

Identification of Systematic Monetary Policy*

Lukas Hack

Klodiana Istrefi

Matthias Meier

This version: January 2024

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 Francesco Fusari, Anna Matzner, Karel Mertens, Alexander Meyer-Gohde, Filippo Palotti, and Josha van Spronsen for insightful discussions, and Klaus Adam, Regis Barnichon, 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 financial support from the German Research Foundation (DFG) through CRC TR 224 (Project C05). Matthias Meier acknowledges financial support 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 Banque de France or the Eurosystem.

Lukas Hack (University of Mannheim): lukas.hack@uni-mannheim.de. Klodiana Istrefi (Banque de France, CEPR, European Central Bank): klodiana.istrefi@banque-france.fr. Matthias Meier (University of Mannheim): 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, 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 below 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 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.

¹A 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, 2022](#)), we propose to directly estimate the causal effects of systematic monetary policy.

Our measure of systematic monetary policy is the aggregate Hawk-Dove balance for each FOMC meeting.² 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; Abrams and Butkiewicz, 2012). 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 levers exogenous variation in the Hawk-Dove balance arising from the FOMC rotation of voting rights. The rotation is an annual mechanical scheme that shuffles four out of twelve voting rights among eleven Federal Reserve Bank presidents.³ We construct an FOMC rotation instrument that is the Hawk-Dove balance among the four FOMC member which the rotation assigns voting rights in a given year. Importantly, the mechanic nature of the rotation renders it orthogonal to economic and political developments. Moreover, the rotation is considered relevant by Fed watchers in the media, the correlation between rotation instrument and overall Hawk-Dove balance is 0.64, and the instrument passes multiple weak instrument tests.

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

²Istrefi (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. Instead, we use the Hawk-Dove classification to study the effects of systematic monetary policy on the propagation of macroeconomic shocks. More specifically, we construct an instrument for the Hawk-Dove balance, propose a novel identification design, and apply it to study the effects of government spending shocks.

³Relatedly, Ehrmann et al. (2022) studies how voting rights affect the communication of Federal Reserve Bank presidents and the market reaction to this communication.

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 dynamics of a New Keynesian model with time-varying systematic monetary policy. However, different from a New Keynesian model, our design identifies the effects of systematic monetary policy without imposing strong structural assumptions. Instead, we leverage historical variation in the composition of policy preferences among FOMC members. This allows us to study the effects of counterfactual Hawk-Dove balances.⁴

We apply our identification design 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 1960-2014.⁵ We find that the real GDP response depends significantly on systematic monetary policy. 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, quarterly GDP increases by up to 0.7% in response to a military spending 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.⁶ 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 to assess 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 dependency. 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 nominal interest rates rise under a hawkish FOMC. Under a dovish FOMC, nominal rates initially fall, and rise only with substantial delay. In more detail, when the

⁴This means we can study counterfactual interest rate responses that are associated with historical variation in the Hawk-Dove balance. In contrast, we cannot study counterfactual interest rate responses that did not occur in the data.

⁵In 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.

⁶For comparison, an increase of the Hawk-Dove balance by two doves or two hawks roughly corresponds to one standard deviation in the change of the Hawk-Dove balance.

Hawk-Dove balance exceeds the sample average by two hawks, the federal funds rate (FFR) starts to increase within one year after the shock, and increases by up to 50 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 50 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) and that the monetary policy response primarily transmits to real GDP through private consumption. Finally, we complement our quantitative analysis with narrative evidence from the historical records of the FOMC meetings. These 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. We show that a hawkish FOMC indeed tightens faster after military buildups, whereas a dovish FOMC delays 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 \(2022\)](#) and [Barnichon and Mesters \(2022\)](#) who use multiple monetary policy (news) shocks to estimate the effects of counterfactual monetary policy rules.⁷ Under the assumption that systematic monetary policy affects private agents only through changes in the policy instrument, their approach allows identifying the effects of a large set of counterfactual interest rate paths. Instead, our approach leverages historical variation in systematic monetary policy, which avoids potential problems related to the identification and size of monetary policy shocks. The identification of monetary policy shocks is subject to a long-running and ongoing debate (e.g., [Ramey, 2016](#); [Bauer and Swanson, 2023](#)). A key concern is that empirical monetary policy shocks may be contaminated by other business cycle shocks. In fact, one reason for contamination may be time variation in systematic monetary policy.⁸ In addition, the effects of monetary policy shocks are typically small, particularly in more recent decades ([Ramey, 2016](#)). 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

⁷[McKay and Wolf \(2022\)](#) focus on constructing policy counterfactuals, whereas [Barnichon and Mesters \(2022\)](#) uses a similar approach to study optimal policy. Relatedly, [Wolf \(2023\)](#) uses the approach of [McKay and Wolf \(2022\)](#) to provide fiscal policy shock counterfactuals for a strict inflation-targeting central bank.

⁸For example, if the rule changes over time, [Romer and Romer \(2004\)](#) and high-frequency identified shocks may be contaminated, see Appendix D.

policy shocks (e.g., [Bernanke et al., 1997](#); [Kilian and Lewis, 2011](#)).⁹ 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 to estimate the effects of time-varying systematic monetary policy uses non-linear VAR models (e.g., [Primiceri, 2005](#); [Sims and Zha, 2006](#)). A key advantage of our approach is that it requires weaker identifying assumptions and addresses the potential endogeneity of systematic monetary policy. Our paper also relates to a literature studying macroeconomic models with exogenous changes in systematic monetary policy (e.g., [Davig and Leeper, 2007](#); [Bianchi, 2013](#); [Leeper et al., 2017](#)) or endogenous changes (e.g., [Davig and Leeper, 2008](#); [Barthélemy and Marx, 2017](#)). Our time series approach requires fewer structural assumptions and provides moments to discipline such models.

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](#); [Barro and Redlick, 2011](#); [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](#); [Ben Zeev et al., 2023](#)); exchange-rate regime, trade openness, and public debt ([Ilzetzki et al., 2013](#)); foreign holdings of debt ([Broner et al., 2022](#)); and tax progressivity ([Ferrière and Navarro, 2018](#)). Compared to this literature, our analysis tackles the

⁹A further related paper on the intersection of shocks and systematic policy is [Arias et al. \(2019\)](#) which identifies monetary policy shocks via sign restrictions on systematic monetary policy.

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 provides a simple New Keynesian model to discuss the identification challenge. Section 3 introduces the identification design for systematic U.S. monetary policy. Section 4 contains the main empirical results on the effects of fiscal spending shocks. Section 5 provides evidence for understanding 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 $E_0 \sum_{t=0}^{\infty} \beta^t (\log C_t - N_t^{1+\varphi})$ 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

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

where $\tilde{\phi}_t \in (1, \infty)$ 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 $\tilde{\phi}_t = \phi + \phi_t$ and ϕ denotes the unconditional mean of $\tilde{\phi}_t$. Importantly, we allow systematic monetary policy to be endogenous, as ϕ_t may respond to macroeconomic shocks

$(\varepsilon_t^s, \varepsilon_t^a)$.¹⁰ 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 . 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

$$y_t = a + b_s x_t^s + b_a x_t^a + c_s x_t^s \phi_t + c_a x_t^a \phi_t + d \phi_t, \quad (2.3)$$

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.

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

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

for $h = 0, \dots, H$ 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 $(\varepsilon_t^s, \varepsilon_t^a)$ and policy shifter (η_t) occurring between t and $t + h$. The estimands in (2.4) are

$$\beta^h = b_s (\rho_s)^h, \quad \gamma^h = c_s (\rho_s \rho_\phi)^h, \quad \delta^h = d (\rho_\phi)^h. \quad (2.5)$$

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 .

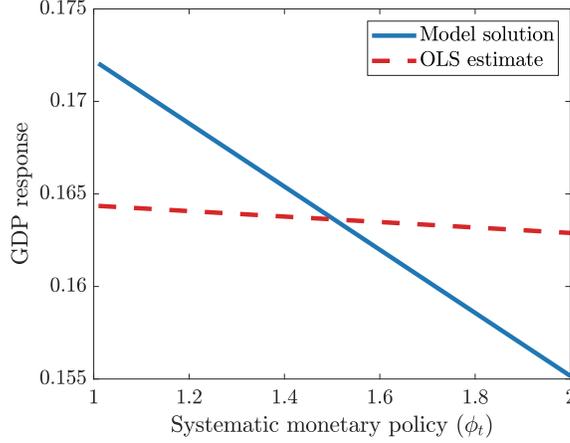
We next ask whether the OLS estimates of $(\beta^h, \gamma^h, \delta^h)$ are consistent, i.e., whether they asymptotically recover the estimands in (2.5).¹¹ Consistency holds under the strong exogeneity assumption $\zeta^s = \zeta^a = 0$, that is if ϕ_t is independent of the macroeconomic shocks. In contrast, if ϕ_t correlates with at least one of the shocks, the OLS estimates do *not* consistently estimate $(\beta^h, \gamma^h, \delta^h)$.¹² If, for example, ϕ_t responds to a spending shock, the OLS

¹⁰For 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 \(2017\)](#).

¹¹We 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.

¹²If 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 (2.3) when the terms in u_t 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 $\hat{\gamma}^0 = -0.002$.

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, present, and future macroeconomic shocks ε_t^s and ε_t^a and that is uncorrelated with all past and future policy shifters η_t (exogeneity). 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.

Illustration. To illustrate the effects of systematic monetary policy and the identification challenge, we focus on a special case of our economy in which $\rho_s = \rho_a = \rho_\phi = 0$. To understand how ϕ_t affects the GDP response to fiscal spending shock ε_t^s , we need to know

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

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.

3 Identification design

In this section, we propose 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 crucial 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 Federal Open Market Committee (FOMC) members into hawks and doves by [Istrefi \(2019\)](#) and argue that the Hawk-Dove balance captures well variation in systematic monetary policy over time.

The FOMC. The FOMC is the committee of the Federal Reserve that sets U.S. monetary policy. The FOMC consists of 12 members: the seven members of the Board of Governors of the Federal Reserve System, 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 rotating basis.¹³

Individual policy preferences. To measure the policy preferences of FOMC members we use the [Istrefi \(2019\)](#) classification of FOMC members as hawks and doves, for the period 1960-2023.¹⁴ Underlying this classification are more than 20,000 real-time media articles from over 30 newspapers and business reports of Fed watchers (available in news archives like ProQuest Historical Newspapers and Factiva) mentioning individual FOMC members. [Istrefi \(2019\)](#) uses these articles to categorize individual FOMC members as hawks or doves for each FOMC meeting based on the news information available up until the meeting. So, the Hawk-Dove classification is a panel that tracks FOMC members over time, at FOMC

¹³While non-voting FRB presidents attend the FOMC meetings and participate in the discussions, we focus on the voting FOMC, the decision-making body, in line with the literature that studies central bank decision making by committees (e.g., [Belden, 1989](#); [Blinder, 2007](#); [Riboni and Ruge-Murcia, 2010, 2022](#); [Bordo and Istrefi, 2023](#)).

¹⁴The data in [Istrefi \(2019\)](#) covers 1960 through 2014. The data is currently extended up to the first meeting of 2023. Thus, our sample covers all 634 (scheduled) FOMC meetings between 1960 and 2023.

meeting frequency. Hawks are perceived to be more concerned with inflation, while doves are more concerned with employment and growth.¹⁵ Through the lens of our model in Section 2, we can think about hawks as preferring a larger inflation coefficient ϕ_t than doves. However, the Hawk-Dove classification we use is not tied to assuming a specific policy rule.

Overall, 129 of the 147 FOMC members between 1960 and 2023 are classified as hawk or dove. The news coverage for the remaining 18 members does not allow classification (as hawk or dove) for any meeting, as some served in the early 1960s with sparse media coverage and others are very recent appointments in the FOMC. The majority (95) of the classified FOMC members are consistently hawks or doves over time while the rest switches camps at least once. Swings are equally split in either direction and quite uniformly distributed over time. On average, the 34 swinging FOMC members switch camps at only 1.8% of the member-meeting pairs.

While true policy preferences are unobserved, [Istrefi \(2019\)](#) shows that perceived preferences match well with policy tendencies that are unknown in real-time to the public, as expressed by preferred interest rates, with forecasting patterns of individual FOMC members, and with dissents. In addition, [Bordo and Istrefi \(2023\)](#) show that the FOMC members' educational background, e.g., whether they graduated from a university related to the Chicago school of economics, and early life experience, i.e., whether they grew up during the Great Depression, predicts the Hawk-Dove classification. The long lasting effect of the early life experience in the formation of policy preferences is consistent with the very few swings in our sample.

Aggregate Hawk-Dove balance. To measure variation in systematic monetary policy over time, we aggregate the cross-section of individual FOMC member preferences into an aggregate Hawk-Dove balance for each meeting (cf. [Istrefi, 2019](#)). We do so because the nature of monetary policy-making by committee involves the aggregation of diverse individual policy preferences in a collective decision.¹⁶

We adopt a symmetric numerical scale for the qualitative Hawk-Dove classification in order to aggregate the preferences. We define $Hawk_{i\tau}$ as the policy preference of FOMC member

¹⁵A 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.

¹⁶Relatedly, [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.*

i at FOMC meeting τ :

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

A consistent hawk is an FOMC member that has not been categorized as a dove in the past. In contrast, a swinging hawk has been a dove at some point in the past. The definition of a consistent dove and a swinging dove is analogous. We assign a lower weight to swingers as they are often perceived as ‘middle-of-the-roaders’ with more moderate leanings to the hawkish or dovish side (Istrefi, 2019).¹⁷ Finally, we assign $Hawk_{i\tau} = 0$ when the policy preference of the FOMC member is (yet) unknown.

We next aggregate the individual policy preferences in (3.1). We compute the aggregate Hawk-Dove balance by

$$Hawk_{\tau} = \frac{1}{|\mathcal{M}_{\tau}|} \sum_{i \in \mathcal{M}_{\tau}} Hawk_{i\tau} \quad (3.2)$$

where \mathcal{M}_{τ} denotes the set of FOMC members at meeting τ . 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.¹⁸ The Hawk-Dove balance in (3.2) is the arithmetic average across individual preferences. This is our baseline aggregation of the Hawk-Dove balance in the FOMC and conforms well with the consensual mode in which the FOMC typically operates.^{19,20} In Section 4.5, we show that our empirical findings are robust to alternatively using the median of preferences or putting a higher weight on the Fed Chair’s preference. Finally, we aggregate

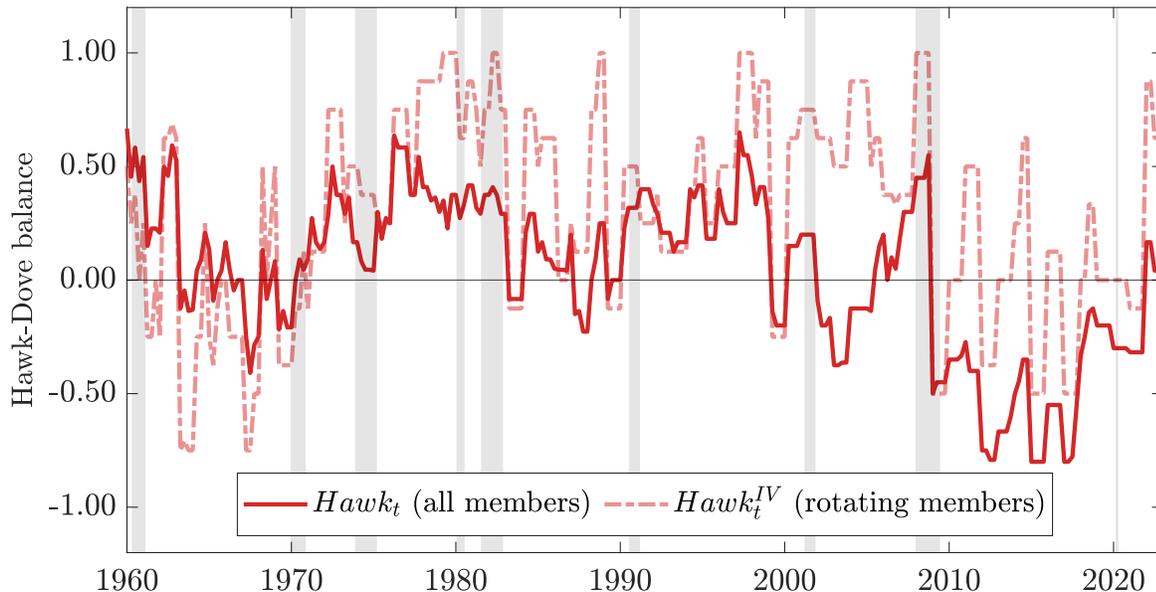
¹⁷Our empirical findings are robust to not distinguishing between consistent and swinging preferences, see Section 4.5.

¹⁸When 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 for details. This assumption affects less than one percent of all observations and is not important for our results.

¹⁹Riboni and Ruge-Murcia (2010) argue that a consensus model fits actual policy decisions of the Federal Reserve. In addition, Riboni and Ruge-Murcia (2022) provide evidence suggesting that policy proposals of the Fed Chair are the result of a compromise, reflecting a balance of power within the FOMC.

²⁰Cieslak et al. (2022) construct a Hawk-Dove score based on the language in FOMC meeting transcripts. In contrast to our measure which captures FOMC members preferences about monetary policy, their measure captures (a hawkish or dovish) sentiment on current direction of policy changes. Furthermore, Ferguson et al. (2023) classify central bank governors in 80 countries as hawks and doves, with respect to financial sector support, for the periods preceding banking crises.

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 ($Hawk_t^{IV}$). Grey bars indicate NBER dated recessions.

$Hawk_\tau$ from meeting frequency to quarterly frequency. We compute the Hawk-Dove balance $Hawk_t$ for 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 evolution 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 FOMC members, the turnover of 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 Subsection 3.2.

Systematic monetary policy. The aggregate Hawk-Dove balance $Hawk_t$ represents our measure of systematic U.S. monetary policy. It accounts for the diversity of views within the FOMC on how policy should be adjusted to promote both, price stability and maximum employment. This diversity is usually expressed in FOMC meetings through different forecasts of individual members, through dissents, and in public through speeches. While the Fed’s response to macroeconomic shocks is sophisticated and depends on various economic factors, we argue that our Hawk-Dove balance matches well with narratives of monetary policy in the U.S. (Istrefi, 2019). For example, the dovish leaning of $Hawk_t$ in the mid-1960s

coincides with a period of delays and hesitation from the FOMC to take 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; Romer and Romer, 2002) 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.²¹ During the 1980s, the perception of a 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). Overall, this suggests that the Hawk-Dove balance captures important aspects of the Fed’s systematic policy-making.

Our approach of measuring systematic policy via $Hawk_t$ has several advantages to alternative approaches such as calibrating or estimating policy rules (e.g., Clarida et al., 2000; Bauer et al., 2022). 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.²² We further avoid the well-known identification issues that plague the estimation of monetary policy rules (Cochrane, 2011; Carvalho et al., 2021). Independently 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. In addition, 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).

Comparability over time. A potential concern with the classification of FOMC members into hawks and doves is that the meaning of being a hawk or dove might have changed over time. We argue this is likely no major concern. First, Istrefi (2019) has classified each member as a hawk or dove based on a common and time-invariant definition, that is the policy leaning with regard to the dual mandate of the Fed: maximum employment and stable prices. Second, given that preferences tend to be stable, we would expect many swings after large changes in the meaning of hawks and doves. However, swings in measured preferences

²¹Moreover, Orphanides (2003) shows that a dovish Taylor rule with a sufficiently large weight on the output gap would have resulted in substantially higher inflation.

²²For a summary of alternative policy rules that the FOMC consults, see here: <https://www.federalreserve.gov/monetarypolicy/policy-rules-and-how-policymakers-use-them.htm>. Policy instruments have been changing over our sample, from targeting monetary aggregates to targeting the Fed Funds rate, conducting balance sheets policy, and through forward guidance communication.

are rare suggesting that the meaning of being a hawk or dove is relatively stable over time. Third, the fact that we observe large and persistent fluctuations in $Hawk_t$ is incompatible with the Hawk-Dove classification being a relative ranking, according to which hawks are those FOMC members which are more hawkish than the contemporaneous average policy preference among FOMC members, and analogously for doves. Finally, in a robustness exercise in Section 4, we show that our results are robust to using an alternative Hawk-Dove balance which accounts for potential trends in the meaning of hawks and doves.

Relation to monetary policy shocks. Empirically identified monetary policy shocks are often considered to reflect changes in central bank preferences (Christiano et al., 1999; Ramey, 2016). Hence, they may be related to the Hawk-Dove balance, our measure of systematic monetary policy. In Appendix D, we characterize this relationship based on a stylized Romer and Romer (2004) identification strategy of monetary policy shocks. Because their identification strategy assumes a time-invariant policy rule, the identified monetary policy shocks may indeed capture time variation in systematic monetary policy. However, the relationship between identified monetary policy shocks and systematic monetary policy is non-linear and also depends on the state of the economy (e.g., the inflation rate). Instead, our Hawk-Dove balance provides a cleaner measure of systematic monetary policy.

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, even if monetary policy is endogenous to the state of the economy (cf. Section 2).

Potential endogeneity. Systematic monetary policy may change depending on the state of the economy. 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). Empirically, Chang et al. (2021) find that the parameters of the monetary policy rule respond to macroeconomic shocks. Changes in systematic monetary policy may also be driven by political pressure. For example, Abrams (2006) and Abrams and Butkiewicz (2012) document the influence of the Nixon administration on the FOMC.²³ Political pressure may lead to endogenous fluctuations in the Hawk-Dove balance through swings of incumbent FOMC members and through new appointments. In this context, note that members of the

²³More recently, Bianchi et al. (2023) and Camous and Matveev (2021) document that President Trump exerted pressure on the Fed and Drechsel (2023) identifies the effects of political pressure on the Fed.

Board of Governors and the Fed Chair require a nomination from the U.S. President for their first and any subsequent term.

FOMC rotation instrument. To address the endogeneity of the Hawk-Dove balance we propose an instrument which 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 FOMC rotation instrument. Formally, the instrument is given by

$$Hawk_{\tau}^{IV} = \frac{1}{|\mathcal{R}_{\tau}|} \sum_{i \in \mathcal{R}_{\tau}} Hawk_{i\tau}, \quad (3.3)$$

where \mathcal{R}_{τ} denotes the set of rotating FOMC members at FOMC meeting τ . A full set of rotating members consists of $|\mathcal{R}_{\tau}| = 4$ members.²⁴ We aggregate the FOMC rotation instrument to quarterly frequency analogously to the Hawk-Dove balance.

In Figure 2, the dashed line presents the FOMC rotation instrument over time. On average, the rotating presidents are more hawkish than the overall FOMC Hawk-Dove balance, reflecting the fact that FRB presidents tend to be more hawkish than governors (Chappell et al., 2005; Istrefi, 2019; Bordo and Istrefi, 2023). Both series display sizable variation over time, but fluctuations in the instrument $Hawk_t^{IV}$ are more short-lived, with a year-over-year autocorrelation of 0.19 compared to 0.66 for $Hawk_t$, see Table 1.

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	253
$Hawk_t^{IV}$	0.28	0.33	0.45	0.19	0.64	-0.75	1.00	253

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. $Hawk_t^{IV}$ is the FOMC rotation instrument. Autocorr refers to the year-over-year autocorrelation. Corr refers to the correlation with $Hawk_t$.

²⁴In 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 $Hawk_t^{IV}$ aggregates the policy preferences of one-third of the FOMC members, capturing a significant part of the variation in the overall Hawk-Dove balance $Hawk_t$. In fact, the correlation between $Hawk_t$ and $Hawk_t^{IV}$ is 0.64. We further study the explanatory power of the FOMC rotation instrument via a stylized first-stage regression by projecting $Hawk_t$ on $Hawk_t^{IV}$ and a constant. Applying the weak instrument test from [Montiel Olea and Pflueger \(2013\)](#) yields an effective F-statistic of 46.13 which is above 37.42, the critical value for rejecting a relative weak instrument bias exceeding 5%.²⁵ 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.

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 entry and exit, see Appendix C for details. We find that extensive margin changes in the FOMC composition due to the rotation account for 53% of the variance in yearly changes of $Hawk_t$. The turnover of non-rotating FOMC members accounts for almost another quarter of the variance, and the remainder is due to preference changes of incumbent FOMC members and various covariance terms. Both the first-stage regression and the variance decomposition strongly suggest that our instrument is relevant for $Hawk_t$. 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 $Hawk_t^{IV}$ is quasi-exogenous. First, the rotation scheme is mechanical and time-invariant and therefore unrelated to the

²⁵We also reject the null of the weak instrument bias exceeding 5% when adding four lags of $Hawk_t$ and $Hawk_t^{IV}$ to control for serial correlation in both variables. In either case, we use Newey-West standard errors with automatic bandwidth selection.

state of the economy. Second, new appointments of FRB presidents are relatively infrequent and unlikely to be influenced by the federal government. FRB presidents are appointed by the Board of Directors of the respective Federal Reserve district. The directors are to represent the financial institutions and the broader public in the district.²⁶ In contrast, members of the Board of Governors (including the Fed Chair) are nominated by the U.S. president and confirmed by the Senate. Furthermore, the average tenure of an FRB president is eleven years but only seven years for a governor in our sample. Relatedly, [Bordo and Istrefi \(2023\)](#) show 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, some regional FRBs have persistent leanings toward either the dovish or the hawkish camp. For example, the Cleveland FRB president is typically a hawk whereas the president of the San Francisco FRB is typically a dove.

Third, of potential concern are swings of FRB presidents between being a hawk or dove. If swings are driven by macroeconomic shocks this will introduce endogeneity in the FOMC rotation instrument. Yet, 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-meetings pairs.²⁷ In addition, we find that swings account for a negligible fraction of the variance of the rotation instrument. In particular, we decompose $Hawk_t^{IV}$ into intensive margin changes of preferences (swings) and extensive margin changes of the composition of rotating FOMC members due to either the rotation or appointments, see Appendix C for details. The rotation accounts for 93% of the variance in yearly changes of $Hawk_t^{IV}$, appointments for 7% and swings for less than 1%. In addition, among the few swings that did happen, some do not appear linked to the state of the economy.²⁸ To address residual concerns about swings, our sensitivity analysis considers an alternative Hawk-Dove balance which mutes the effects of swings. Our results are robust to these alternatives, see Section 4.5.

Fourth, $Hawk_t^{IV}$ displays relatively short-lived time series fluctuations that are unlikely to be correlated with slow-moving macroeconomic trends, such as increasing market power, female labor force participation, and various technological innovations. Similarly, $Hawk_t^{IV}$ is uncorrelated with business cycle fluctuations. For example, the correlation between $Hawk_t^{IV}$

²⁶Formally, 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](#)).

²⁷Specifically, in 2533 member-meeting observation, we observe only 34 swings.

²⁸[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.

and yearly real GDP growth is -0.02 and statistically insignificant. In contrast, the correlation between $Hawk_t$ and GDP growth is 0.15 and significant at the 5% level.

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.

A validation exercise for $Hawk_t$ and $Hawk_t^{IV}$. Given our definition of hawkish policy makers and conventional wisdom about hawkish monetary policy, we should expect a hawkish FOMC to respond more aggressively to inflation. As validation exercise, we empirically test this correlation via a dynamic Taylor rule regression. We use $Hawk_t^{IV}$ as instrument in a local projection of the federal funds rate on the Greenbook inflation forecast interacted with $Hawk_t$. We find that a hawkish FOMC indeed raises the federal funds rate significantly more aggressively in the presence of higher inflation forecasts.²⁹ For more details on the exercise, the results, and a weak instrument test, see Appendix E. Overall, this exercise suggests that $Hawk_t$ and $Hawk_t^{IV}$ capture important variation in systematic monetary policy.

3.3 Local projection framework

Finally, we propose to combine $Hawk_t$ and $Hawk_t^{IV}$ in a state-dependent local projection framework that permits causal identification of how systematic monetary policy shapes the propagation of various macroeconomic shocks. The setup of the local projection is consistent with the New Keynesian model discussed in Section 2.

We regress an outcome variable of interest, x_{t+h} , on a macroeconomic 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,

$$x_{t+h} = \alpha^h + \beta^h \varepsilon_t^s + \gamma^h \varepsilon_t^s (Hawk_t - \overline{Hawk}) + \delta^h (Hawk_t - \overline{Hawk}) + \zeta^h Z_{t-1} + v_{t+h}^h, \quad (3.4)$$

for $h = 0, \dots, H$ forecast horizons. \overline{Hawk} denotes the arithmetic sample mean of $Hawk_t$. To address the potential endogeneity of $Hawk_t$, we use the instrument vector

$$q_t = \left[1, \varepsilon_t^s, \varepsilon_t^s (Hawk_t^{IV} - \overline{Hawk}^{IV}), (Hawk_t^{IV} - \overline{Hawk}^{IV}), Z_{t-1} \right] \quad (3.5)$$

for the regressors in (3.4). The two key coefficients in (3.4) are β^h and γ^h , which capture the

²⁹This is in line with the findings in [Bordo and Istrefi \(2023\)](#) who provide OLS estimates of a Taylor rule regression augmented by the Hawk-Dove balance.

average response, when the Hawk-Dove balance equals its sample average, and the differential response, when the FOMC is more or less hawkish than the sample average.³⁰

Based on Section 2, the IV estimator is consistent if the instrument $Hawk_t^{IV}$ 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 $Hawk_t^{IV}$) and the identifying assumptions for our analysis of government spending shocks.

Variables. We first specify the local projection framework (3.4)-(3.5). Our baseline shock of interest, ε_t^s in (3.4), 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

³⁰Formally, we define state-dependent impulse responses as

$$\mathbb{E}[x_{t+h}|\varepsilon_t^s = \varepsilon, Hawk_t = Hawk] - \mathbb{E}[x_{t+h}|\varepsilon_t^s = 0, Hawk_t = Hawk] = [\beta^h + \gamma^h (Hawk - \overline{Hawk})] \varepsilon,$$

where both expectations additionally condition on the control vector Z_{t-1} .

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 outcome variables of interest, x_{t+h} in (3.4), are real GDP and real government spending, both expressed relative to real potential GDP.³¹ Finally, the vector of control variables, Z_{t-1} in (3.4), 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 (3.4) corresponds to equation (1) of [Ramey and Zubairy \(2018\)](#). This facilitates the comparability of our results with the literature.³²

Sample. 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 \(2018\)](#).³³ In the context of studying the response of monetary policy to fiscal spending shocks, however, it may be desirable to exclude these events because monetary policy was less autonomous from fiscal policy prior to the Treasury-Fed Accord in 1951. Between 1942 and 1951, the Fed was constrained to support government bond prices by pegging short-term interest rates.

Identifying assumptions. Two key identifying assumptions are necessary for the causal interpretation of the estimates of β^h and γ^h in (3.4). The first assumption is that the FOMC rotation instrument is orthogonal to all macroeconomic shocks at all leads and lags. This is plausible for various reasons as discussed in Section 3.2. More specifically, given that fluctuations in $Hawk_t^{IV}$ are relatively short-lived and uncorrelated with real GDP growth, it is unlikely that our estimates capture differences in the response across booms and busts (e.g., [Auerbach and Gorodnichenko, 2012](#); [Ramey and Zubairy, 2018](#)).

The second assumption is that military spending shocks are random shocks. In particular,

³¹Detrending by potential GDP is the so-called [Gordon and Krenn \(2010\)](#) transformation. Compared to using log variables, this avoids using an ex-post multiplication with the GDP/G ratio, which substantially varies over time, to obtain the fiscal spending multiplier.

³²In Section 4.5 we present various sensitivity checks, including additional control variables such as lags of $Hawk_t$ or interactions of $Hawk_t$ with the control vector.

³³[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 military spending shocks are a weak instrument for contemporaneous government spending because the shocks largely pertain to future spending. Therefore, we do not use military spending shocks as an instrument but as shocks in our local projection framework (3.4) and find a significant dynamic government spending response, see Section 4.2.

the distribution of military spending shocks may not depend on systematic monetary policy. According to [Ramey and Shapiro \(1998\)](#) and [Ramey and Zubairy \(2018\)](#), military spending shocks are unanticipated changes in spending plans triggered by geopolitical events and are therefore exogenous to the economy. This argument similarly applies when conditioning on systematic monetary policy. We provide three additional arguments as to why the military spending shocks are independent of systematic monetary policy: (i) the response of military spending to the shock does not depend on systematic monetary policy; (ii) the news quotes used to construct military spending shocks as described in the supplementary appendix to [Ramey and Zubairy \(2018\)](#) do not mention monetary policy, the Federal Reserve, or the FOMC for our sample; and (iii) the Hawk-Dove balance does not predict spending shocks. The specific concern the last point addresses is that military spending shocks might be timed to episodes with a more dovish FOMC. To test this concern we regress future military spending shocks on $Hawk_t$ and use $Hawk_t^{IV}$ as an instrument. We find no significant effects of the Hawk-Dove balance on contemporaneous or future military spending shocks, see Figure F.1 in Appendix I.

4.2 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.

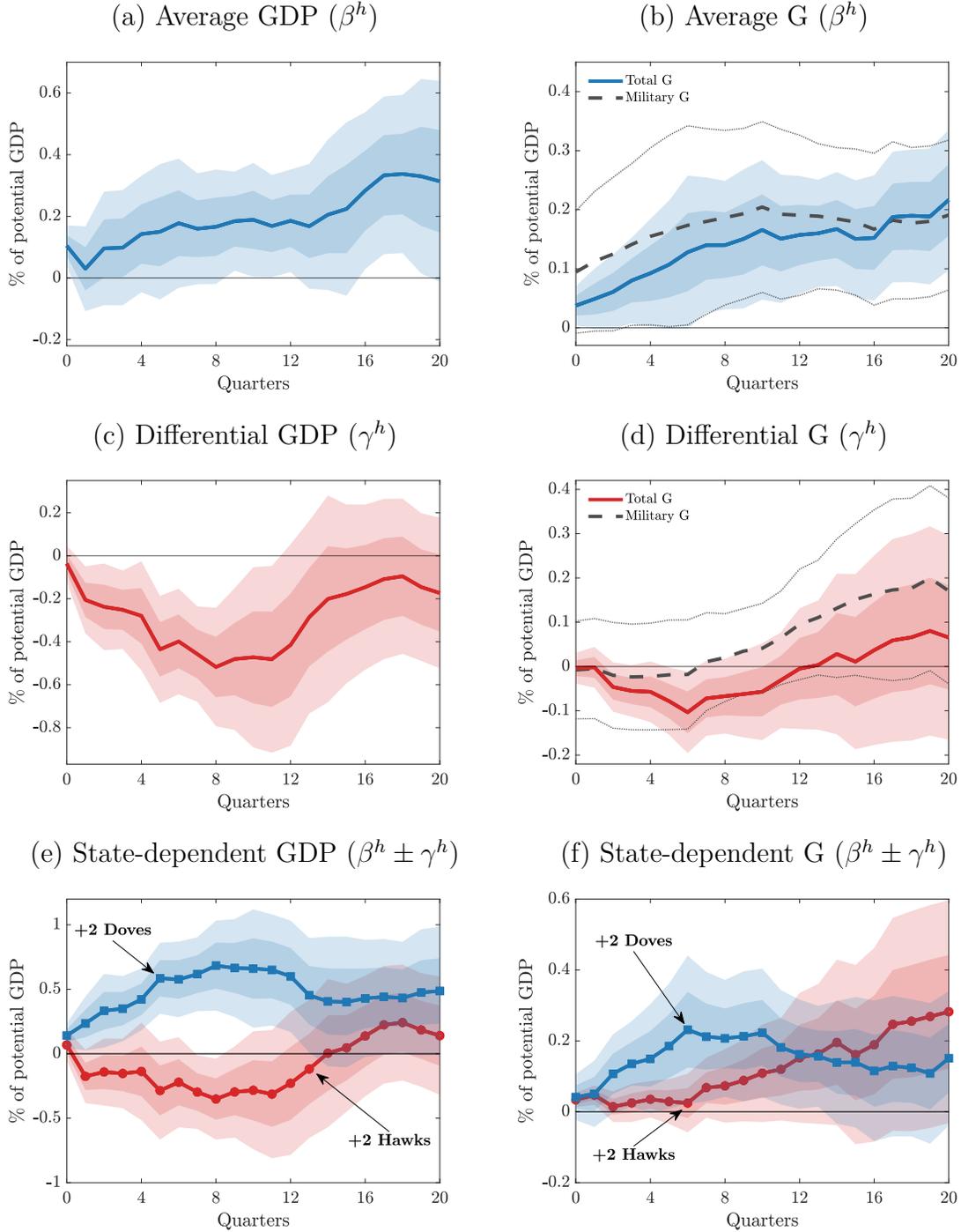
Baseline IV estimates. Figure 3 shows the responses of real GDP and real government spending (G) to a military spending shock conditional on systematic monetary policy ($Hawk_t$). The estimates are based on the local projection framework (3.4)-(3.5) as specified in Section 4.1. The solid lines show the 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 GDP.³⁵

Panels (a) and (b) show the IV 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 0.15% for GDP and 0.11% for total G after one

³⁴For the Newey-West standard errors, we set the bandwidth to $h + 1$, where h is the horizon in (3.4). A truncation parameter rule ([Lazarus et al., 2018](#)) or automatic bandwidth selection leads to similar results.

³⁵Normalizing the responses to a shock size of 1% of GDP approximately normalizes to one standard deviation of the shock series, which is 1.17% of GDP.

Figure 3: Responses to spending shocks conditional on monetary policy



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 GDP, conditional on systematic monetary policy ($Hawk_t$). We show IV estimates based on the local projection framework (3.4)-(3.5) 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.

year. The response of military G (dashed line) resembles total G, meaning the expansion of total G primarily reflects higher military G.³⁶

Panels (c) and (d) show the estimates of γ^h , which capture the differential responses of GDP and G when the FOMC exceeds the average Hawk-Dove balance by two hawks. 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. 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 magnitudes are sizeable. Between two and three years after the shock the GDP response is more than 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 government spending (G) is also negative at horizons until three years after the shock, albeit smaller in absolute terms and less significant.

The differential response of military G is insignificant at all horizons. This results supports our identifying assumption that the military spending shock does not depend on the Hawk-Dove balance. In contrast, the negative γ^h for total G means non-military G falls in response to more hawkish monetary policy. This fiscal policy response is unsurprising in an environment of tighter monetary policy and constitutes a part of the transmission of systematic monetary policy.

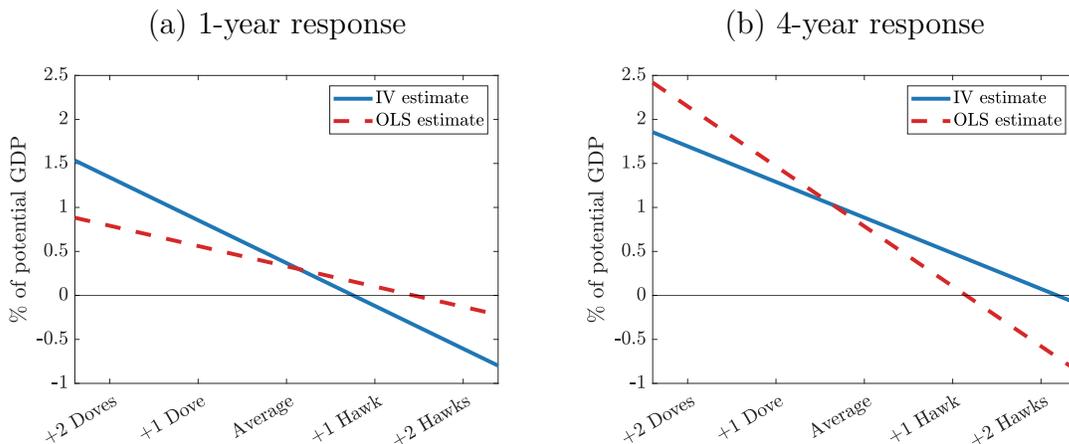
Panels (e) and (f) of Figure 3 show $\beta^h \pm \gamma^h$, the state-dependent responses when $Hawk_t$ exceeds the sample average either by two hawks (+2 Hawks) or by two doves (+2 Doves). The GDP response strongly varies between the dovish and the hawkish FOMC. The dovish FOMC supports the GDP expansion while the hawkish FOMC undoes the GDP expansion. Quantitatively, GDP increases by up to 0.68% 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 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 presented above with the OLS counterparts that do not use the FOMC rotation instrument. Figure 4 shows the response

³⁶Military G is defined relative to real potential GDP, analogous to total G, see Appendix B for details.

Figure 4: GDP responses for OLS and IV



Notes: The figure shows the yearly real GDP response to an expansionary military spending shock, corresponding to one percent of GDP, conditional on systematic monetary policy ($Hawk_t$). We show IV and OLS estimates based on the local projection framework (3.4)-(3.5) 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).

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.³⁷ 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.

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-

³⁷Figure F.3 in the Appendix presents the same responses for two and three years after the shock. Figure F.2 presents the OLS estimates of β^h and γ^h .

dependent fiscal multiplier as

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

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 ($Hawk_t - \overline{Hawk}$).³⁸ We estimate the responses for cumulative GDP and government spending jointly by seemingly unrelated regressions, see Appendix G.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).³⁹

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 at both horizons, and significantly different from zero at the 10% level. Analogous to Figure 3, 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 are similar across the two horizons. Moreover, the difference is statistically significant at the 5% level for the four-year horizon, see Table F.1 in Appendix F. 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 F.1. This result is analogous to the findings in Figure 3.

Comparison with linear model. We explicitly 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 $\widehat{FM}^H = (\sum_{h=0}^H \beta_{\text{GDP}}^h) / (\sum_{h=0}^H \beta_{\text{G}}^h)$ and the estimates are presented in the last column of Table 2. We find average multipliers

³⁸Alternatively, one could discount future horizons in (4.1). For common discount rates, this will have a minor impact on our estimated fiscal multipliers.

³⁹Our 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.

Table 2: Government spending multipliers and monetary policy

Outcome	Baseline model					Linear model
	+2 Hawks	+1 Hawk	Average	+1 Dove	+2 Doves	
<i>Two-year horizon</i>						
Multiplier	-4.825 (5.229)	-0.476 (1.418)	1.348 (0.708)	2.351 (0.934)	2.986 (1.239)	0.860 (1.427)
GDP (<i>cum</i>)	-1.689 (0.989)	-0.282 (0.768)	1.124 (0.649)	2.531 (0.689)	3.937 (0.865)	0.616 (1.057)
G (<i>cum</i>)	0.350 (0.250)	0.592 (0.300)	0.834 (0.395)	1.076 (0.510)	1.319 (0.634)	0.716 (0.338)
<i>Four-year horizon</i>						
Multiplier	-1.790 (2.637)	-0.001 (0.862)	1.308 (0.475)	2.307 (0.808)	3.095 (1.162)	0.838 (1.449)
GDP (<i>cum</i>)	-2.735 (2.498)	-0.002 (1.557)	2.731 (0.842)	5.465 (1.045)	8.198 (1.892)	1.494 (2.747)
G (<i>cum</i>)	1.528 (1.010)	1.808 (0.804)	2.088 (0.734)	2.368 (0.848)	2.649 (1.079)	1.782 (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 (3.4)-(3.5) 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 (3.4). Driscoll-Kraay standard errors are in parenthesis, see Appendix G for details.

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. Moreover, one potential reason for the broad range of multiplier estimates in the literature is not accounting for time variation in systematic monetary policy.

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.

First-stage results. Our local projection framework (3.4) contains two endogenous regressors, $\varepsilon_t^s(Hawk_t - \overline{Hawk})$ and $(Hawk_t - \overline{Hawk})$. The estimates of the two associated first-stage regressions are shown in Table G.1 in the Appendix. We find that the instrumental variable $\varepsilon_t^s(Hawk_t^{IV} - \overline{Hawk}^{IV})$ has a positive effect on the endogenous variable $\varepsilon_t^s(Hawk_t - \overline{Hawk})$ that is significant at the one percent level. Similarly, $(Hawk_t^{IV} - \overline{Hawk}^{IV})$ 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).

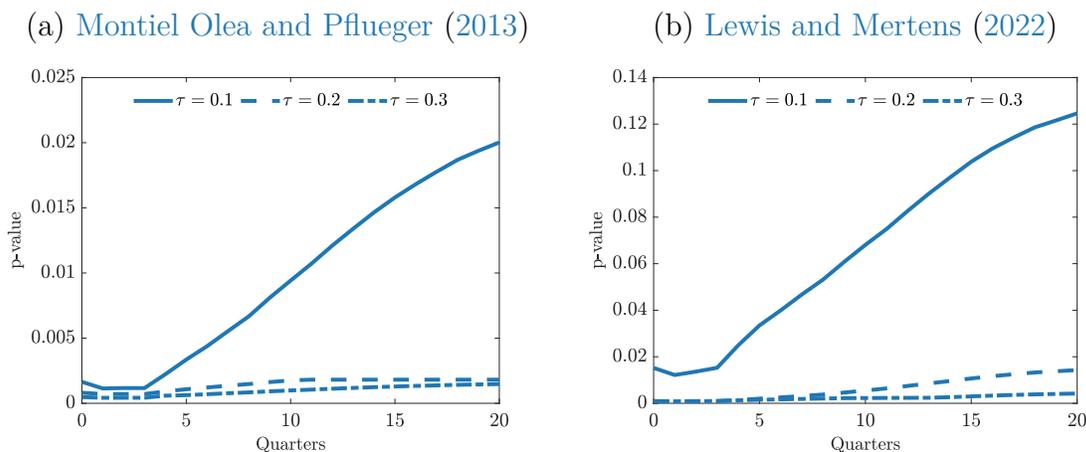
Weak instrument tests. We use three statistical tests to assess the strength of our instrument more formally. First, we use the Montiel Olea and Pflueger (2013) test of weak instruments, which is popular in time series settings because it is robust to autocorrelation and heteroskedasticity. Formally, 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 5 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 (2022) 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 5. 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 the same conclusion, see Figure G.1 in Appendix G.

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.

⁴⁰We 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 $Hawk_t^{IV}$ we obtain very similar results.

Figure 5: Weak instrument tests



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 (3.4)-(3.5) 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 \(2022\)](#) 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.

Figure G.2 in the Appendix shows that the reduced-form estimates for γ^h are significant, as in Figure 3. To summarize, all three tests indicate that our instruments are not weak.

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 G.3 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 a detailed description of the implementation in Appendix G.2. The robust confidence sets are presented in Figure G.4. They leave our conclusions about fiscal multipliers in Table 2 broadly unchanged. In particular, we estimate dovish fiscal multipliers with p-values of 0.08 and 0.12 for +1 Dove and +2 Doves, respectively. The hawkish multipliers are highly insignificant. Finally, the average multiplier is significant with a p-value of 0.06, whereas the estimate of the multiplier in the liner model remains highly 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. The multiplier estimates for all specifications are provided in Appendix H.⁴¹

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 multipliers similar to the baseline, see Table H.2. 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 baseline multipliers, see Table H.1.

We also address potential endogeneity concerns due to preference swings of policymakers by alternative rotation instruments. 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 fiscal multipliers in Table H.2 is slightly muted compared to the baseline.⁴²

Another potential concern is that the meaning of being a hawk or dove might have changed over time, see the discussion in Section 3.1. To account for trends in the Hawk-Dove balance, we consider an alternative Hawk-Dove balance which subtracts from the baseline $Hawk_t$ its backward-looking 5, 10, or 15-year moving average. The average and dovish multipliers have similar magnitudes as the baseline while the hawkish multiplier is similarly imprecise, see Table H.2 in the Appendix. Overall, our results reinforce the arguments in Section 3.1 that the classification of hawks and doves is indeed comparable over time.

⁴¹For the impulse responses of GDP and G associated to the sensitivity analysis, see Appendix H of the CEPR Discussion Paper version (Hack et al., 2023, see <https://cepr.org/publications/dp17999>).

⁴²The finding of muted state-dependence in the multipliers does not necessarily imply that our results are partly driven by endogenous swings. The alternative rotation instrument also takes out variation from swings that are exogenous, see the discussion in Section 3.2.

Alternative spending shock. Our baseline shock is specific to military spending. 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. This is in line with the nature of the BP shock. More importantly, we find that a hawkish FOMC significantly dampens the expansionary effect on GDP. The average fiscal multiplier is around 1.4 for the four-year horizon, see Table H.2, which is remarkably similar to our baseline multiplier. The fiscal multiplier ranges from 0.88 to 1.74 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 significant.⁴³

Great Recession and ZLB. Our baseline results are estimated using the 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 to exclude the GR and ZLB period or in 2008Q4 to exclude the ZLB period. For both of these subsamples, our multiplier estimates are very similar to the baseline, see Table H.2.

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 Table H.2. Nevertheless, we estimate dovish multipliers around 2 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 Table H.2.

⁴³The compressed variation in the multiplier appears consistent with the interest rate responses to BP shocks. Initially, interest rates significantly rise under a more hawkish FOMC, but the magnitude is smaller than for the baseline spending shocks. Starting two years after the shock, the differential interest rate response flips sign and interest rates are lower under a more hawkish FOMC. An important reason for the different interest rate responses across spending shocks may be the fact that G rises only temporarily in response to BP shocks, but rises persistently after military spending shocks.

5 Inspecting the mechanism

In this section, we inspect the mechanism behind our findings in the previous section. 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, supporting a crowd in of consumption and investment.

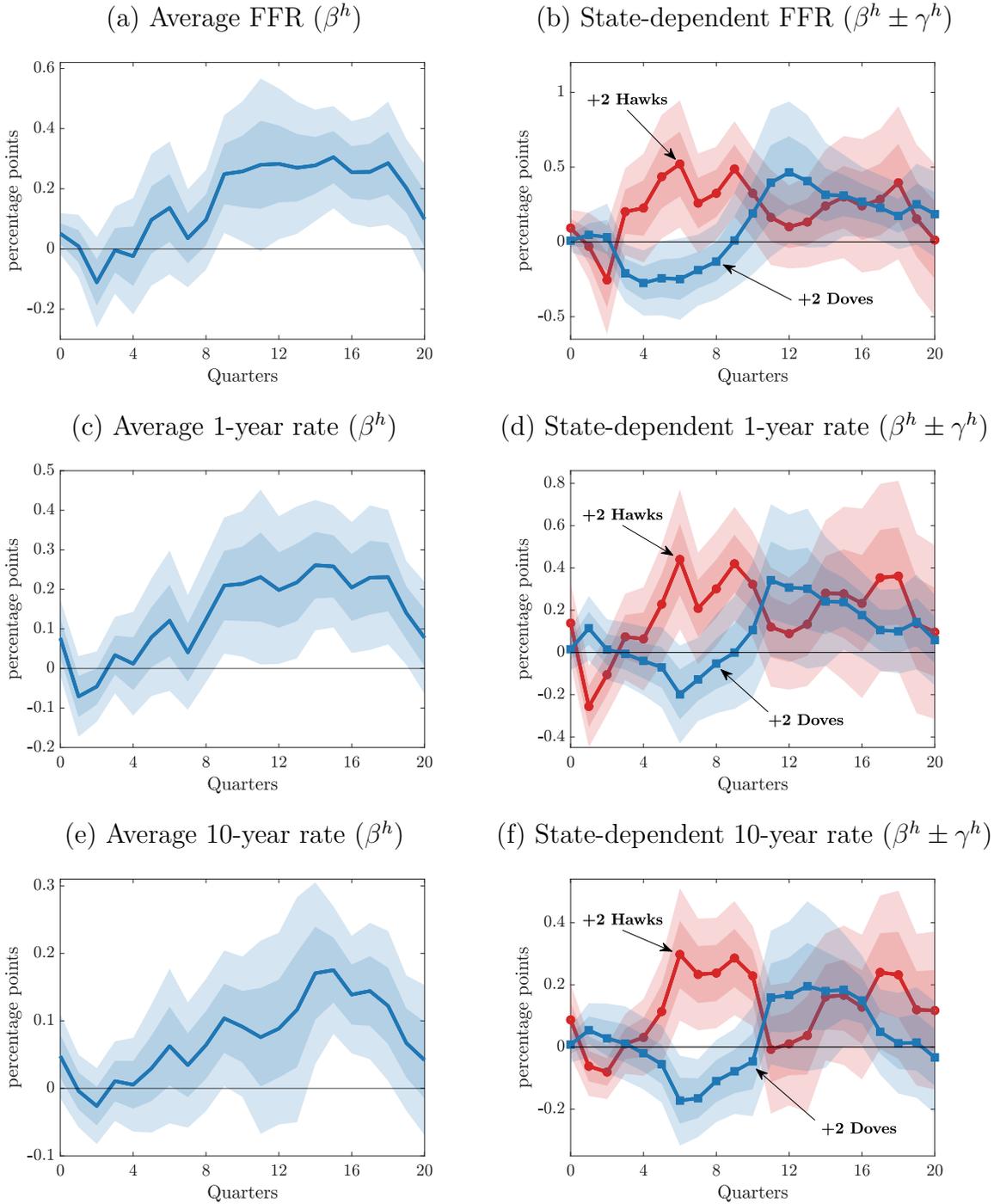
5.1 Additional responses

Conventional wisdom says that monetary policy tightens in response to higher government spending in order to mitigate the inflationary pressure. The Federal Reserve can use a range of tools, including the target federal funds rate, the discount rate, balance sheet policies and communication including forward guidance. These tools can affect short- and long-term interest rates, and hence inflation, consumption, and investment.

Nominal interest rates. We study the response of the federal funds rate (FFR) and the annualized yield on 1-year and 10-year Treasury securities to government spending shocks by using our local projection framework (3.4)-(3.5) with interest rates as outcome variable x_{t+h} . We 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 control for pre-trends in these outcomes.

Panels (a), (c) and (e) of Figure 6 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 30 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 56 basis points). In contrast, under a dovish FOMC, the FFR falls for almost two years and a reversion to a higher FFR is observed only three years after the shock. Similarly, both 1-year and 10-year Treasury yields increase after two years under a dovish FOMC, suggesting that the monetary regimes also differ in their effects on expected future policy at long horizons. The delayed FFR response is consistent with the initial uncertainty surrounding the military spending shock and the gradually evolving macroeconomic effects of the shock, see Figure 3.

Figure 6: Responses of nominal interest rates



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 GDP, conditional on systematic monetary policy ($Hawk_t$). All outcomes are annualized interest rates. We show IV estimates based on the local projection framework (3.4)-(3.5) 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.

Section 6 provides narrative evidence from the FOMC historical records suggesting that indeed the FOMC delays action until some uncertainty about the spending plans and their potential effect on the economy and inflation is resolved. Furthermore, a delayed differential policy response that extends for several quarters beyond the term of the FOMC and the associated rotation present at the time of the shock, is consistent with the decision dynamics in the FOMC. For example, Laurence Meyer, member of the Board of Governors from 1996 to 2002, describes these dynamics during his term at the Fed as follows:

So was the FOMC meeting merely a ritual dance? No. I came to see policy decisions as often evolving over at least a couple of meetings. The seeds were sown at one meeting and harvested at the next. [...] Similarly, while in my remarks to my colleagues it sounded as if I were addressing today's concerns and today's policy decisions, in reality I was often positioning myself, and my peers, for the next meeting.

Laurence Meyer (2004), *A Term at the Fed: An Insider's View*, Harper Business

Consistent with Meyer's view that it takes time to influence policy strategies in the FOMC, we find that the FOMC rotation ($Hawk_t^{IV}$) is more important for the policy response to the spending shock and its real effects when the shock occurs closer to the beginning of the FOMC rotation, which takes place in the first quarter of the year. When we drop spending shocks in the second half of the year, we obtain similar findings compared to the baseline, see Figures I.1-I.2 in Appendix I. Conversely, the dependence on monetary policy becomes weaker and less significant when dropping spending shocks in the first half of the year.

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 inflation.⁴⁴ We estimate the inflation responses using the specification of our local projection framework (3.4)-(3.5) for nominal interest rates and control for four lags of the inflation measure under consideration. The results are shown in Figure 7. 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

⁴⁴We 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.

pattern, suggesting that policy tightening is successful in containing inflationary pressures. Compared to the interest rate responses, the inflation response appear 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 real interest rate response. We consider real interest rates constructed by subtracting the expected CPI inflation from the three nominal interest rates considered in Figure 6. Figure I.3 in Appendix I presents the IV estimates of the average and state-dependent responses.

Investment and consumption. We examine the underlying components of the responses of real GDP. The fiscal spending multiplier can be above one when GDP components other than G are crowded in by the spending shock. Conversely, crowding out may lead to multipliers below one. We find that the differential GDP effects are primarily driven by private consumption and somewhat less by private investment, see Figure I.4 in Appendix I.⁴⁵ For the average Hawk-Dove balance, we find a mild but insignificant crowding out of private consumption and crowding-in of private investment in the short run. In contrast, the crowding out of consumption is strong and significant under a hawkish FOMC. For investment, we find a similar albeit smaller and less significant pattern. Overall, the strong state-dependence of fiscal multipliers appears to be mainly driven by private consumption.

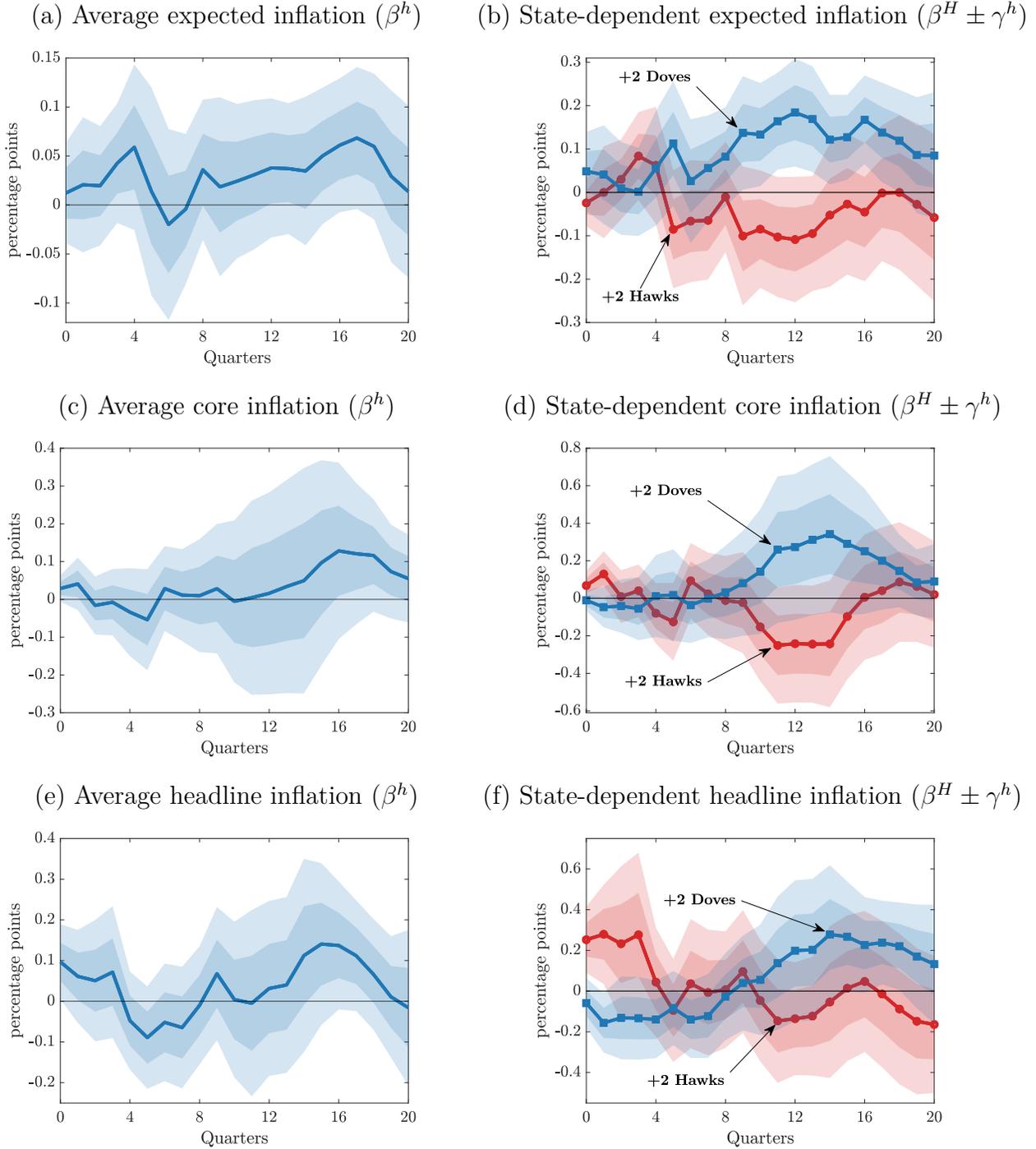
5.2 Relation to the literature

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

Relation to monetary policy shocks. Most of the related empirical literature estimates the effects of monetary policy shocks in the economy. Therefore, it may be interesting to

⁴⁵For details on the definition of consumption and investment, see Appendix B.

Figure 7: Responses of inflation rates



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 GDP, conditional on systematic monetary policy ($Hawk_t$). All outcomes are annualized inflation rates. We show IV estimates based on the local projection framework (3.4)-(3.5) 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.

compare the effects of such shocks with our estimates.⁴⁶ To this aim, we compare the ratio of the peak output to 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 $(\min_h \gamma_y^h)/(\max_h \gamma_i^h)$, from our estimation. For U.S. monetary policy shocks, recursively identified shocks in Coibion (2012) imply a ratio of -1.56 , Romer and Romer (2004) 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 .⁴⁷ Hence, a peak interest rate hike of one percentage point coincides with a peak decline of real GDP between 1 and 2 percentage points. In comparison, our estimates imply a ratio of -1.35 , which is within the range implied by evidence on monetary policy shocks.

Relation to fiscal multipliers. The interest rate responses further allow us to relate our fiscal spending multiplier estimates in Table 2 with the findings in the related literature. Our spending multiplier is between two and three under the dovish FOMC which is associated with a weak 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.

Our findings also relate to an empirical literature that estimates fiscal spending multipliers. For example, Nakamura and Steinsson (2014) estimate 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 to our two-year multiplier estimates. 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.⁴⁸ This requires a χ slightly below the “+1 Dove” case in Table 2. The associated two-year spending multiplier is 1.9, which is 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,

⁴⁶As we discuss in the introduction, an advantage of our approach is that it circumvents potential concerns related to the identification of monetary policy shocks and their size.

⁴⁷We compute these numbers based on the (baseline) estimates reported in each article, using the respective replication package.

⁴⁸Formally, we solve $\min_{\chi} \sum_{h=0}^8 (\beta_{FFR}^h - \chi \cdot \gamma_{FFR}^h)^2$, where χ indicates a level of the Hawk-Dove balance in deviation from the sample mean ($Hawk_t - \overline{Hawk}$).

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.

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.⁴⁹

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 appears on average 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, therefore they wait until "all the evidence was in". Third, the effects on inflation become

⁴⁹Former 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 decision of the Federal Reserve to raise the discount rate on December 1965. Lyndon B. Johnson was the President of the United States from 1963 to 1969.

visible but the FOMC delays action. The first two are common for hawkish and dovish committees while the third phase is more pronounced under a dovish one, broadly in line with our empirical findings.

We provide a short version of these case studies in Appendix J and an extended version in the Appendix J of the CEPR Discussion Paper version of this paper, see [Hack et al. \(2023\)](#). 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 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 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.

Finally, while our identification design is specific to U.S. monetary policy, a promising avenue for future research is to study other countries or currency areas in which committees decide monetary policy. In fact, since 2015 the European Central Bank's governing council allocates voting rights to its members through a rotation mechanism.

References

- ABRAMS, B. (2006): “How Richard Nixon Pressured Arthur Burns: Evidence from the Nixon Tapes,” *Journal of Economic Perspectives*, 20, 177–188.
- ABRAMS, B. A. AND J. L. BUTKIEWICZ (2012): “The Political Business Cycle: New Evidence from the Nixon Tapes,” *Journal of Money, Credit and Banking*, 44, 385–399.
- ANDERSON, T. W. AND H. RUBIN (1949): “Estimation of the parameters of a single equation in a complete system of stochastic equations,” *The 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.
- ARIAS, J. E., D. CALDARA, AND J. F. RUBIO-RAMIREZ (2019): “The systematic component of monetary policy in SVARs: An agnostic identification procedure,” *Journal of Monetary Economics*, 101, 1–13.
- AUERBACH, A. AND Y. GORODNICHENKO (2012): “Measuring the Output Responses to Fiscal Policy,” *American Economic Journal: Economic Policy*, 4, 1–27.
- BAILY, M. N., C. HULTEN, AND D. CAMPBELL (1992): “Productivity Dynamics in Manufacturing Plants,” *Brookings Papers on Economic Activity*, 23, 187–267.
- 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 (2022): “A Sufficient Statistics Approach for Macro Policy Evaluation,” Working Paper Series 2022, Federal Reserve Bank of San Francisco.
- BARRO, R. J. AND C. J. REDLICK (2011): “Macroeconomic Effects from Government Purchases and Taxes,” *Quarterly Journal of Economics*, 126, 51–102.
- 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. PFLUEGER, AND A. SUNDERAM (2022): “Perceptions about Monetary Policy,” NBER Working Paper 30480.
- 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.
- BEN ZEEV, N., V. A. RAMEY, AND S. ZUBAIRY (2023): “Do Government Spending Multipliers Depend on the Sign of the Shock?” *AEA Papers and Proceedings*, forthcoming.
- 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 O. F. HUMPAGE (2014): “Federal Reserve policy and Bretton Woods,” NBER Working Paper 20656.
- 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.
- CHAPPELL, H., R. R. MCGREGOR, AND T. A. VERMILYEA (2005): *Committee Decisions on Monetary Policy: Evidence from Historical Records of the Federal Open Market Committee*, vol. 1, MIT Press, 1 ed.
- 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.
- 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.
- CHRISTIANO, L. J., M. EICHENBAUM, AND C. L. EVANS (1999): “Monetary policy shocks: What have we learned and to what end?” *Handbook of Macroeconomics*, 1, 65–148.
- CIESLAK, A., S. HANSEN, M. MCMAHON, AND S. XIAO (2022): “Policymakers’ Uncertainty,” *mimeo*.
- 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.
- (2022): “The fiscal roots of inflation,” *Review of Economic Dynamics*, 45, 22–40.
- 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.
- DRECHSEL, T. (2023): “Estimating the Effects of Political Pressure on the Fed: A Narrative Approach with New Data,” *CEPR Discussion Paper*.
- 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.
- 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.
- FERGUSON, N., M. KORNEJEV, P. SCHMELZING, AND M. SCHULARICK (2023): “The Safety Net: Central Bank Balance Sheets and Financial Crisis, 1987-2020,” Economics working paper 23102, Hoover Institution.
- FERRIÈRE, A. AND G. NAVARRO (2018): “The Heterogeneous Effects of Government Spending : It’s All About Taxes,” *International Finance Discussion Papers*, 1237.
- 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.
- GORDON, R. J. AND R. KRENN (2010): “The End of the Great Depression: VAR Insight on the Roles of Monetary and Fiscal Policy.” NBER Working Paper 16380.
- HACK, L., K. ISTREFI, AND M. MEIER (2023): “Identification of Systematic Monetary Policy,” CEPR Discussion Paper 17999 (<https://cepr.org/publications/dp17999>).
- 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.
- LEEPER, E. M., N. TRAUM, AND T. B. WALKER (2017): “Clearing up the Fiscal Multiplier Morass,” *American Economic Review*, 107, 2409–54.
- LEWIS, D. J. AND K. MERTENS (2022): “A Robust Test for Weak Instruments with Multiple Endogenous Regressors,” Staff reports, no. 1020, Federal Reserve Bank of New York.
- LIU, Y. AND J. C. WU (2021): “Reconstructing the yield curve,” *Journal of Financial Economics*, 142, 1395–1425.
- LUCAS, R. E. (1976): “Econometric policy evaluation: A critique,” *Carnegie-Rochester Conference Series on Public Policy*, 1, 19–46.
- MCKAY, A. AND C. K. WOLF (2022): “What Can Time-Series Regressions Tell Us About Policy Counterfactuals?” NBER Working Paper 30358.
- 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.
- 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 M. D. SHAPIRO (1998): “Costly capital reallocation and the effects of government spending,” *Carnegie-Rochester Conference Series on Public Policy*, 48, 145–194.
- 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*, National Bureau of Economic Research, Inc, 121–184.
- ROMER, C. D. AND D. H. ROMER (2002): “The Evolution of Economic Understanding and Postwar Stabilization Policy,” NBER Working Paper 9274.
- (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.
- SIMS, C. A. AND T. ZHA (2006): “Were There Regime Switches in US Monetary Policy?” *American Economic Review*, 96, 54–81.
- WOLF, C. K. (2023): “Fiscal Stimulus and the Systematic Response of Monetary Policy,” *AEA Papers and Proceedings*, forthcoming.
- WOODFORD, M. (2011): “Simple Analytics of the Government Expenditure Multiplier,” *American Economic Journal: Macroeconomics*, 3, 1–35.

Appendix

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:

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

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

$$i_t = \tilde{\phi}_t \pi_t, \quad (\text{A.3})$$

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

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

Combining (A.1) and (A.4), the model dynamics follow $\mathcal{Y}_t = A(\phi_t) \mathbb{E}_t [\mathcal{Y}_{t+1}] + B(\phi_t) \mathcal{X}_t$, 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

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

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

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

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:

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

Defining as instrument vector $Q_t = [\varepsilon_t^s, \varepsilon_t^s \phi_t^{IV}, \phi_t^{IV}]'$, 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.⁵⁰ Hence, the IV estimator consistently estimates $(\beta^h, \gamma^h, \delta^h)$ even absent strong exogeneity assumptions for ϕ_t .

⁵⁰Note 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.

Appendix B Data

B.1 Narrative data

We use [Istrefi's \(2019\)](#) data set which is a panel containing the policy preferences of voting FOMC members at each FOMC meeting for 1960-2023.

The news coverage of FOMC members is relatively sparse during the first six years in our sample, leaving us with relatively more unclassified FOMC members in this period. For example, we observe the preferences for 115 out of 195 member-meeting pairs in 1960. Fortunately, the share of observed preferences increases quickly and from 1966 onward, we reach an average share of 88 percent. Specifically for the first six years, we account for some of the 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 believe a plausible assumption is that short-term substitutes act in the best interest of the person that is substituted, partly because substitutes are often direct subordinates of the original voting member. More specifically, we assume that short-term substitutes act as if the original member attended the meeting if the following three criteria hold: (i) the substitution period is no longer than six months when the substitute is from the same Federal Reserve bank, (ii) the substitution period is no longer than three months if the substitute is not from the same Federal Reserve bank, (iii) the substitution does not take place at the beginning or the end of a rotation cycle within a rotation group.⁵¹ However, it frequently holds that the preferences of the substitute and the original voter coincide which implies that the procedure above does not change the data. We change less than 1% of preferences when a substitution occurs 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 steps to create the aggregate series as in their paper.⁵²

From FRED, we use headline CPI ($CPIAUCSL$) and CPI core ($CPILFESL$) inflation defined

⁵¹For 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 $\tau + 1$.

⁵²The 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 $ngov_t / (pgdp_t \times rgdp_pott6_t) \times 100$.

as the year-over-year growth rate of the respective price index, and the effective federal funds rate (*DFE*). The 10-year treasury market yield (*DGS10*) starts only in 1962q1 and is therefore combined with the very 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 impute the first four observations (1960q1 to 1960q4) with a similar 1-year treasury market yield from Fred (*DTB1YR*). Personal consumption expenditures (*PCE*), gross private domestic investment (*GPDI*), and federal government defense expenditures (*FDEFX*) is divided by the GDP deflator and by real potential GDP, both taken from [Ramey and Zubairy \(2018\)](#), see above. We compute non-military government spending by subtracting the defense spending from total government spending. 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. Formally, we let $\pi_t^e = \pi_{t-1}^e$, whenever there is no survey conducted in quarter t . The (ex-ante) real rates are computed as $i_t^r = i_t^n - \pi_t^e$ where i_t^n is a nominal rate of interest.

The validation exercise in Appendix E is based on forecasts from the Fed’s Greenbook. We use the average of the one- and two-quarter ahead inflation forecast, following [Coibion and Gorodnichenko \(2011\)](#). For the [Blanchard and Perotti \(2002\)](#) shock, we account for anticipation in government spending by including the one-quarter projected growth rate of government spending from [Ramey’s \(2011\)](#) data.⁵³ We further consider as control variable 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 decompose fluctuations in $Hawk_t$ and $Hawk_t^{IV}$ finding that the FOMC rotation is a key source of variation for both time series.

Decomposition of $Hawk_t$. We derive a decomposition of the aggregate Hawk-Dove balance similar to the aggregate productivity decomposition in [Baily et al. \(1992\)](#). We first rewrite

⁵³The SPF provides the government spending forecasts only from 1981q3 onward. [Ramey \(2011\)](#) imputes the government spending forecasts with defense spending forecasts to extend the sample until 1968q4.

the aggregate Hawk-Dove balance in equation (3.2) as⁵⁴

$$Hawk_t = \sum_{i \in \mathcal{M}_t} s_t Hawk_{it}, \quad s_t = \frac{1}{|\mathcal{M}_t|}. \quad (\text{C.1})$$

We define a decomposition over p-period changes in the balance:

$$\Delta^p Hawk_t = Hawk_t - Hawk_{t-p} = \sum_{i \in \mathcal{M}_t} s_t Hawk_{it} - \sum_{i \in \mathcal{M}_{t-p}} s_{t-p} Hawk_{it-p} \quad (\text{C.2})$$

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 to rewrite:

$$\begin{aligned} \Delta^p Hawk_t &= \sum_{i \in S_t} (s_t Hawk_{it} - s_{t-p} Hawk_{it-p}) + \sum_{i \in E_t} s_t Hawk_{it} - \sum_{i \in X_t} s_{t-p} Hawk_{it-p} \\ &= \sum_{i \in S_t} s_{t-p} (Hawk_{it} - Hawk_{it-p}) + \sum_{i \in S_t} (s_t - s_{t-p}) Hawk_{it} \\ &\quad + \sum_{i \in E_t} s_t Hawk_{it} - \sum_{i \in X_t} s_{t-p} Hawk_{it-p} \end{aligned} \quad (\text{C.3})$$

The first term captures changes in preferences of surviving FOMC members, the second term captures changes in the number of FOMC members, the third term captures entry into the FOMC, and the last term captures exit from the FOMC.

Finally, we further distinguish between the rotating and non-rotating FOMC members in the set of entering and exiting FOMC members, denoted E_t^R , E_t^N , X_t^R and X_t^N to obtain our decomposition of interest:

$$\begin{aligned} \Delta^p Hawk_t &= \sum_{i \in S_t} s_{t-p} (Hawk_{it} - Hawk_{it-p}) + \sum_{i \in S_t} (s_t - s_{t-p}) Hawk_{it} + \sum_{i \in E_t^N} s_t Hawk_{it} \\ &\quad - \sum_{i \in X_t^N} s_{t-p} Hawk_{it-p} + \sum_{i \in E_t^R} s_t Hawk_{it} - \sum_{i \in X_t^R} s_{t-p} Hawk_{it-p} \end{aligned} \quad (\text{C.4})$$

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 0.083. The variance of the first term of (C.4), which captures intensive margin changes of preferences, corresponds to 9% of the total variance. Changes in the weights, the second term, are

⁵⁴To be precise, we consider the first FOMC meeting in each quarter t .

negligible in size. The variance of the third and fourth term, capturing extensive margin changes of non-rotating FOMC members, corresponds to 22% of the total variance. The variance of the fifth and sixth term, capturing extensive margin changes of rotating FOMC members, corresponds to 53% of the total variance. Finally, the covariances between these terms account for 15% of the total variance. The results differ little for quarterly changes ($p = 1$). Notably, extensive margin changes of rotating FOMC members still account for 52% of the total variance.

Decomposition of $Hawk_t^{IV}$. Analogously, we propose a decomposition for the FOMC rotation instrument

$$\begin{aligned} \Delta^p Hawk_t^{IV} = & \sum_{i \in S_t^R} s_{t-p}^R (Hawk_{it} - Hawk_{it-p}) + \sum_{i \in S_t^R} (s_t^R - s_{t-p}^R) Hawk_{it} \\ & + \left(\sum_{i \in E_t^{RA}} s_t^R Hawk_{it} - \sum_{i \in X_t^{RA}} s_{t-p}^R Hawk_{it-p} \right) \\ & + \left(\sum_{i \in E_t^{RI}} s_t^R Hawk_{it} - \sum_{i \in X_t^{RI}} s_{t-p}^R Hawk_{it-p} \right), \end{aligned} \quad (C.5)$$

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. Yearly changes mechanically mute the importance of intensive margin changes, because current rotating FOMC members are typically not FOMC members a year later. Therefore, we also study quarterly changes ($p = 1$). Intensive margin changes now explain 4% of the variance, appointments account for 23%, and rotations of incumbent members account for 71%. Appointments become relatively more important for $p = 1$ because only every fourth quarter of $\Delta^1 Hawk_t^{IV}$ features a rotation. Compared to $\Delta^4 Hawk_t^{IV}$ for which the rotation affects all quarters, we mechanically lower the importance of rotations and the overall variance for $p = 1$.

Appendix D Relation to monetary policy shocks

We show that empirically identified monetary policy may blur the distinction between what in our New Keynesian model is systematic monetary policy (ϕ_t) and what would be monetary

policy shocks (a deviation from the Taylor rule).

Romer and Romer (2004), RR henceforth, identify monetary policy shocks as the residual from a regression of changes in the target federal funds rate on various Greenbook forecasts. To interpret their regression through the lens of our model in Section 2, we consider a stylized version of the RR regression

$$i_t = \phi^{\text{RR}} \pi_t^{\text{GB}} + \varepsilon_t^{\text{RR}}, \quad (\text{D.1})$$

in which π_t^{GB} denotes the Greenbook inflation forecast before a change in monetary policy, and $\varepsilon_t^{\text{RR}}$ is the RR monetary policy shock.⁵⁵ For simplicity, we put estimation and identification concerns aside and further assume the data is generated from the following policy rule

$$i_t = \phi_t \pi_t^{\text{GB}} + \varepsilon_t^m, \quad (\text{D.2})$$

where ε_t^m is a true monetary policy shock and systematic monetary policy satisfies $\phi_t > 1$. Combining this with the RR regression yields the RR monetary policy shock

$$\varepsilon_t^{\text{RR}} = (\phi_t - \phi^{\text{RR}}) \pi_t^{\text{GB}} + \varepsilon_t^m. \quad (\text{D.3})$$

Hence, the empirical shock $\varepsilon_t^{\text{RR}}$ captures variation in systematic monetary policy ϕ_t , variation in inflation forecasts π_t^{GB} , and monetary policy shocks ε_t^m . The empirical shock captures the model shock $\varepsilon_t^{\text{RR}} = \varepsilon_t^m$ in the special case when systematic monetary policy is time-invariant. In general, the empirical shock also captures joint time-variation in systematic monetary policy ϕ_t and inflation forecasts π_t^{GB} , where the latter naturally depends on the state of the economy. Finally, high-frequency identified monetary policy shocks may reflect changes in systematic monetary policy in a similar fashion (Bauer and Swanson, 2023).

Appendix E Validation exercise

We use the Hawk-Dove balance and the FOMC rotation instrument to estimate the federal funds rate FFR response to inflation forecasts as a function of the hawkishness of the FOMC. We find that a hawkish FOMC is associated with a more pronounced hike of the federal funds rate in the face of inflationary pressure. We estimate a state-dependent local projection specification that is akin to a forward-looking Taylor rule. Formally, we estimate a set of

⁵⁵The stylized regression omits any lags or leads from the original regression in Romer and Romer (2004). This is inconsequential if the DGP features *iid* fluctuations.

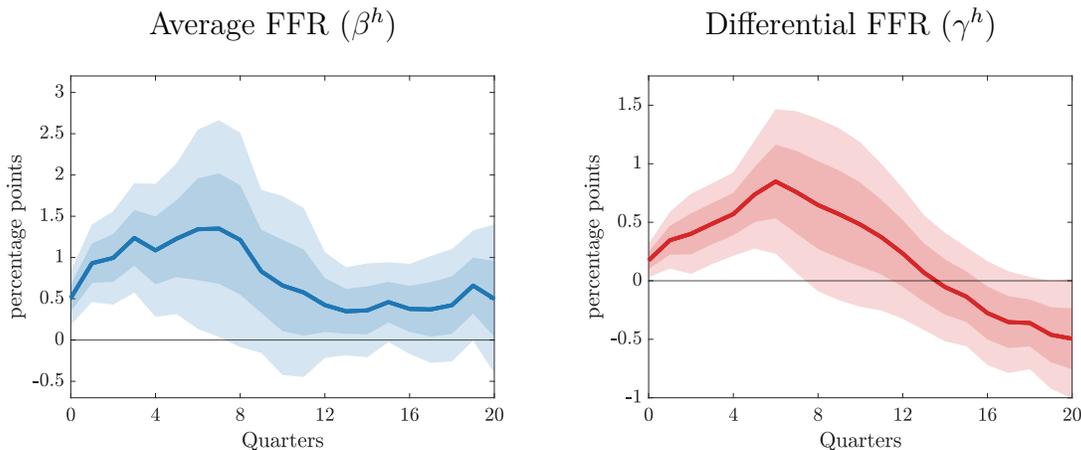
regressions

$$FFR_{t+h} = \alpha^h + \beta^h \hat{\pi}_t + \gamma^h \hat{\pi}_t (\text{Hawk}_t - \overline{\text{Hawk}}) + \zeta^h Z_{t-1} + v_{t+h}^h, \quad (\text{E.1})$$

for $h = 0, 1, \dots, H$, and FFR_{t+h} and $\hat{\pi}_t$ denote the federal funds rate and the average of the one- and two-quarter ahead Greenbook inflation forecast, respectively. The control vector includes four lags of the federal funds rate and the inflation forecast. The data is at a quarterly frequency and the sample runs from 1969 to 2008, due to the availability of inflation forecasts and the reaching of the zero lower bound in 2008.

Figure E.1 presents IV estimates where we use the FOMC rotation instrument interacted with the inflation forecast as an instrument for the interaction term in the specification above. We show estimates that are normalized to represent the inflation forecast being one percentage point above the sample average. The left panel displays the response under the average FOMC (β^h). The right panel displays the differential response (γ^h) when there are 2 more hawks in the FOMC relative to the average composition.

Figure E.1: FFR response to inflation and the FOMC hawkishness



Notes: The figure shows responses of the federal funds rate to an inflation Greenbook forecast that is one percentage point above its sample average, conditional on systematic monetary policy (Hawk_t). We show IV estimates based on (E.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 shaded areas indicate 68% and 95% confidence bands using Newey-West standard errors.

On average, the FOMC reacts with a federal funds rate hike. The response is statistically significant at the five percent level for six quarters. The response builds up over time, consistent with interest rate smoothing. Incidentally, it satisfies the Taylor principle for almost two years and peaks at 1.48 percentage points. The response turns stronger when the FOMC is more hawkish, as indicated by the differential effects in Panel (b). The estimates of the interaction coefficient γ^h are hump-shaped and peak after two years at 0.92 percentage

points. The response is significant at five percent for almost two years. This result suggests that a more hawkish FOMC is associated with a stronger and more persistent federal funds rate hike. Conversely, a more dovish FOMC implies a substantially weaker response.

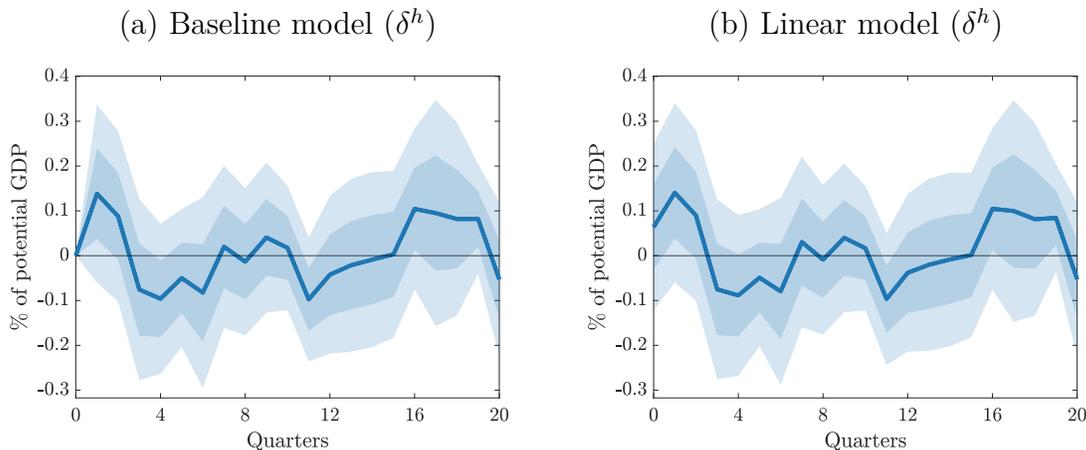
Finally, this validation exercise lends itself to assessing the relevance condition of our instrument more formally. We use the weak instruments test from [Montiel Olea and Pflueger \(2013\)](#).⁵⁶ We can reject the null of weak instruments. More formally, we compute p-values for the bias exceeding 10% percent of the benchmark, see [Montiel Olea and Pflueger \(2013\)](#) for details. The p-values are bounded from above by 0.055 and are below the 0.05 level at most horizons. Moreover, for a test of whether the bias exceeds 20%, we can reject the null at 1% for all horizons.

Overall, we show that the federal funds rate response to inflation correlates positively with the hawkishness of the FOMC, $Hawk_t$. The responses are consistent with our measurement of the stance of systematic monetary policy and are further in line with [Bordo and Istrefi \(2023\)](#). We see this result as a validation that our measurement of systematic monetary policy, through $Hawk_t$, captures important aspects of the Federal Reserve’s monetary policy-making.

Appendix F Additional results for Section 4

This appendix contains additional results for Section 4.1-4.3 in the main text.

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

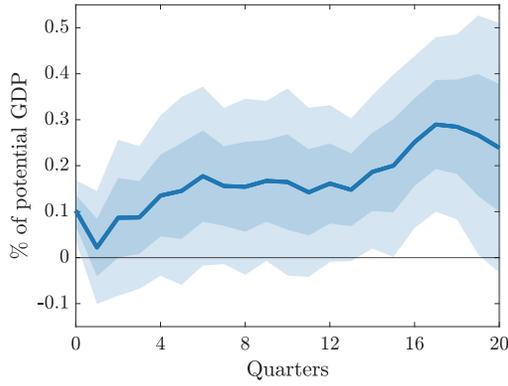


Notes: The figure shows responses of the military spending shock to systematic monetary policy ($Hawk_t$). We show IV estimates based on the local projection framework (3.4)-(3.5) as specified in Section 4.1. The δ^h captures the response when $Hawk_t$ exceeds the sample average by two hawks. Panel (a) 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 (3.4). The shaded areas indicate 68% and 95% confidence bands using Newey-West standard errors.

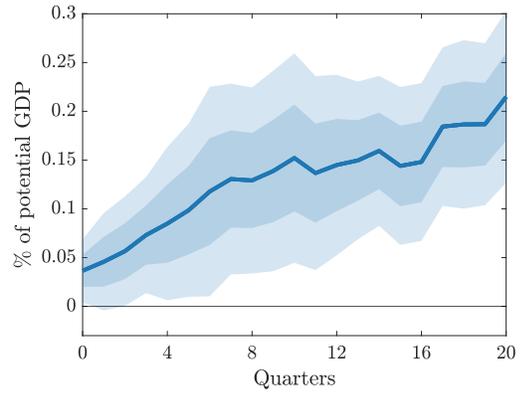
⁵⁶With a single endogenous regressor, this is equivalent to the [Lewis and Mertens \(2022\)](#) test.

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

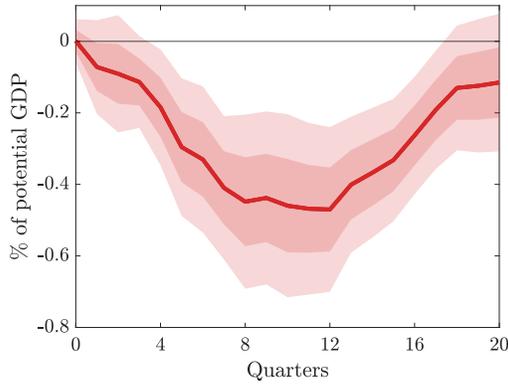
(a) Average GDP (β^h)



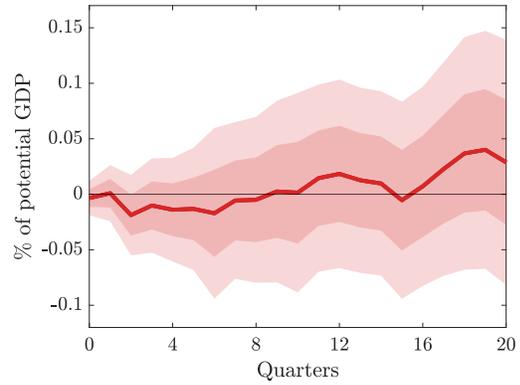
(b) Average G (β^h)



(c) Differential GDP (γ^h)

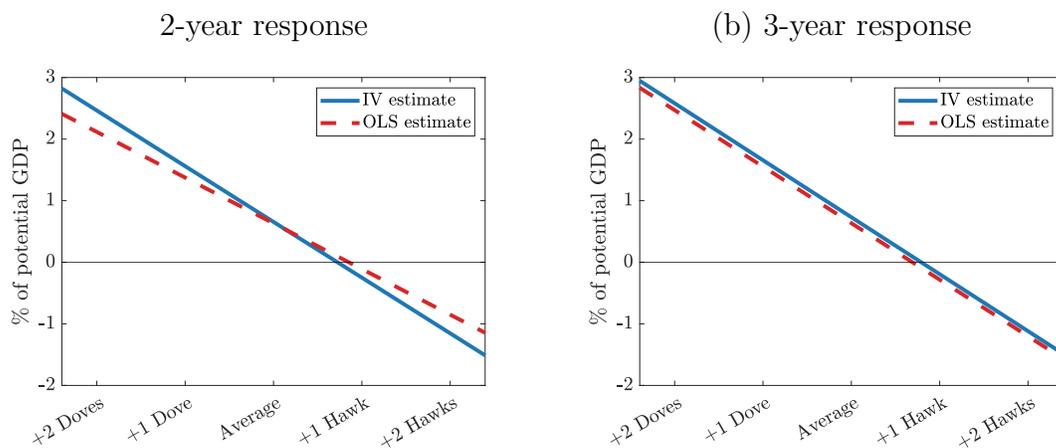


(d) Differential G (γ^h)



Notes: The figure shows 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 OLS estimates based on the local projection framework (3.4) 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 shaded areas indicate 68% and 95% confidence bands using Newey-West standard errors.

Figure F.3: GDP responses for OLS and IV



Notes: The figure shows the yearly real GDP response to an expansionary military spending shock, corresponding to one percent of GDP, conditional on systematic monetary policy ($Hawk_t$). We show IV and OLS estimates based on the local projection framework (3.4)-(3.5) 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 $H = 8$ quarters in Panel (a) and $H = 12$ quarters in Panel (b).

Table F.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	0.223	0.119	0.102	0.104
GDP (<i>cum</i>)		0.000		
G (<i>cum</i>)		0.080		
Four-year horizon				
Multiplier	0.245	0.122	0.041	0.041
GDP (<i>cum</i>)		0.008		
G (<i>cum</i>)		0.448		

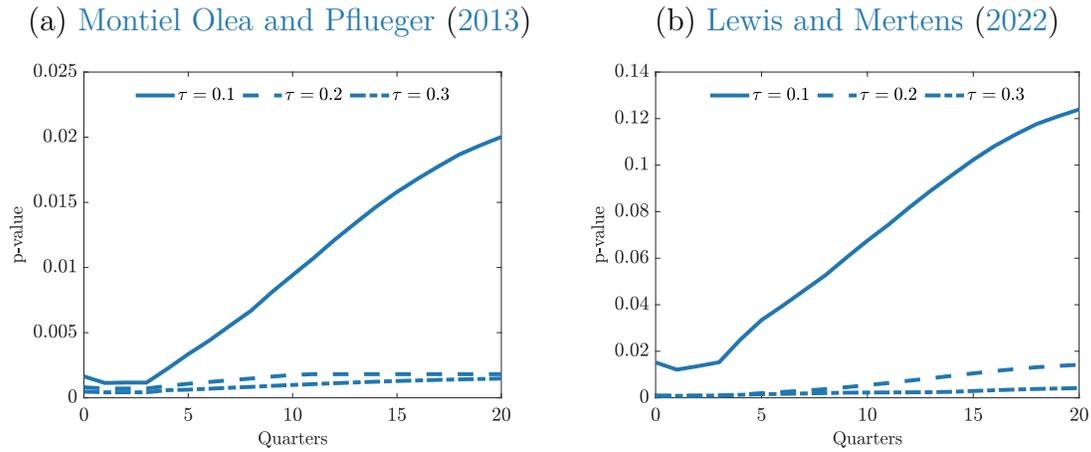
Notes: The table shows p-values corresponding to statistical tests for whether the fiscal multiplier or its components are significantly different across monetary regimes ($Hawk_t$). The tests are based on the multiplier estimates reported in Table 2 in Section 4.3, using Driscoll-Kraay standard errors,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

Appendix G Weak instruments and robust inference

This appendix contains additional results for Section 4.4. The first subsection presents diagnostics on instrument strength. The second section presents robust inference regarding weak instruments for impulse responses and fiscal multipliers.

G.1 Weak instrument tests

Figure G.1: Weak instrument tests



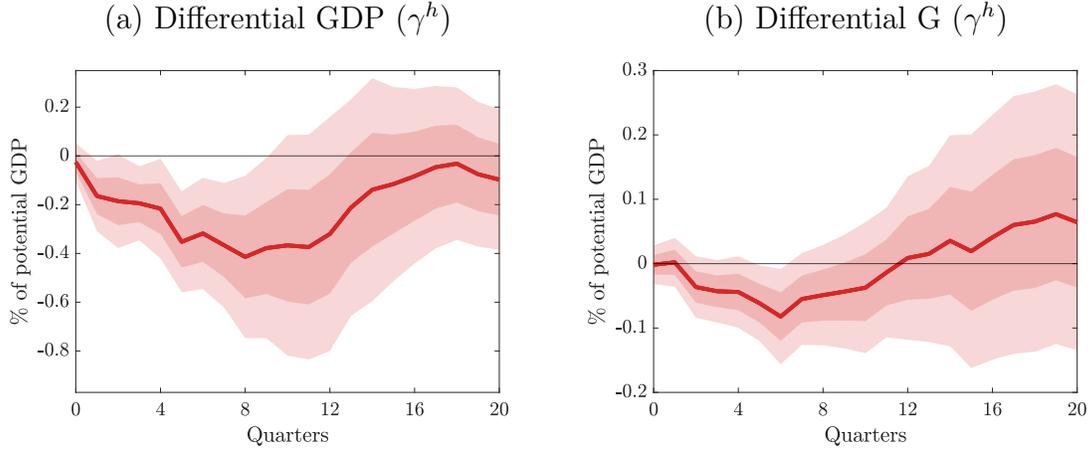
Notes: The figure shows p-values for rejecting the null of weak instruments for the responses of real government spending (G), based on the local projection framework (3.4)-(3.5) as specified in Section 4.1. The [Montiel Olea and Pflueger \(2013\)](#) test evaluates the null of the bias in γ^h exceeding a threshold τ . Similarly, the [Lewis and Mertens \(2022\)](#) 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.

Table G.1: Responses of GDP and government spending, incl. first-stage

Regressors	GDP responses				G responses				First-stage results	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
ε_t^s	0.142 (0.096)	0.166 (0.095)	0.185 (0.085)	0.283 (0.130)	0.092 (0.047)	0.140 (0.056)	0.157 (0.051)	0.152 (0.054)	0.050 (0.039)	0.010 (0.007)
$\varepsilon_t^s(Hawk_t - \overline{Hawk_t})$	-1.672 (0.775)	-3.099 (0.841)	-2.485 (1.433)	-0.873 (1.174)	-0.342 (0.209)	-0.401 (0.258)	-0.030 (0.416)	0.220 (0.653)		
$Hawk_t - \overline{Hawk_t}$	-2.770 (1.220)	-3.698 (1.728)	-4.247 (2.216)	-4.562 (2.217)	-0.593 (0.322)	-0.985 (0.650)	-1.389 (1.020)	-0.948 (1.135)		
$\varepsilon_t^s(Hawk_t^{IV} - \overline{Hawk_t}^{IV})$									0.290 (0.053)	-0.019 (0.021)
$Hawk_t^{IV} - \overline{Hawk_t}^{IV}$									-0.008 (0.017)	0.402 (0.042)
ε_{t-1}^s	0.024 (0.157)	0.057 (0.216)	0.086 (0.221)	0.245 (0.153)	0.044 (0.033)	0.076 (0.046)	0.092 (0.043)	0.124 (0.043)	0.007 (0.003)	0.011 (0.006)
ε_{t-2}^s	0.110 (0.125)	0.035 (0.185)	0.078 (0.205)	0.150 (0.160)	0.032 (0.030)	0.052 (0.030)	0.063 (0.041)	0.092 (0.049)	-0.012 (0.011)	0.007 (0.008)
ε_{t-3}^s	0.045 (0.149)	0.036 (0.163)	0.126 (0.153)	0.188 (0.144)	0.038 (0.018)	0.036 (0.028)	0.037 (0.045)	0.073 (0.052)	-0.000 (0.006)	0.008 (0.008)
ε_{t-4}^s	0.001 (0.141)	0.033 (0.125)	0.152 (0.117)	0.224 (0.144)	0.023 (0.022)	0.037 (0.027)	0.060 (0.041)	0.139 (0.038)	-0.018 (0.012)	0.004 (0.010)
GDP_{t-1}	1.314 (0.182)	0.777 (0.243)	0.424 (0.252)	0.037 (0.282)	0.033 (0.053)	0.103 (0.075)	0.135 (0.100)	0.124 (0.121)	-0.000 (0.013)	-0.012 (0.017)
GDP_{t-2}	-0.406 (0.190)	-0.166 (0.209)	-0.110 (0.159)	0.149 (0.197)	0.006 (0.054)	0.060 (0.072)	0.035 (0.077)	0.039 (0.094)	-0.016 (0.020)	0.013 (0.014)
GDP_{t-3}	-0.240 (0.203)	-0.012 (0.180)	-0.093 (0.223)	0.081 (0.171)	0.062 (0.055)	0.034 (0.068)	0.044 (0.068)	0.084 (0.062)	0.004 (0.016)	-0.005 (0.010)
GDP_{t-4}	-0.164 (0.183)	-0.440 (0.267)	-0.284 (0.313)	-0.444 (0.333)	-0.103 (0.051)	-0.218 (0.095)	-0.279 (0.138)	-0.355 (0.167)	0.003 (0.007)	-0.026 (0.011)
G_{t-1}	-0.639 (0.714)	0.012 (1.012)	0.336 (0.909)	0.864 (0.940)	1.340 (0.195)	1.311 (0.241)	1.121 (0.308)	1.138 (0.387)	0.028 (0.022)	-0.073 (0.055)
G_{t-2}	1.177 (0.617)	0.596 (0.734)	0.194 (0.602)	-0.223 (0.479)	0.008 (0.195)	-0.042 (0.220)	0.078 (0.277)	0.097 (0.278)	0.008 (0.039)	0.062 (0.047)
G_{t-3}	-0.391 (0.651)	-0.347 (0.706)	-0.233 (0.618)	-0.519 (0.491)	-0.079 (0.203)	-0.106 (0.248)	-0.136 (0.273)	-0.138 (0.272)	0.017 (0.054)	-0.091 (0.042)
G_{t-4}	0.022 (0.920)	0.020 (0.911)	0.018 (0.888)	0.162 (0.791)	-0.346 (0.202)	-0.308 (0.314)	-0.268 (0.437)	-0.343 (0.486)	-0.039 (0.049)	0.140 (0.060)
Observations	196	196	196	196	196	196	196	196	196	196
R^2	0.577	0.347	0.201	0.138	0.934	0.843	0.730	0.646	0.452	0.547
R^2 excl. IVs									0.036	0.154
F-statistic	16.398	4.243	3.418	2.630	94.688	22.683	11.316	15.287	43.691	28.077
F-statistic excl. IVs									4.935	5.804

Notes: The table shows 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 (3.4)-(3.5) as specified in Section 4.1. Columns (1) to (4) and (5) to (8) display the one, two, three, and four-year ahead responses, respectively. Regressor ε_t^s captures the responses when $Hawk_t$ equals its sample average and $\varepsilon_t^s(Hawk_t - \overline{Hawk_t})$ captures the differential responses. Columns (9) and (10) display the first-stage results for $\varepsilon_t^s(Hawk_t - \overline{Hawk_t})$ and $(Hawk_t - \overline{Hawk_t})$, respectively. Newey-West standard errors are in parenthesis.

Figure G.2: Differential responses of GDP and government spending, reduced-form

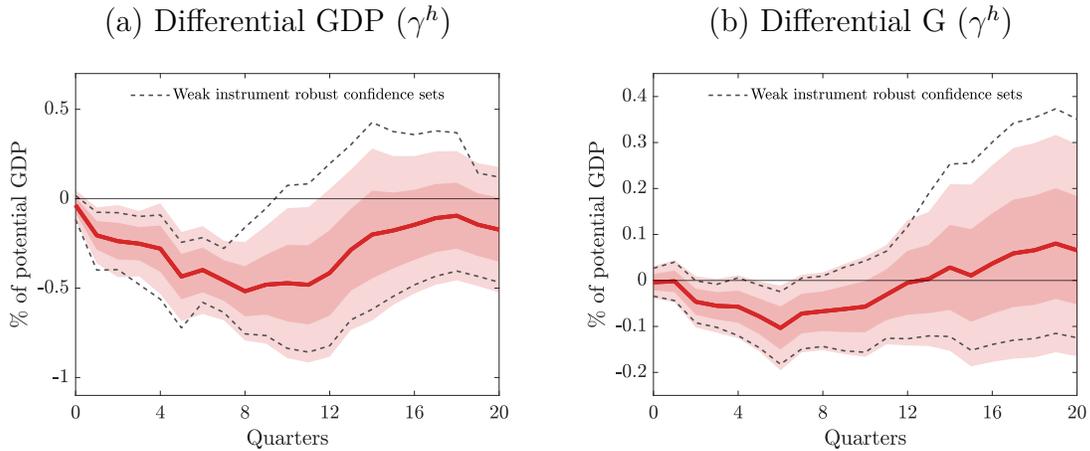


Notes: The figure shows differential responses of real GDP and real government spending (G) to an expansionary military spending shock, corresponding to one percent of GDP, conditional on systematic monetary policy ($Hawk_t$). We show reduced-form estimates based on the local projection framework (3.4)-(3.5) 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.

G.2 Robust inference

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

Figure G.3: Responses of GDP and government spending, robust inference



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 (3.4)-(3.5) 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).

Baseline multiplier inference. To obtain multiplier estimates and conduct inference about them, we first estimate the responses of cumulative GDP and cumulative government spending (G). Formally, we estimate

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

where \tilde{x}_t is either cumulative GDP ($\tilde{x}_t = \sum_{h=0}^H \text{GDP}_{t+h}$) or cumulative G ($\tilde{x}_t = \sum_{h=0}^H G_{t+h}$). This yields estimates $\tilde{\beta}_{\text{GDP}} = \sum_{h=0}^H \beta_{\text{GDP}}^h$, $\tilde{\beta}_G = \sum_{h=0}^H \beta_G^h$, with β_{GDP}^h and β_G^h being the coefficients in (3.4). The coefficients $\tilde{\alpha}_x$, $\tilde{\gamma}_x$, $\tilde{\delta}_x$, $\tilde{\zeta}_x$ are analogously related to (3.4). These estimates allow us to estimate the fiscal multiplier in (4.1).

To obtain a covariance matrix for the IV estimates $\hat{\vartheta} = (\tilde{\beta}_{\text{GDP}}, \tilde{\beta}_G, \tilde{\gamma}_{\text{GDP}}, \tilde{\gamma}_G)'$, we estimate the two regressions (i.e., for GDP and G) jointly via seemingly unrelated regressions. For our baseline inference, we use the [Driscoll and Kraay \(1998\)](#) covariance estimator, allowing for serial correlation and cross-correlation between GDP and G. We use the covariance matrix to compute standard errors for the fiscal multiplier by applying the Delta method to the fiscal multiplier in (4.1).

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. \(2019\)](#). We build the test based on two sets of regressions. First, consider the reduced-form regressions

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

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

$$\begin{aligned} \varepsilon_t^s &= \tilde{\alpha}^{fs1} + \tilde{\beta}^{fs1} \varepsilon_t^s + \tilde{\gamma}^{fs1} \varepsilon_t^s (\text{Hawk}_t^{IV} - \overline{\text{Hawk}}^{IV}) \\ &\quad + \tilde{\delta}^{fs1} (\text{Hawk}_t^{IV} - \overline{\text{Hawk}}^{IV}) + \tilde{\zeta}_x^{fs1} Z_{t-1} + \tilde{v}_{t+j}^{fs1}, \end{aligned} \quad (\text{G.3})$$

$$\begin{aligned} \varepsilon_t^s (\text{Hawk}_t - \overline{\text{Hawk}}) &= \tilde{\alpha}^{fs2} + \tilde{\beta}^{fs2} \varepsilon_t^s + \tilde{\gamma}^{fs2} \varepsilon_t^s (\text{Hawk}_t^{IV} - \overline{\text{Hawk}}^{IV}) \\ &\quad + \tilde{\delta}^{fs2} (\text{Hawk}_t^{IV} - \overline{\text{Hawk}}^{IV}) + \tilde{\zeta}_x^{fs2} Z_{t-1} + \tilde{v}_{t+j}^{fs2}, \end{aligned} \quad (\text{G.4})$$

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

$$AR(\vartheta_0) = \hat{g}(\vartheta_0)' \hat{\Omega}(\vartheta_0)^{-1} \hat{g}(\vartheta_0), \quad (\text{G.5})$$

$$\text{with } \hat{g}(\vartheta_0) = \hat{\rho} - \hat{\Pi}\vartheta_0, \quad (\text{G.6})$$

$$\text{and } \hat{\Omega}(\vartheta_0) = \hat{\mathbb{E}}[\varrho \varrho'] - \hat{\mathbb{E}}[\varrho \vartheta_0' \Pi'] - \hat{\mathbb{E}}[\Pi \vartheta_0 \varrho'] + \hat{\mathbb{E}}[\Pi \vartheta_0 \vartheta_0' \Pi'], \quad (\text{G.7})$$

where hats denote respective 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 $AR(\vartheta_0) \xrightarrow{d} \chi^2(4)$, since ϑ is 4×1 , 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.

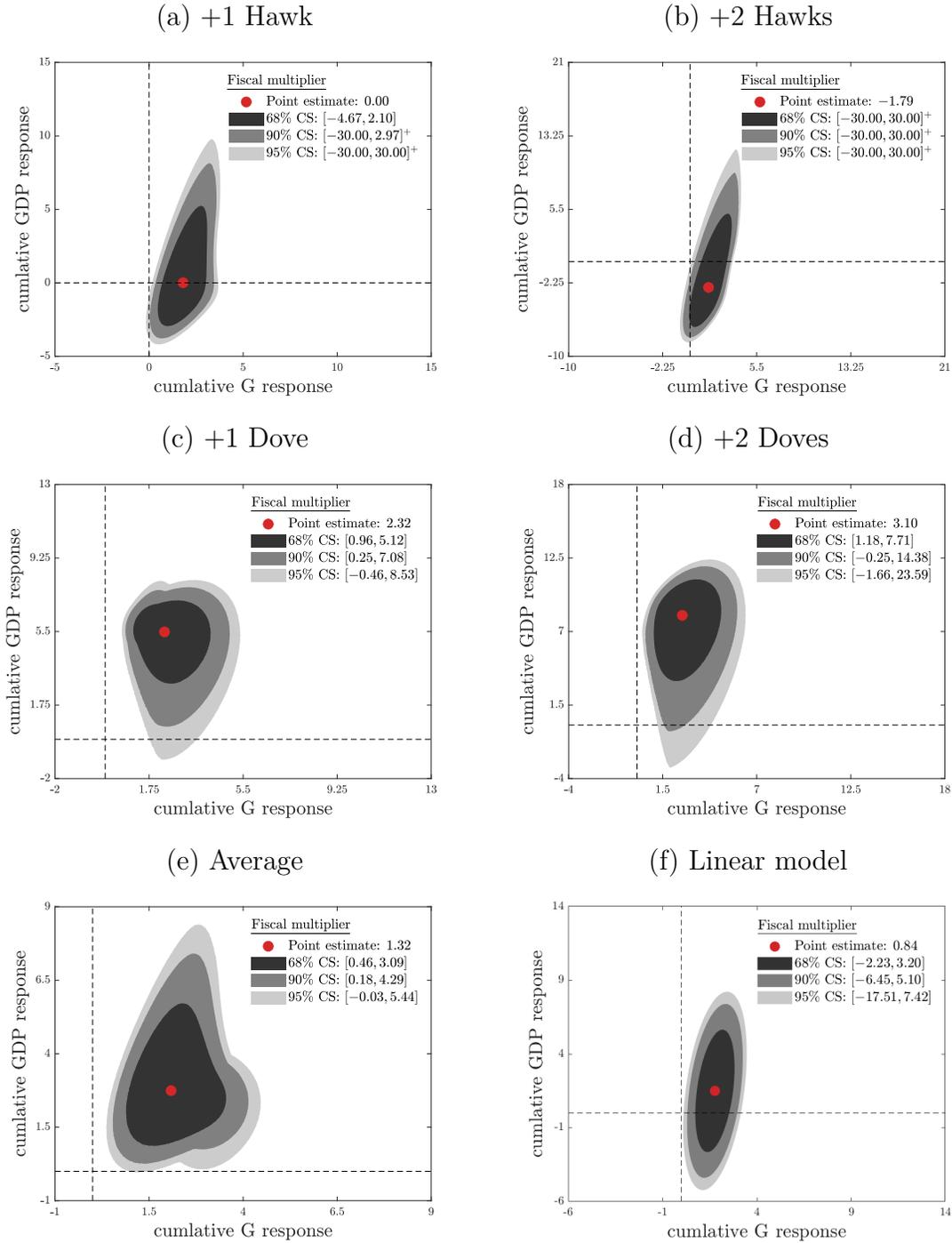
1. Define set Θ that contains the confidence region of $FM^H(\chi)$.
2. Define discrete set $\Theta_N \subset \Theta$ that contains N vectors of ϑ .
3. Construct the set $CS^\vartheta = \{\vartheta \in \Theta_N \mid AR(\vartheta) \leq c_{1-\alpha, \chi^2(4)}\}$.
4. Compute the confidence set for the fiscal multiplier as

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

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 . The set Θ is then defined by 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 (G.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 .⁵⁷ 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.

⁵⁷For the multiplier in 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 $\hat{\vartheta}_i < 0$.

Figure G.4: Anderson-Rubin confidence sets for four-year fiscal multipliers

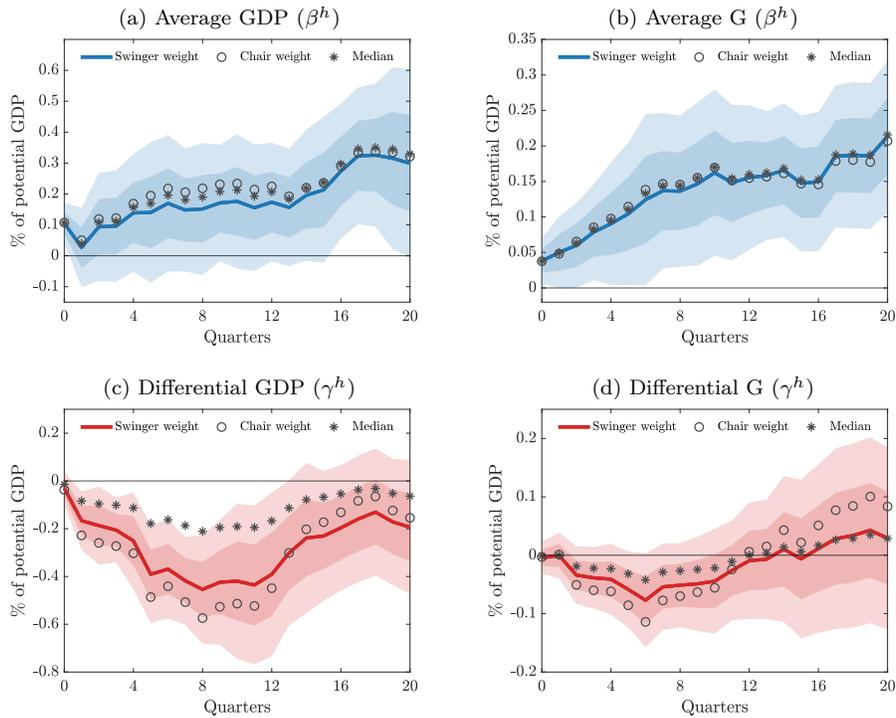


Notes: This figure shows Anderson-Rubin type confidence sets for the cumulative four-year fiscal multiplier. We depict the numerator and denominator of the multiplier on the vertical and horizontal axis, respectively. The shaded areas depict the confidence sets and various levels of significance. The 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 $\tilde{\gamma}_{GDP} = \tilde{\delta}_{GDP} = \tilde{\gamma}_G = \tilde{\delta}_G = 0$.

Appendix H Sensitivity analysis

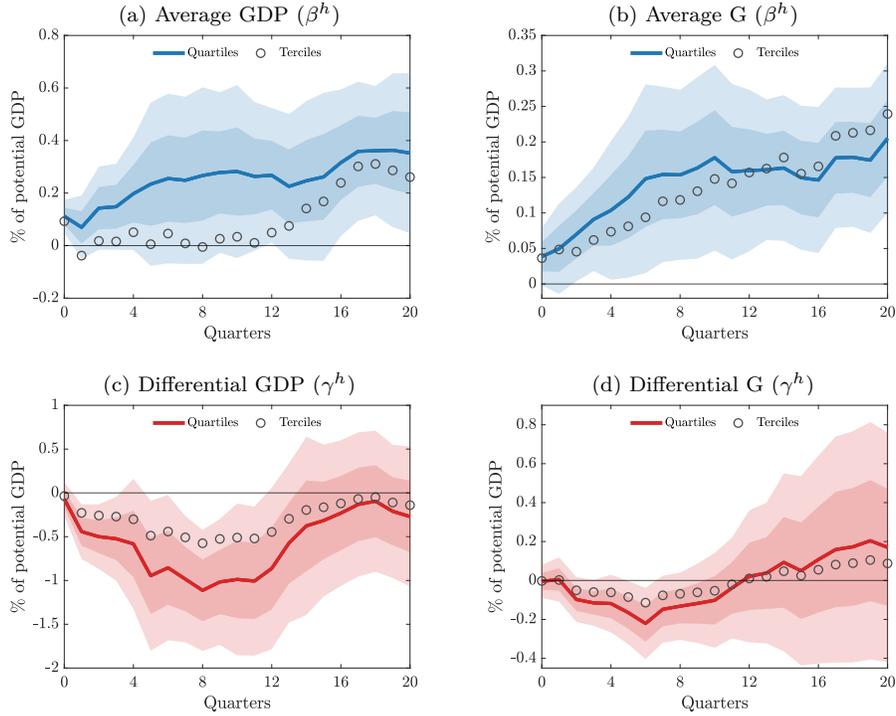
This appendix contains the results of our sensitivity analysis in Section 4.5. Unless specified otherwise, we present the responses of real GDP and real government spending (G) to an expansionary military spending shock, corresponding to one percent of GDP, conditional on systematic monetary policy ($Hawk_t$). We show IV estimates based on the local projection framework (3.4)-(3.5) 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 shaded areas indicate 68% and 95% confidence bands using Newey-West standard errors.

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



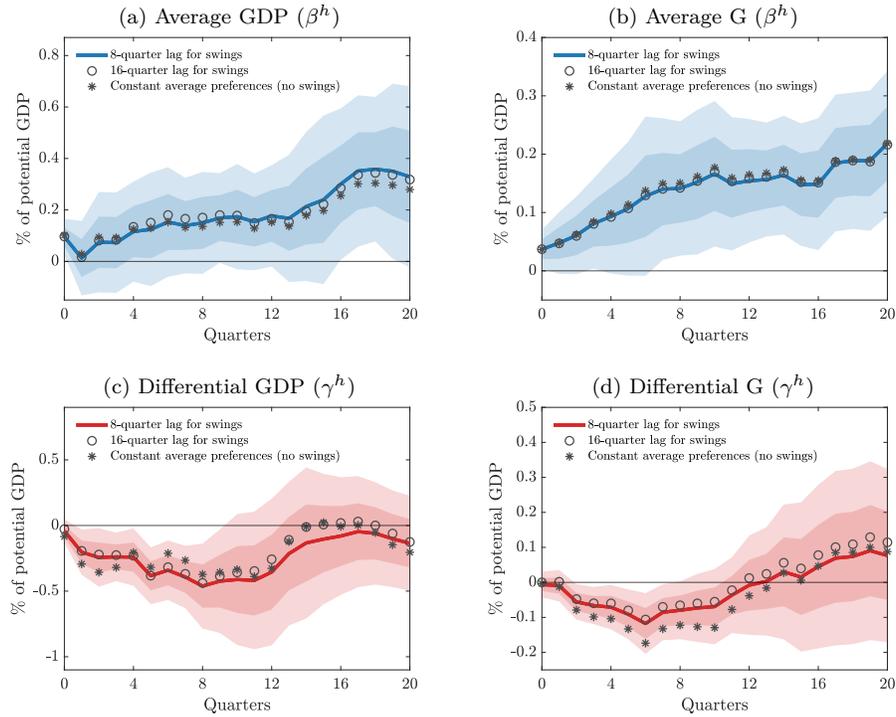
Notes: We use three variants of $Hawk_t$. Swinger weight: We do not discriminate between swingers and consistent members. Chair weight: We assign the preferences of the Fed Chair twice the weight of an ordinary member when aggregating to $Hawk_t$. Median: We aggregate the cross-section of FOMC members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

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



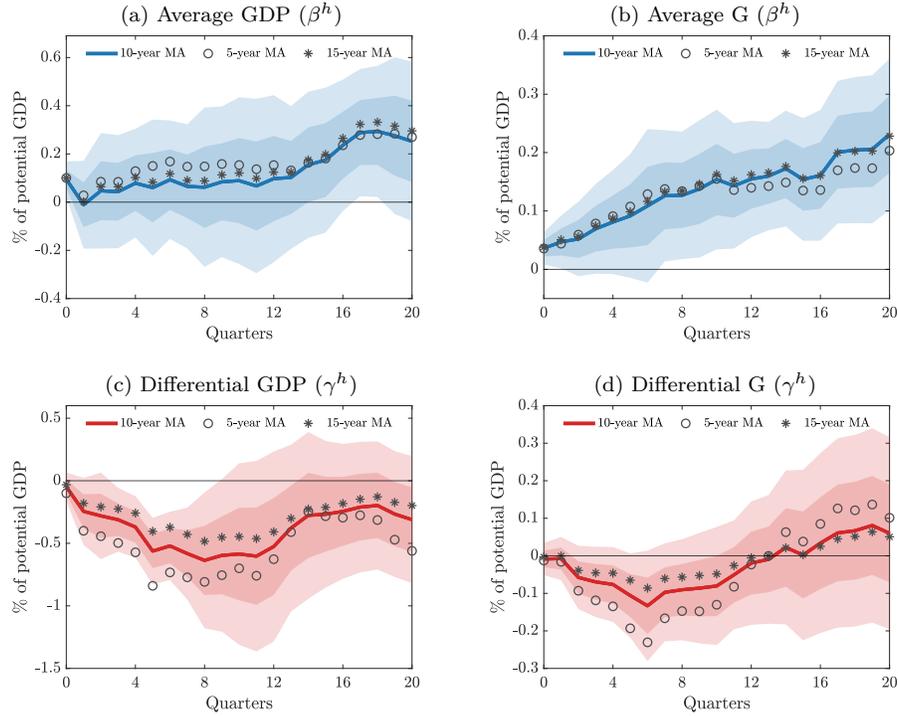
Notes: We use two discrete variants of $Hawk_t$. We define that the discrete $Hawk_t$ equals -1 if $Hawk_t$ falls below the first quartile or tertile of the distribution of $Hawk_t$ over time, +1 if above the highest quartile or tertile, and zero else.

Figure H.3: Responses of GDP and government spending, alternative IVs



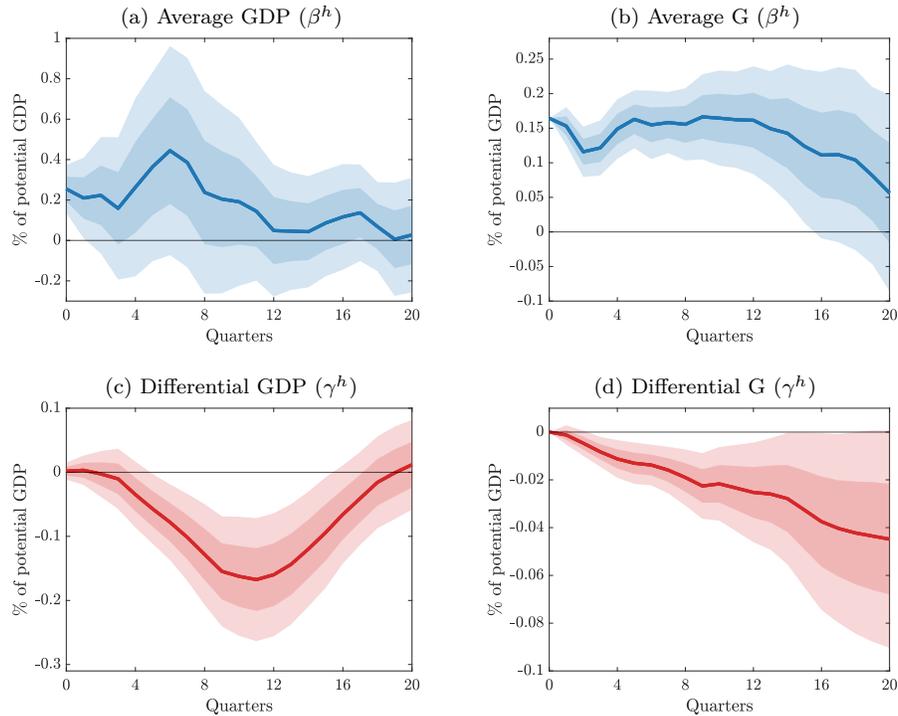
Notes: We use an alternative definition of the instrumental variable $Hawk_t^{IV}$ where swings affect the individual preference only 8 or 16 quarters after the date of the swing, or where no swing occurs because we set the individual preference to the average, rendering them time-invariant.

Figure H.4: Responses of GDP and government spending, accounting for trends



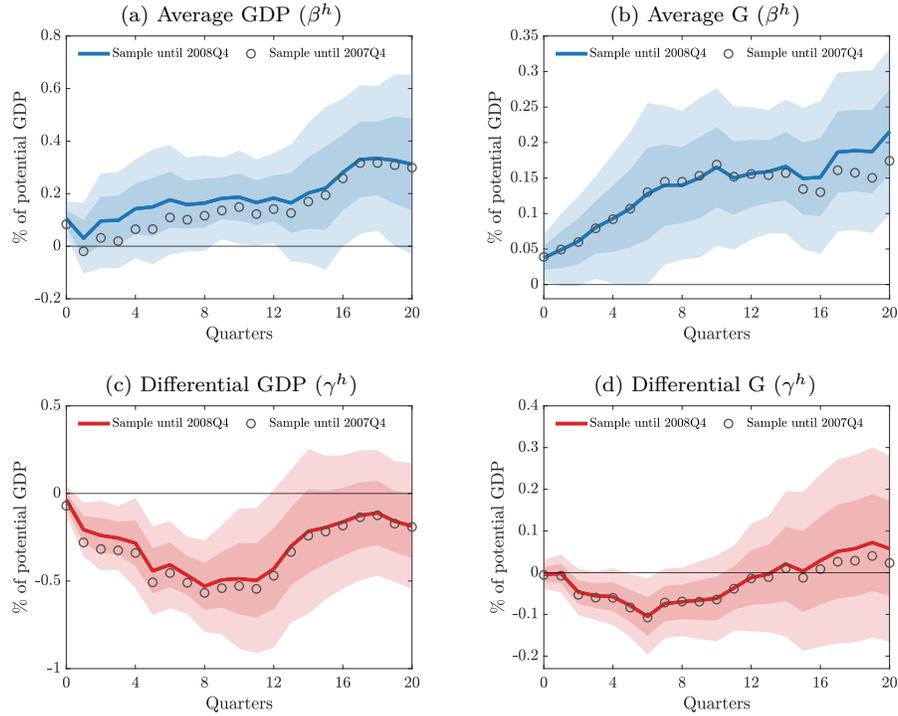
Notes: We use three variants of $Hawk_t$ where we subtract the backward-looking 5, 10, or 15-year moving average from $Hawk_t$ prior estimation.

Figure H.5: Responses of GDP and government spending, Blanchard-Perotti shock



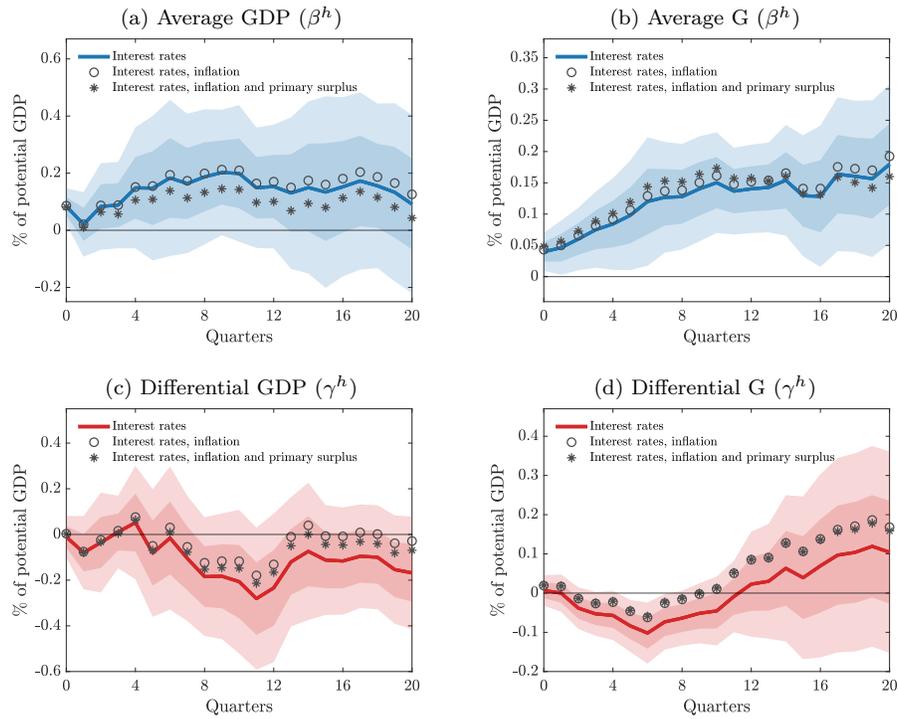
Notes: The shock is contemporaneous G, conditional on controls that include four lags of real GDP and real government spending, as well as the projected growth rate of real government spending. The projected growth rate is taken from the Survey of Professional Forecasters and is available from 1969 onward,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

Figure H.6: Responses of GDP and government spending, accounting for the ZLB



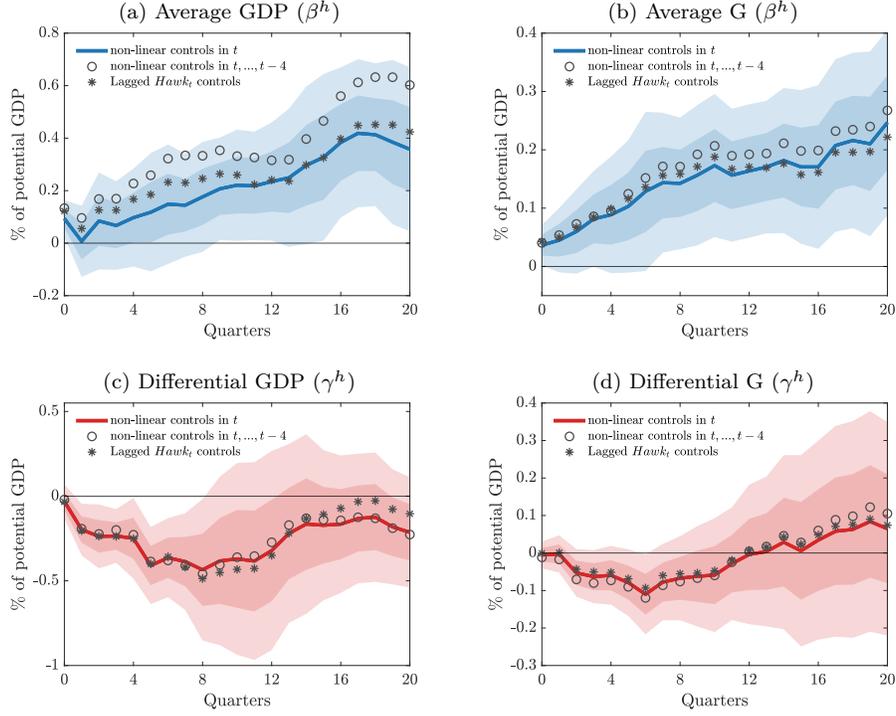
Notes: We use a sub-sample that ends either in 2008Q4 or 2007Q4 to exclude the ZLB, or both the ZLB and the Great Recession.

Figure H.7: Responses of GDP and government spending, additional controls



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 H.8: Responses of GDP and government spending, non-linear controls



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 $t, \dots, t-4$: Augments the control vector by also including and instrumenting lagged interaction terms, i.e. $Hawk_{t-i} \times C_{t-i}$ with $i = 1, \dots, 4$ 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 H.1: Cumulative 4-year government spending multipliers, Discrete Hawk-Dove balance

Specification	Multipliers across regimes			p-values for differences		
				across regimes		
	Hawkish	Average	Dovish	Hawkish vs. Dovish	Hawkish vs. Average	Average vs. Dovish
Quartiles	-6.002 (10.343)	1.727 (0.775)	4.814 (2.774)	0.264	0.460	0.201
Tertiles	-3.481 (6.227)	0.490 (0.772)	2.835 (1.083)	0.336	0.488	0.047

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 coefficients are estimated using a cumulative version of the local projection framework (3.4)-(3.5) 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).

Table H.2: Cumulative 4-year government spending multipliers, Robustness

Specification	Multipliers across regimes			p-values for differences across regimes		
	+2 Hawks	Average	+2 Doves	+2 Hawks	+2 Hawks	Average
				vs.	vs.	vs.
				+2 Doves	Average	+2 Doves
Baseline	-1.786 (2.636)	1.315 (0.478)	3.105 (1.167)	0.122	0.245	0.041
BP shock	0.849 (1.068)	1.342 (0.839)	1.730 (0.699)	0.077	0.091	0.062
Aggregation schemes						
Median	0.426 (0.569)	1.419 (0.546)	2.232 (0.816)	0.043	0.065	0.036
Chair weight	-1.671 (2.222)	1.538 (0.664)	3.468 (1.518)	0.070	0.175	0.078
Swinger weight	-1.597 (2.168)	1.267 (0.554)	3.046 (1.144)	0.090	0.180	0.043
Accounting for trends						
5-year MA	-11.458 (59.772)	1.290 (1.143)	3.770 (2.106)	0.801	0.829	0.210
10-year MA	-4.716 (10.012)	0.844 (0.868)	3.238 (1.224)	0.439	0.553	0.085
15-year MA	-2.093 (3.051)	0.987 (0.430)	2.977 (0.884)	0.137	0.279	0.035
Accounting for swings in the IV						
8-quarter lag	-1.534 (2.977)	1.220 (0.541)	2.622 (1.069)	0.233	0.345	0.071
16-quarter lag	-0.728 (2.675)	1.239 (0.599)	2.560 (1.455)	0.338	0.449	0.239
Average preferences	-1.556 (4.946)	1.070 (0.656)	1.872 (1.224)	0.549	0.580	0.494

(Table continues on the next page)

Table H.2 (continued): Cumulative 4-year government spending multipliers, Robustness

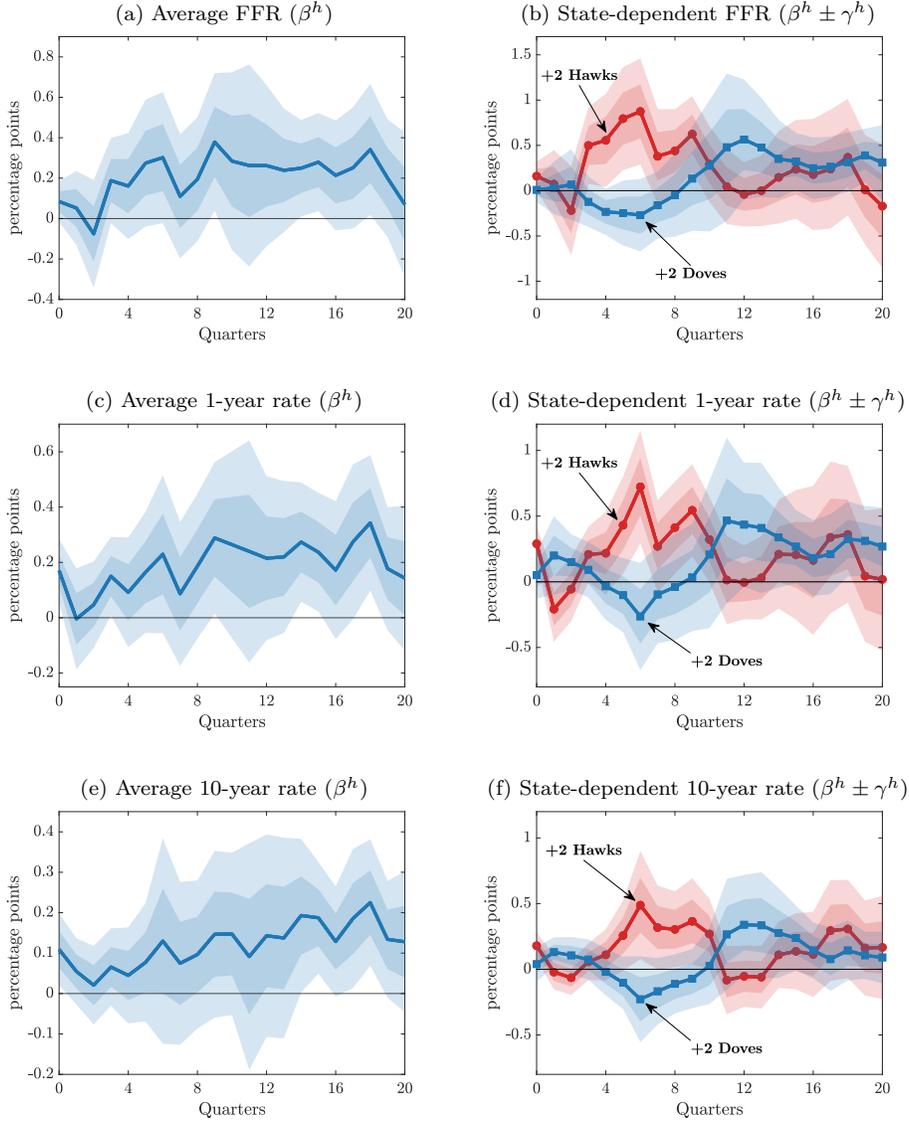
Specification	Multipliers across regimes			p-values for differences		
				across regimes		
	+2 Hawks	Average	+2 Doves	+2 Hawks vs. +2 Doves	+2 Hawks vs. Average	Average vs. +2 Doves
Accounting for the ZLB						
End sample '08	-1.999 (2.672)	1.306 (0.513)	3.099 (1.138)	0.107	0.225	0.032
End sample '07	-3.378 (4.513)	0.922 (0.531)	3.016 (1.137)	0.204	0.344	0.031
Additional controls						
Interest rates	0.390 (1.269)	1.258 (0.690)	1.861 (0.760)	0.301	0.322	0.351
Interest rates, inflation	0.738 (0.871)	1.260 (0.646)	2.055 (1.046)	0.327	0.306	0.380
Interest rates, inflation, surplus	0.654 (1.107)	1.324 (0.855)	2.210 (1.437)	0.373	0.328	0.463
Non-linear controls						
in t	-1.019 (2.642)	1.436 (0.553)	2.830 (1.129)	0.206	0.356	0.067
in $t, \dots, t - 4$	0.371 (2.293)	2.026 (0.646)	3.033 (1.157)	0.344	0.486	0.141
Lagged $Hawk_t$	-0.575 (1.734)	1.659 (0.549)	3.216 (1.213)	0.118	0.230	0.070

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 (3.4)-(3.5) 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 G for details. The various exercises correspond to the impulse responses presented in Figures H.1-H.8, see the respective figure notes for details.

Appendix I Additional results for Section 5

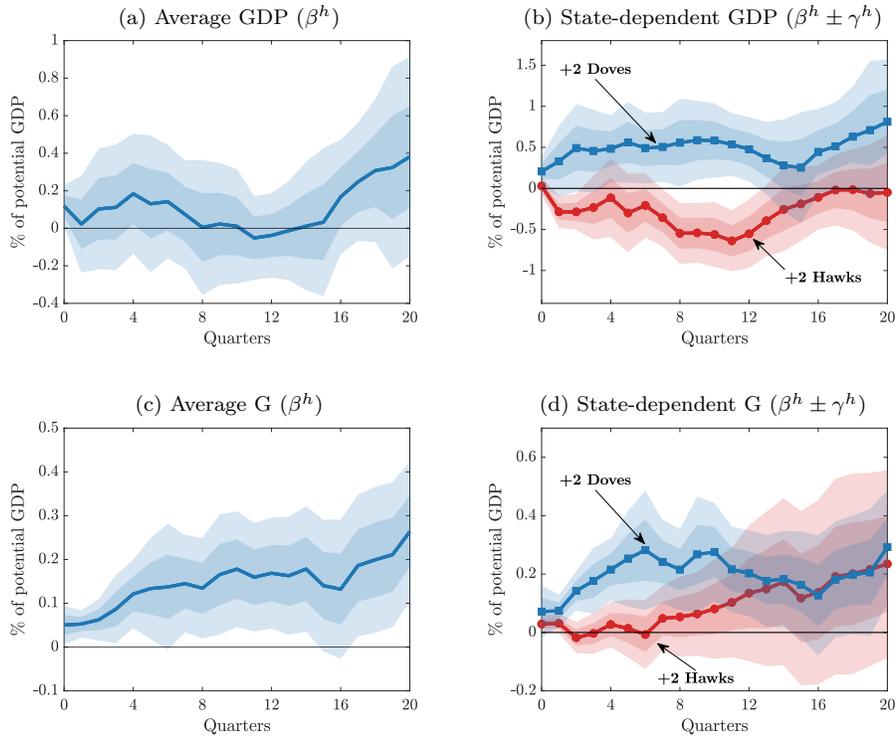
This appendix contains additional findings discussed in the main text.

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



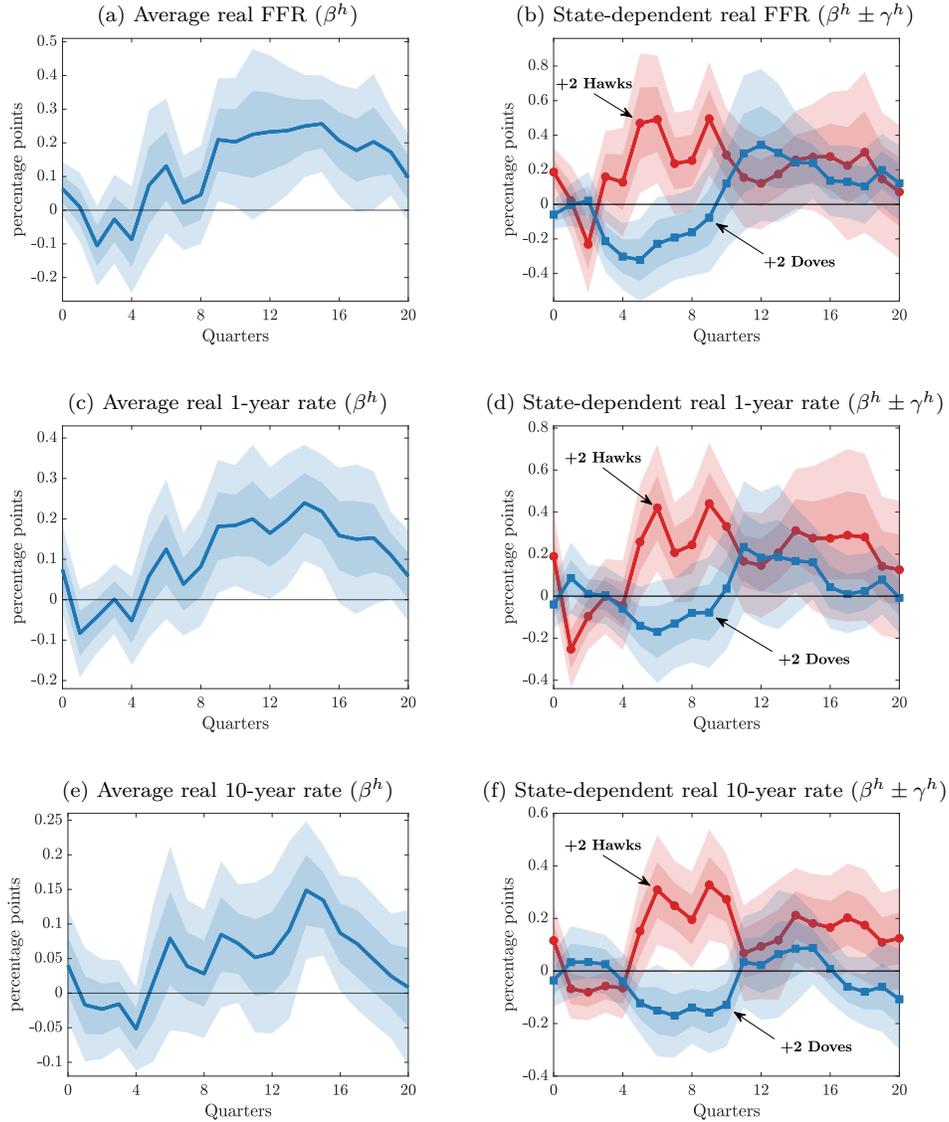
Notes: The figure shows responses of the federal funds rate (FFR), as well as the 1-year and 10-year treasury yields to an expansionary military spending shock, corresponding to one percent of GDP, conditional on systematic monetary policy ($Hawk_t$). All outcomes are annualized interest rates. We show IV estimates based on the local projection framework (3.4)-(3.5) 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. We set the military spending shocks occurring in Q3 or Q4 to zero.

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



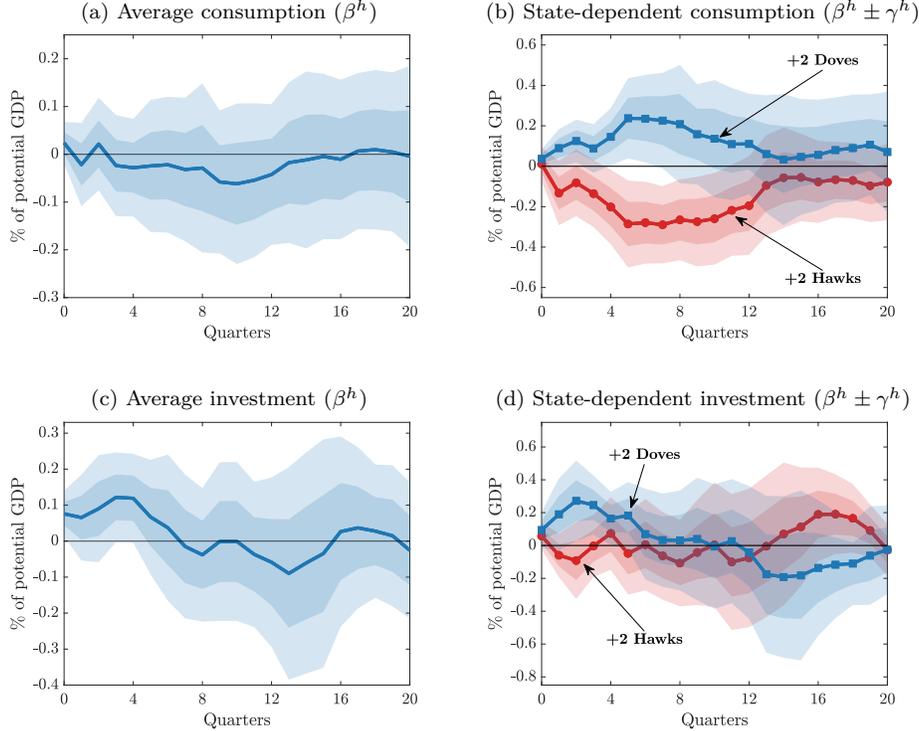
Notes: The figure shows responses of real GDP and real government spending (G) to an expansionary military spending shock, corresponding to one percent of GDP, conditional on systematic monetary policy ($Hawk_t$). We show IV estimates based on the local projection framework (3.4)-(3.5) as specified in Section 4.1. 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. We set the military spending shocks occurring in Q3 or Q4 to zero.

Figure I.3: Responses of real interest rates



Notes: The figure shows responses of the real federal funds rate (FFR), as well as the 1-year and 10-year real treasury yields 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 (3.4)-(3.5) as specified in Section 5.1. All outcomes are annualized ex-ante real interest rates which we compute as nominal rate minus one-year ahead inflation expectations according to the Livingston Survey, see Appendix B for details. 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.

Figure I.4: Decomposing the GDP response, private spending



Notes: The figure shows responses of real private consumption and real private investment 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 (3.4)-(3.5) as specified in Section 4.1. 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. We modify the control vector to include four lags of consumption, investment, and government spending, as well as the shock and a residual component of GDP, which we compute as GDP minus consumption, investment, and government spending.

Appendix J 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 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.⁵⁸ 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)

In this period, the FOMC composition was hawkish on average. This helped the hawkish Chairman Martin to reach a consensus for tighter policy to act preemptively against inflationary pressures.

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

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,

⁵⁸The 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).

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, in their judgment, 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 in order to prevent the risk of inflation. However, the tightening signal by the Fed 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 the 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 hitting 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 the policy on December 12, 1967.

In the period between 1965 and 1967, the FOMC is categorized as dovish on average. Both, 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 modestly but soon erred toward loose policy.