. On May 7, 1963, the FOMC voted to firm policy as a preemptive move against inflation.[56] In this meeting, Chairman Martin said:

> *If the Committee waited too long, however, it might have to deal with an active problem of inflationary pressures. In his opinion, there was already a good bit of pressure in some areas that could build up rapidly.* (p.61)

The FOMC composition was hawkish on average in this period, helping Chairman Martin build consensus for preemptive tightening against inflationary pressures.

The Vietnam War. In 1965, the U.S. entered the ground war in Vietnam leading to a series of expansions in military spending lasting until 1967Q1. In the FOMC meeting of August 10, 1965, the staff's presentation explicitly accounted for the intended increase of military spending:

---

[56] The FOMC shifted the emphasis of monetary policy toward slightly less ease and toward maintaining a moderately firm tone in the money market in June 1962, mentioning balance-of-payments concerns. In this period, FOMC members interested in a tighter, inflation-focused monetary policy often cited the balance-of-payments criterion to bolster their case (Bordo and Humpage, 2014).

> *Further stimulus to the economy will come from expanded Government procurement for Vietnam hostilities. [...] the increases in spending and in the armed forces now proposed do not appear significant enough to touch off [...] widespread price increases. [...] The market response to Vietnam developments doesn't suggest any widespread fears of shortages, rationing, or inflation. On balance, then, the domestic evidence isn't clear enough to me to justify a significant policy move in either direction at this juncture.* (p. 28-29).

Several FOMC members agreed with the staff's assessment and argued for an unchanged policy due to significant uncertainties related to the developments in Vietnam. In contrast, few hawkish FOMC members noted that the Vietnam hostilities were already affecting industrial prices. Two meetings later, on September 28, the dovish members dissented against the "status quo", arguing that evidence of inflationary pressure was lacking and hence they preferred an easier policy. In contrast, Alfred Hayes (New York Fed), a hawk, argued in the meeting of October 12, 1965 that: *Looking ahead, I think we have a real basis for concern about potential inflationary pressures* (p.25). Chairman Martin shared similar thinking on inflation while sensing that he did not have a majority to firm policy:

> *While the evidence was not clear, he thought there were many signs of inflation and of inflationary psychology in the economy. [...] But the Committee had a tendency to feel that it was best to wait until all the evidence was in before making a policy change. The difficulty was that when all the evidence was in it was likely to be too late. [...] With a divided Committee and in face of strong Administration opposition he did not believe it would be appropriate for him to lend his support to those who favored a change in policy now.* (p.68-69)

On December 5, 1965, the discount rate was raised with a narrow majority. However, the tightening was not enough to contain the buildup of inflationary pressures. While this had become clear for most members, the U.S. President had promised an anti-inflationary fiscal program and the FOMC delayed action in support of promised fiscal restraint. On September 13, 1966, Governor James Robertson summarized the situation as follows: *Inflationary pressures are persisting, as the staff materials have underlined. [...] To counter these inflationary pressures, we now have the promise of help from a somewhat greater degree of fiscal restraint.* (p.72). Hoping on legislative action to raise taxes in 1967, by the last quarter of 1966 and throughout the first part of 1967, the FOMC eased policy, despite two large expansionary military spending shocks in 1966Q4 and 1967Q1. In the FOMC meeting of September 12, 1967, Chairman Martin acknowledged that tightening had been delayed for too long because of the tendency to *underestimate the strains being put on economic resources by the hostilities in Vietnam. A "guns and butter" economy was not feasible; the country's resources were not sufficient for that.* (p.73). The FOMC decided to tighten on December 12, 1967.

In the period between 1965 and 1967, the FOMC is categorized as dovish. The dovish committee and the political pressure against tighter policy made it more difficult for Chairman Martin to reach a consensus for firm policy within the FOMC. Indeed, we observe that even when the expansionary effects of military spending related to the Vietnam War became evident, the FOMC initially hesitated, then tightened only modestly.

Appendix References

Anderson, T. W. and H. Rubin (1949): "Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations," *Annals of Mathematical Statistics*, 20, 46–63.

Andrews, I. (2018): "Valid Two-Step Identification-Robust Confidence Sets for GMM," *Review of Economics and Statistics*, 100, 337–348.

Andrews, I., J. H. Stock, and L. Sun (2019): "Weak Instruments in Instrumental Variables Regression: Theory and Practice," *Annual Review of Economics*, 11, 727–753.

Angeletos, G.-M., F. Collard, and H. Dellas (2018): "Quantifying Confidence," *Econometrica*, 86, 1689–1726.

——— (2020): "Business-Cycle Anatomy," *American Economic Review*, 110, 3030–3070.

Baker, S. R., N. Bloom, and S. J. Davis (2016): "Measuring Economic Policy Uncertainty," *Quarterly Journal of Economics*, 131, 1593–1636.

Barsky, R. and E. Sims (2011): "News Shocks and Business Cycles," *Journal of Monetary Economics*, 58, 273–289.

Baumeister, C. and P. Guérin (2021): "A Comparison of Monthly Global Indicators for Forecasting Growth," *International Journal of Forecasting*, 37, 1276–1295.

Baumeister, C. and J. D. Hamilton (2019): "Structural Interpretation of Vector Autoregressions With Incomplete Identification: Revisiting the Role of Oil Supply and Demand Shocks," *American Economic Review*, 109, 1873–1910.

Beaudry, P. and F. Portier (2014): "News-Driven Business Cycles: Insights and Challenges," *Journal of Economic Literature*, 52, 993–1074.

Ben-Zeev, N. and H. Khan (2015): "Investment-Specific News Shocks and US Business Cycles," *Journal of Money, Credit and Banking*, 47, 1443–1464.

Bilbiie, F. O. (2020): "The New Keynesian cross," *Journal of Monetary Economics*, 114, 90–108.

Blanchard, O. and R. Perotti (2002): "An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output," *Quarterly Journal of Economics*, 117, 1329–1368.

Bordo, M. D. and O. F. Humpage (2014): "Federal Reserve policy and Bretton Woods," NBER Working Paper 20656.

Caldara, D., C. Fuentes-Albero, S. Gilchrist, and E. Zakrajšek (2016): "The

Macroeconomic Impact of Financial and Uncertainty Shocks," *European Economic Review*, 88, 185–207.

Caldara, D. and M. Iacoviello (2022): "Measuring Geopolitical Risk," *American Economic Review*, 112, 1194–1225.

Caldara, D., M. Iacoviello, P. Molligo, A. Prestipino, and A. Raffo (2020): "The Economic Effects of Trade Policy Uncertainty," *Journal of Monetary Economics*, 109, 38–59.

Chaudhuri, S. and E. Zivot (2011): "A new method of projection-based inference in GMM with weakly identified nuisance parameters," *Journal of Econometrics*, 164, 239–251.

Cochrane, J. H. (2022): "The fiscal roots of inflation," *Review of Economic Dynamics*, 45, 22–40.

Coibion, O. and Y. Gorodnichenko (2011): "Monetary Policy, Trend Inflation, and the Great Moderation: An Alternative Interpretation," *American Economic Review*, 101, 341–70.

Driscoll, J. C. and A. C. Kraay (1998): "Consistent covariance matrix estimation with spatially dependent panel data," *Review of Economics and Statistics*, 80, 549–560.

Fernald, J. G. (2014): "A Quarterly, Utilization-Adjusted Series on Total Factor Productivity," Working Paper 2012-19, Federal Reserve Bank of San Francisco.

Gabaix, X. (2020): "A Behavioral New Keynesian Model," *American Economic Review*, 110, 2271–2327.

Gilchrist, S. and E. Zakrajšek (2012): "Credit Spreads and Business Cycle Fluctuations," *American Economic Review*, 102, 1692–1720.

Hack, L., K. Istrefi, and M. Meier (2024): "The Systematic Origins of Monetary Policy Shocks," CEPR Discussion Paper 19063.

Istrefi, K. (2019): "In Fed Watchers' Eyes: Hawks, Doves and Monetary Policy," Banque de France Working Paper.

Jurado, K., S. C. Ludvigson, and S. Ng (2015): "Measuring Uncertainty," *American Economic Review*, 105, 1177–1216.

Känzig, D. R. (2021): "The Macroeconomic Effects of Oil Supply News: Evidence from OPEC Announcements," *American Economic Review*, 111, 1092–1125.

Liu, Y. and J. C. Wu (2021): "Reconstructing the yield curve," *Journal of Financial Economics*, 142, 1395–1425.

Mertens, K. and M. O. Ravn (2013): "The dynamic effects of personal and corporate income tax changes in the United States," *American Economic Review*, 103, 1212–47.

Pfäuti, O. and F. Seyrich (2026): "A Behavioral Heterogeneous Agent New Keynesian Model," Working Paper 2026-01, EMPCT Working Paper Series.

Ramey, V. A. (2011): "Identifying Government Spending Shocks: It's all in the Timing," *Quarterly Journal of Economics*, 126, 1–50.

Ramey, V. A. and S. Zubairy (2018): "Government Spending Multipliers in Good Times and in Bad: Evidence from US Historical Data," *Journal of Political Economy*, 126, 850–901.

Romer, C. D. and D. H. Romer (2010): "The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks," *American Economic Review*, 100, 763–801.