Journal of Macroeconomics 86 (2025) 103721

Contents lists available at ScienceDirect

Journal of Macroeconomics

journal homepage: www.elsevier.com/locate/jmacro

Research Paper

What do we know about estimating government spending multipliers?[☆]

Taewoong Jo [a,1], Jihye Kang [b,*], Joonyoung Hur [a]

[a] Department of Economics, Sogang University, 35, Baekbeom-ro, Mapo-gu, Seoul, 04107, Republic of Korea
[b] Research Department, Bank of Korea, 39, Namdaemun-ro, Jung-gu, Seoul, 04531, Republic of Korea

ARTICLE INFO

JEL classification:
C11
E20
E62

Keywords:
Government spending
Vector autoregressions
Local projections

ABSTRACT

Using the DSGE model as the data-generating process (DGP), we assess how three key modeling choices influence government spending multiplier estimates: (1) the econometric method—vector autoregressions (VARs) versus local projections (LPs); (2) the identification strategy for government spending shocks—such as recursive, Blanchard–Perotti (BP), or forecast error (FE) methods; and (3) the variable transformation—log versus Gordon–Krenn (GK). Our results demonstrate that even when using the same data set, these choices can lead to substantially different multiplier estimates. Furthermore, we find that the choice of econometric method should align with the shock identification strategy and targeted estimation horizon. For the short-run, LP method produces the most accurate government spending multipliers when the true shock sequence is known. When there is no strong candidate for the shock, BP-type shocks are preferable, with both VAR and LP methods being more suitable for short-run analysis, while VAR models yield more reliable estimates for long-run horizons. Additionally, using the GK transformation instead of the log transformation reduces the upward bias commonly observed in VAR and LP estimates.

1. Introduction

There has been growing interest in the macroeconomic effects of fiscal policy since the onset of the Global Financial Crisis of 2008–09. This is mainly because, as monetary policy hit the zero lower bound, fiscal policy became a crucial proactive measure to prevent the economy from falling into recession or crisis. One of the primary issues in subsequent research has been whether fiscal policy, particularly government spending, stimulates private consumption or investment, and thereby increases output—often referred to as "government spending multipliers". However, as Ramey (2019) notes, despite the rapid advancement of research in this area, there remains a wide range of estimates for government spending multipliers.

One of the primary reasons for the variation in government spending multiplier estimates is that time-series methods, unlike DSGE models, are heavily dependent on both the data and the econometric techniques used. For instance, as illustrated in Ramey's 2019 Table 1, multiplier estimates vary significantly based on the specifications of dataset — whether the analysis focuses on a single country or a panel of countries, the time span of the data, variables included in the model, etc. On the other hand,

☆ We would like to thank Neville Francis (the editor) and two anonymous referees whose comments were particularly helpful. We also thank Jaeho Kim, Sungwon Lee, and seminar participants at Korea University, the Bank of Korea, and the Korea Development Institute for their helpful comments. Any remaining errors are the responsibility of the authors. The views expressed herein are those of the authors and do not necessarily represent those of the Bank of Korea.

* Corresponding author.
E-mail addresses: taewoongjo@sogang.ac.kr (T. Jo), kjh01@bok.or.kr (J. Kang), joonyhur@sogang.ac.kr (J. Hur).
[1] First author.

https://doi.org/10.1016/j.jmacro.2025.103721
Received 7 January 2025; Received in revised form 9 July 2025; Accepted 13 September 2025
Available online 27 September 2025

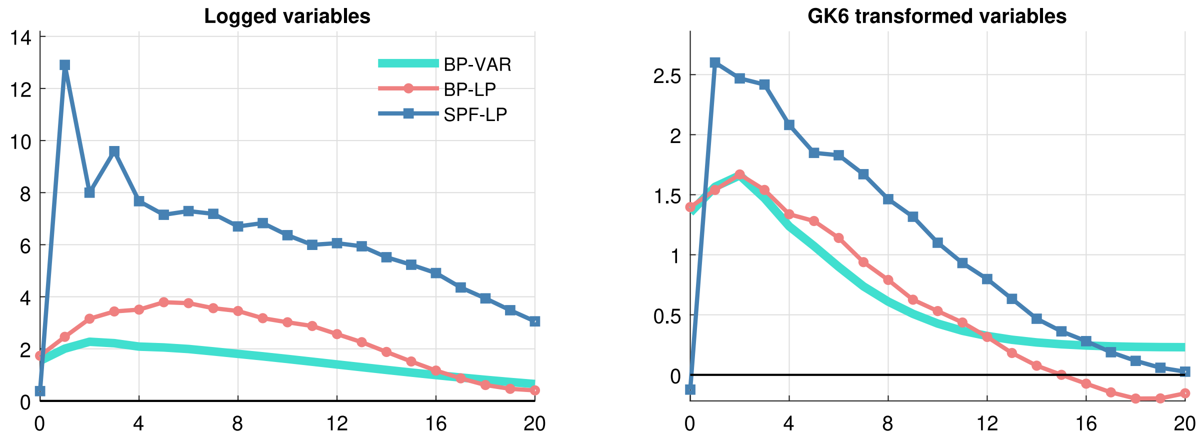

Fig. 1. Present-value government spending multipliers for output, using variables based on actual data: logarithmic transformations (left panel) and Gordon–Krenn transformations (right panel). In each panel, the solid line, the solid line with circles, and the solid line with squares represent the multipliers corresponding to the VAR model with Blanchard–Perotti (BP) identification, LP with the BP-identified shock sequence, and LP with the shock sequence based on the SPF dataset, respectively. The x-axis measures quarters.

methodological choices — such as the econometric method (VARs or LPs) and shock identification strategies (e.g., recursive) — also play a critical role in determining the multiplier estimates. Therefore, time-series approaches are expected to produce comparable multiplier estimates only if applied to the same dataset with similar econometric methodology choices.

However, in practice, there are at least three dimensions of applying time-series approaches that lead to different multiplier estimates. The first dimension relates to the methodological differences between VARs and LPs. Since Jordà's (2005) seminal work, LPs have been widely used in the field of empirical macroeconomics. Accordingly, there has been a surge in research comparing LPs with the VAR methodology, with notable examples including Plagborg-Møller and Wolf (2021) and Li et al. (2024). Nevertheless, these studies primarily focus on either TFP shocks (e.g, Meier, 2005) or monetary policy shocks (e.g., Plagborg-Møller and Wolf, 2021), and there is a lack of literature examining how the two methodologies differ in the context of government spending multipliers.

The second dimension relates to the identification of government spending shocks. Existing studies propose several prominent methods for isolating exogenous changes in government spending. Blanchard and Perotti (2002) use the institutional structure, including legislative and implementation lags, to identify these exogenous changes. Meanwhile, another strand of literature relies on economic agents' forecasts of government spending. Auerbach and Gorodnichenko (2012, 2013) provide a typical example of this approach by utilizing the Survey of Professional Forecasters (SPF) dataset and the OECD's Outlook and Projections Database.

Lastly, It is also crucial to concern how to handle endogenous variables in the model, such as using log transformations or Gordon–Krenn (GK) transformations, as proposed by Gordon and Krenn (2010). Owyang et al. (2013) demonstrate that log-transformations of the model's endogenous variables when calculating government spending multipliers tend to produce an upward bias arising from the ad-hoc ex-post conversion factor, irrespective of the econometric models used. A number of subsequent studies use the GK transformation of variables instead [e.g., Barro and Redlick (2011) and Ramey and Zubairy (2018)].

Fig. 1 illustrates how multiplier estimates vary across different econometric specifications based on actual data. The results are derived from a 5-variable model similar to those in Perotti (2005) and Caldara and Kamps (2008), which extends the specifications of Fatás and Mihov (2001) and Blanchard and Perotti (2002).[2] The figure clearly shows that government spending multiplier estimates are highly sensitive to the choice of these three dimensions. The fact that discrepancies in the results emerge even when using the same dataset raises a natural question about which estimates are more reliable.

In this paper, we aim to address this question by evaluating the empirical significance of each dimension. We accomplish this by setting up a benchmark model economy using the DSGE model framework described in Leeper et al. (2017). The model is sufficiently detailed to incorporate key features relevant to the consequences of exogenous changes in government spending. It includes New Keynesian elements with real frictions, hand-to-mouth agents, and the interplay between private and public consumption, whether as complements or substitutes. Using the estimated model as the data generating process (DGP), we compare the model-implied government spending multipliers with those recovered econometrically.

Our approach builds on previous DGP-based comparisons of VARs and LPs, such as Meier (2005) who examine monetary policy, technology, and labor supply shocks in the Euro area using DSGE models as DGPs, and Brugnolini (2018) who focus on US monetary policy shocks using VAR-based DGPs. We extend this literature by applying the DGP approach to government

[2] More specifically, the endogenous variables are government spending, taxes, output, inflation, and the nominal interest rate, covering the period from 1983:Q3 to 2007:Q4. Government spending shocks are identified either using the method outlined by Blanchard and Perotti (2002) or by relying on the SPF dataset. The lag length is set to 4. Present-value multipliers are reported.

spending multipliers and systematically evaluating three critical dimensions: variable transformation methods (log vs. Gordon–Krenn), econometric approaches (VARs vs. LPs), and shock identification strategies. A salient feature of this approach is that, with the estimated parameters of the model, we can construct the sequence of exogenous shocks that is consistent with the model. The model-implied sequence of government spending shocks is then used to assess the importance of accurate shock identification in calculating multipliers, particularly within the LP framework.

Two key takeaways emerge from our empirical framework. First, when estimating government spending multipliers, one must consider not only the choice of methodology (VAR or LP) but also how closely the identified shocks approximate the true shocks and the targeted estimation horizon (short, medium, or long run). In the short run, GK-transformed LP estimates are closer to the model-implied counterparts compared to GK-transformed VAR estimates. However, for the medium to long run, this conclusion should be revised, as the VAR specification with GK-transformed variables outperforms the LP approach. It is important to note that accurate shock identification is especially critical in the short run. Specifically, the GK-transformed LP method produces the most accurate government spending multipliers when the true shock sequence is known. This finding highlights a key implication for estimating government spending multipliers: having access to a precise shock sequence enhances the accuracy of short-run estimates. However, obtaining an accurate shock sequence remains a significant challenge in practice. Our results suggest that VARs with shock identification methods as described by Blanchard and Perotti (2002) or using the Blanchard–Perotti shock within an LP model can serve as viable alternatives for estimating short-run government spending multipliers.

Second, regarding the transformation methods for the model's endogenous variables, our findings support (Owyang et al., 2013) that the GK transformation approach yields more precise government spending multipliers than the logarithmic transformation method in both VAR and LP frameworks. Using log-transformed variables tends to introduce an upward bias in the multiplier estimates arising from the ad-hoc conversion factor defined by the sample average of the ratio between GDP and government spending, irrespective of the econometric specification. Notably, this upward bias is not confined to the VAR framework; it persists even when the true shock sequence is used in the LP framework.

The remainder of the paper is organized as follows. Section 2 introduces the methodology: data generating process, econometric models, shock identification and variable transformation methods. Section 3 presents empirical results. Section 4 conducts robustness checks. Section 5 concludes.

2. Methodology

In this section, we present our data generating process in Section 2.1. Sections 2.2, 2.3, and 2.4 summarize the three dimensions of the specifications of government spending multiplier estimation: econometric methods (VAR vs. LP/LP-IV), shock identifications (recursive/BP vs. FE), and variable transformation methods (log vs. GK).

2.1. Data generating process

DSGE Model overview. This section describes the new Keynesian DSGE (NK-DSGE) model used as the DGP for this article. The model employed in this paper is a conventional new Keynesian model as utilized by Leeper et al. (2017). Key model features include the presence of hand-to-mouth agents and households' utility from composite consumption—where private and public consumption can function as complements or substitutes. The model also incorporates several real rigidities, such as consumption habit formation, variable capacity utilization, and investment adjustment costs, which are crucial for enhancing the model's fit to the data. The market structure of the model is associated with monopolistic competition in both product and labor markets.[3] This model is equipped with 11 exogenous shocks: technology, investment-specific, preference, price markup, and wage markup shocks, together with five fiscal policy disturbances and a monetary policy disturbance.

Regarding the generation of model-implied series of government spending shocks, government spending G is governed by the following process:

$$\hat{G}_t = \rho_G \hat{G}_{t-1} - (1-\rho_G)\gamma_G \hat{s}^b_{t-1} + \epsilon^G_t, \tag{1}$$

where a hat ($\hat{\ }$) denotes the log deviation from the steady state. $s^b_{t-1} \equiv \left(P^B_{t-1}B_{t-1}\right) / \left(P_{t-1}Y_{t-1}\right)$ and $\epsilon^G_t \sim i.i.d.\ N(0,\sigma_G)$, where B_t is a portfolio of long-term nominal government bonds, and P^B_t denotes their price.

We estimate the DSGE model using Bayesian inference methods to construct the posterior distribution of the parameters, which combines the likelihood function and prior information. The model is estimated using 11 observables, representing US quarterly time series data from 1983:Q3 to 2007:Q4.[4] These observables include consumption growth, investment growth, real wage growth, hours worked, CPI inflation rate, the nominal interest rate, and the growth rates of fiscal variables—capital, labor, and consumption tax revenues, government spending, and transfers. A full description of the data used is provided in Appendix B. Prior and Posterior for the selected model parameters are shown in Table C.1 in Appendix C.

Model-implied government spending shock. The model-implied government spending shock sequence serves as the benchmark "true" shock sequence for our analysis. With the estimated parameters, the sequence of exogenous shocks in the model can be reconstructed

[3] We provide summary and estimation of the model in Appendix C. For a detailed description of the model, we refer to Leeper et al. (2017).

[4] The choice of this sample period aims to avoid potential regime-switching issues in monetary policy. It is well-established in the existing literature that monetary policy in the pre-mid-1980s did not sufficiently react to inflation, leading to equilibrium indeterminacy (see Clarida et al. (2000), Lubik and Schorfheide (2004), and Bianchi (2013), among many others). Additionally, the sample ending before 2008 is to exclude the zero lower bound period in the US.

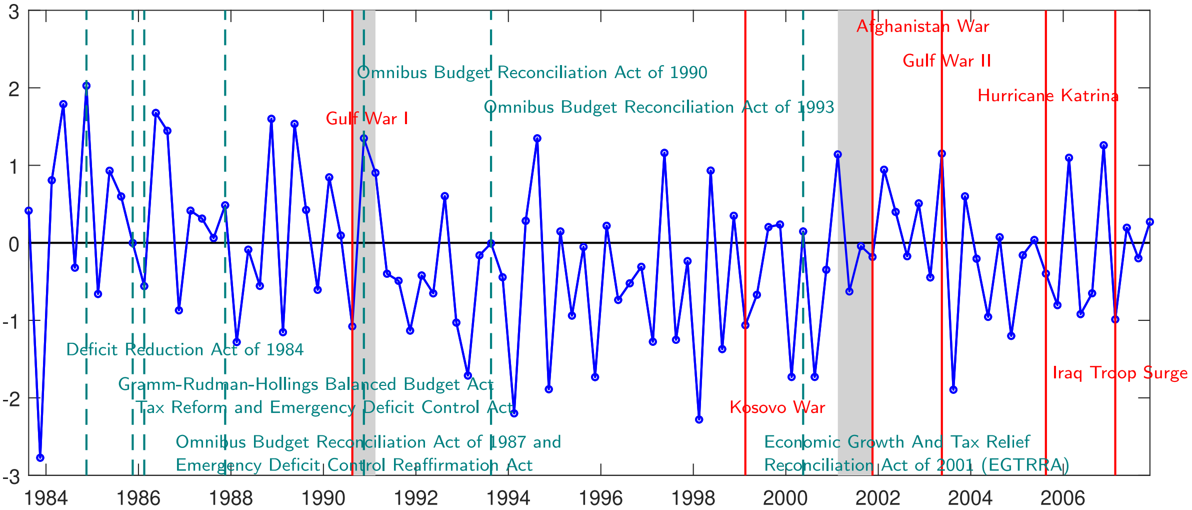

Fig. 2. Model-implied government spending shock sequence. Vertical lines indicating the dates of fiscal and geopolitical episodes that incurred changes in government spending. Shaded areas indicate NBER recession dates.

using a Kalman smoother. Fig. 2 plots the median estimates for the Kalman-smoothed government spending shock sequence, along with notable events in the history of US fiscal policy. Specifically, the figure employs solid and dashed vertical lines to denote the timing of expansionary and contractionary fiscal policies, respectively.

Overall, many of the spikes and dips in the recovered shock sequence align well with the timing of fiscal policy changes. For instance, the shock sequence either begins to spike up in the aftermath of wars (Gulf War I, Kosovo War, Afghanistan War, and Iraq Troop Surge) or is already spiked up, as in the case of Gulf War II. A similar tendency is observed for the contractionary fiscal policy cases.

Government Spending Multiplier. Government spending multipliers are commonly used to summarize the effects of fiscal policy. Following the approach outlined by Mountford and Uhlig (2009), we calculate the cumulative multipliers[5] for output as follows:

$$\text{Cumulative Multiplier } (q) = \frac{\sum_{j=0}^{q} E_t \Delta Y_{t+j}}{\sum_{j=0}^{q} E_t \Delta G_{t+j}}, \tag{2}$$

where E_t denotes an expectation conditional on the information available up to time t. ΔY_t and ΔG_t represent level changes in output and government spending relative to their steady-state values, respectively.

Fig. 3 reports the model-implied cumulative multipliers for output. The size of the multipliers is compatible with that in Leeper et al. (2017) for a similar sample period (1982–2007), with the impact multiplier exceeding unity in terms of mean estimates and monotonically decaying over time.[6]

2.2. Econometric methods for multiplier estimation

In this paper, we estimate two econometric models based on the same dataset, which consists of model-implied series obtained by conditioning on a particular sequence of shocks in the aforementioned NK-DSGE model. More specifically, the model-implied series are constructed by fixing the DSGE model's parameters at their posterior median estimates and applying a Kalman smoother.[7]

[5] While Mountford and Uhlig (2009) calculate present value multipliers that discount future effects, we compute cumulative multipliers to maintain consistency with LP-IV estimates presented in Section 2.2.2.

[6] Note, however, that the multiplier estimates in this paper tend to be slightly lower than those of Leeper et al. (2017). At impact, our median estimate (1.08) is lower than the estimate of 1.25 in Leeper et al. (2017). This tendency persists, as the one-year multipliers are 0.84 for our study and 1.01 for Leeper et al. (2017), respectively. As made explicitly in Leeper et al. (2017), certain parameters are crucial for the size of short-run government spending multipliers, including the degree of habit formation in consumption (θ), the share of non-savers (μ), and the substitutability between private and public consumption (α_G). We find that θ is comparable to the value in Leeper et al. (2017), whereas the fraction of rule-of-thumb households, μ, is estimated to be lower, and the degree of substitutability between private and public consumption is reduced (less complementarity), resulting in smaller increases in private consumption relative to increases in government spending. These differences in parameter estimates account for the discrepancies in the multiplier estimates between the two studies.

[7] However, a more comprehensive Bayesian analysis would perform experiments for each parameter draw, resulting in a set inference rather than a point inference. Accordingly, we implement the full Bayesian methodology and find that our main empirical results remain almost unaltered when thorough Bayesian inference is employed. The results associated with the full Bayesian methodology are reported in the companion Appendix D.1.

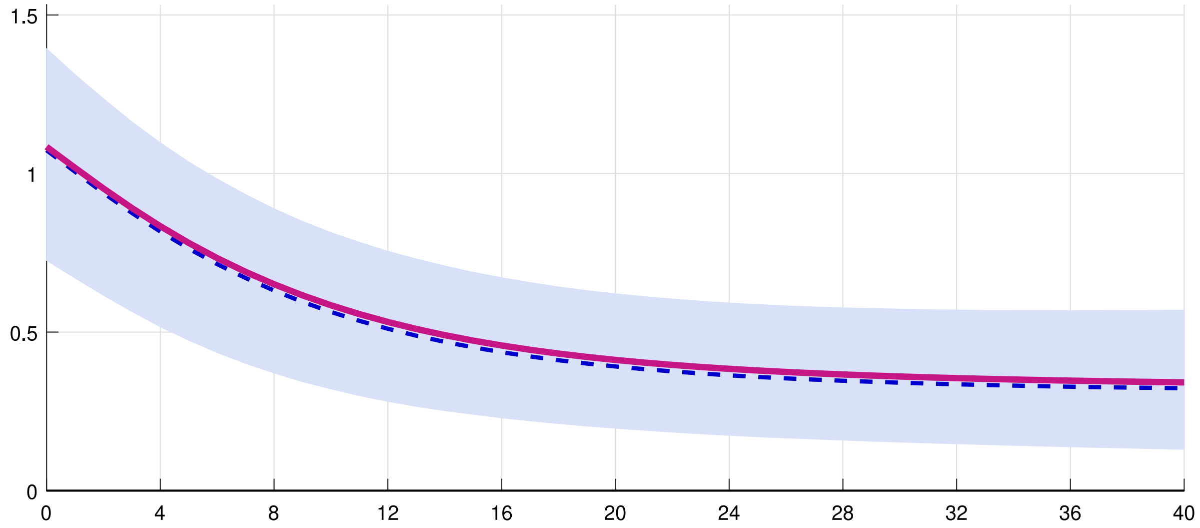

Fig. 3. The solid line indicates the cumulative government spending multipliers for output, evaluated at the median of the posterior parameter estimates. The dashed line and shaded area represent the posterior median cumulative multipliers and their 90 percent probability bands. The x-axis measures quarters.

2.2.1. *Vector Autoregressions (VARs)*

The first approach we consider is a VAR model. The reduced-form VAR model of the n-dimensional endogenous vector X_t is expressed as

$$X_t = \mu_0 + \mu_1 t + A(L)X_{t-1} + u_t, \tag{3}$$

where μ_0 is a constant, t is a linear time trend, $A(L)$ is a finite-order autoregressive lag polynomial, and u_t is an n-dimensional vector of reduced-form innovations with $E[u_t] = 0$, $E[u_t u_t'] = \Sigma_u$, and $E[u_t u_s'] = 0$ for $s \neq t$. Following the work of Blanchard and Perotti (2002), we set the lag length to be four quarters.

The endogenous variable X_t contains five variables: government spending, G_t; taxes, T_t; output, Y_t; inflation, π_t; and the nominal interest rate, R_t. Note that this particular set of variables is utilized in Perotti (2005) and Caldara and Kamps (2008) to assess the effectiveness of government spending in boosting output. In particular, previous VAR studies such as Rossi and Zubairy (2011) and Dupor and Li (2015) underscore the role of monetary policy in determining the size of government spending multiplier estimates. The inclusion of the interest rate is intended to control for monetary policy when fiscal policy changes occur. Based on Caldara and Kamps (2008), government spending shocks are identified through the recursive ordering as $X_t = [G_t, Y_t, \pi_t, T_t, R_t]'$. Arranging taxes following output and inflation can be rationalized by the fact that, when the tax rate remains constant, variations in both output and inflation simultaneously impact the tax base, resulting in changes in tax revenue.

VAR-based Government Spending Multipliers. For cumulative government spending multipliers, we calculate the multipliers following Mountford and Uhlig (2009) and Auerbach and Gorodnichenko (2012) as in Eq. (2).

2.2.2. *Local Projections (LPs)*

For comparison, we establish an LP model. Local projections, introduced by Jordà (2005), provide a flexible alternative to VARs for estimating impulse response functions by directly estimating the response of variables to shocks at each horizon without imposing the dynamic restrictions inherent in VAR models.

Basic LP Framework. The linear LP model for each horizon h can be written as follows:

$$z_{t+h} = \alpha_h + \gamma_h t + \beta_h S_t + \sum_{\ell=1}^{L} \Gamma'_{h,\ell} X_{t-\ell} + \varepsilon_{t+h}, \tag{4}$$

where α_h represents a horizon-specific constant term and t is a linear time trend. z_t denotes the variable of interest, X_t represents a vector of endogenous control variables, and S_t is the identified shock sequence. We utilize the same vector of endogenous variables X_t as in Eq. (3), along with a lag length of four ($L = 4$), to ensure consistency with the VAR model.

The impulse responses of the target variable z_t to the shock S_t, denoted as $\{\beta_h\}_{h\geq 0}$ in Eq. (4), can then be obtained as follows:

$$\beta_h = E(z_{t+h} \mid S_t = 1,\ \{X_{t-\ell}\}_{\ell=1}^{L}) - E(z_{t+h} \mid S_t = 0,\ \{X_{t-\ell}\}_{\ell=1}^{L}). \tag{5}$$

LP-based Government Spending Multipliers. Throughout the empirical analysis using LPs, the two variables of interest are government spending (G_t) and output (Y_t). For each horizon h, β_h's for G_t and Y_t are obtained, and then LP-based government spending multipliers are defined as the ratio between these two impulse responses, analogous to Eq. (2).[8]

LP-IV Method. Ramey and Zubairy (2018) propose an instrumental-variable LP method for estimating government spending multipliers when NIPA variables are transformed using the Gordon–Krenn approach.[9] The LP-IV methodology estimates cumulative multipliers directly in a single-step procedure. For horizon h, the LP-IV regression can be written as:

$$\sum_{k=0}^{h} y_{t+k} = \alpha_h + \gamma_h t + m_h \sum_{k=0}^{h} g_{t+k} + \sum_{\ell=1}^{L} \Gamma'_{h,\ell} x_{t-\ell} + \omega_{t+h}, \tag{6}$$

where y_t and g_t denote the GK-transformed output and government spending, respectively. The endogenous vector x_t consists of the same five variables as in Eqs. (3) and (4), with output, government spending, and taxes obtained through GK transformation. $\sum_{k=0}^{h} y_{t+k}$ and $\sum_{k=0}^{h} g_{t+k}$ represent the cumulative sums of GK-transformed output and government spending from t to $t+h$.

LP-IV based Government Spending Multipliers. By comparing with Eq. (F5), the estimated coefficients m_h in Eq. (6) can directly be interpreted as the cumulative government spending multipliers at horizon h. The LP-IV approach utilizes various sequences of government spending shocks as instrumental variables for $\sum_{k=0}^{h} g_{t+k}$ to estimate the coefficients m_h.[10]

2.3. Shock identification strategies

2.3.1. True shock is known: model-implied shock sequence (benchmark)

In our benchmark analysis, we assume that econometricians have access to the model-implied government spending shock sequence. This assumption not only serves as a benchmark but also allows us to examine the effects of other dimensions, econometrical model and variable transformation methods, while controlling for the shock identification dimension by holding it constant.

2.3.2. True shock is unknown: Recursive/BP and FE identifications

For the case where the true government spending shocks remain unknown, which is the more realistic scenario, we compare the model-implied government spending multiplier with several commonly employed econometric specifications: Recursive identification(recursive VAR and LP-BP) and FE identification. For LP-IV, we employ both of BP and FE shock identification.

Recursive VAR/LP-BP. Initially, we present the findings from employing the Vector Autoregressive (VAR) model with a recursive identification strategy and the LP with Blanchard and Perotti (2002, "BP" hereafter) shocks. In LP-BP specification, we employ BP shocks identified from Structural Vector Autoregressions (SVARs), which resembles the recursive VAR in our setup.

FE identification. The second approach to constructing a government spending shock sequence involves extracting the unanticipated components of government spending based on the forecast errors (FE) of the variable, as demonstrated by Ramey (2011) and Auerbach and Gorodnichenko (2013). A forecast error shock in government spending can be constructed as follows. As the first step, we define the forecast error as $fe_t = G_t - E_{t-1}\left(G_t\right)$, where E_t represents private agents' expectations formed using information up to time t. Subsequently, following the methodology proposed by Auerbach and Gorodnichenko (2013), we regress fe_t on the contemporaneous and lagged values of $[\Delta G, \Delta Y, \pi, \Delta T, R]$, where Δ denotes changes in the respective variables. The forecast error shock is defined as the residual of this regression. In doing so, a practical question arises regarding how to treat the expectation term, E_t. For instance, many previous empirical studies utilize professional forecasts' expectations, often employing survey-measured

8 In order to detail this point, we consider an LP model as follows:

$$y_{t+h} = \alpha_h^y + \beta_h^y g_t + \gamma_h^y(L)x_{t-1} + \epsilon_{t+h}^y, \tag{F1}$$

$$g_{t+h} = \alpha_h^g + \beta_h^g g_t + \gamma_h^g(L)x_{t-1} + \epsilon_{t+h}^g, \tag{F2}$$

where y_t and g_t denote output and government spending divided by the trend of output (Y_p), respectively. Then we have

$$\beta_h^y = \frac{\Delta y_{t+h}}{\Delta g_t} = \frac{\Delta Y_{t+h}/Y_p}{\Delta G_t/Y_p}, \tag{F3}$$

$$\beta_h^g = \frac{\Delta g_{t+h}}{\Delta g_t} = \frac{\Delta G_{t+h}/Y_p}{\Delta G_t/Y_p}. \tag{F4}$$

Finally, cumulative government spending multipliers at horizon h are expressed as

$$m_h \equiv \frac{\Sigma_{k=0}^h \Delta Y_{t+k}}{\Sigma_{k=0}^h \Delta G_{t+k}} = \frac{\Sigma_{k=0}^h \beta_k^y}{\Sigma_{k=0}^h \beta_k^g}. \tag{F5}$$

9 We do not include SVAR with external IV(SVAR-IV) in our analysis for several reasons. First, as Stock and Watson (2018) and Li et al. (2024) demonstrate, SVAR-IV requires invertibility for consistency, while LP-IV does not. Given that fiscal shocks are susceptible to non-invertibility due to the fiscal foresight phenomenon, SVAR-IV would likely produce systematically biased estimates in our setting. Second, to our knowledge, SVAR-IV is rarely used for government spending multiplier estimation, while there are applications to tax shocks such as Mertens and Ravn (2013).

10 A key advantage of the LP-IV methodology is that it provides a single-step estimation procedure for calculating government spending multipliers. This allows for straightforward statistical inference regarding m_h through the standard errors associated with the coefficients. In contrast, the standard LP procedure described in Eqs. (F1) through (F5) requires calculating ratios of coefficients, making statistical inference more complex as standard errors for β_h^y and β_h^g are only available separately.

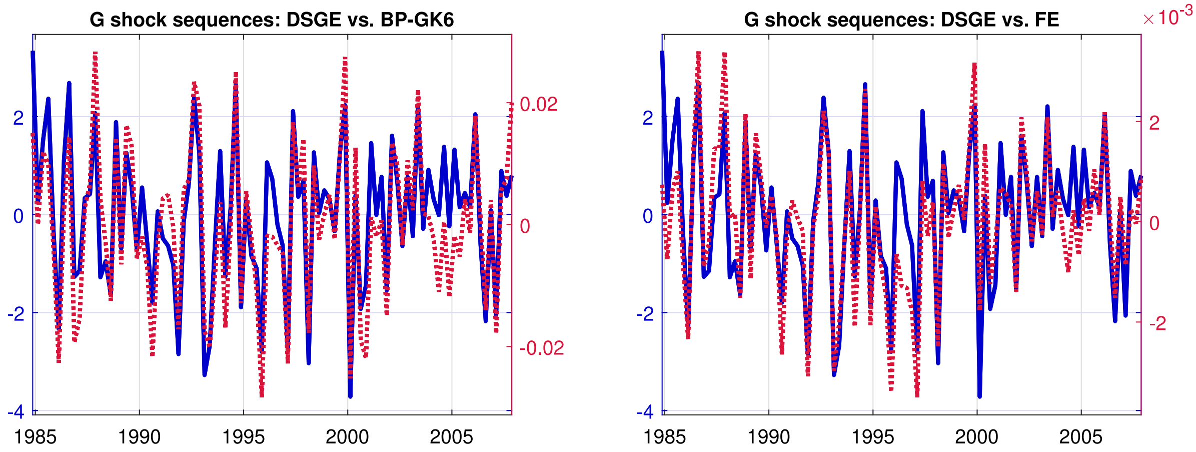

Fig. 4. Model-implied and recovered government spending shock sequences. **[Left panel]** The solid line (with the left y-axis) and the dashed line (with the right y-axis) indicate the model-implied government spending shock sequence and the sequence derived from the Blanchard–Perotti VAR model with GK-transformed variables, respectively. **[Right panel]** The solid line (with the left y-axis) and the dashed line (with the right y-axis) indicate the model-implied government spending shock sequence and the sequence derived using the forecast error on government spending, respectively.

forecasts such as the Survey of Professional Forecasters (SPF). This paper uses expectations associated with the DSGE model in Section 2.1 to maintain consistency with other model-based empirical analyses.[11]

Quality of Recovered Shock Sequences. To assess the validity of our identification strategies, we evaluate how well the BP and FE methods can recover the true government spending shocks from our simulated data. Fig. 4 plots the recovered government spending shock sequences obtained from both BP and FE methods, along with the true sequence implied by the DSGE model. Aside from differences in magnitude, the econometrically recovered shock sequences track the model-implied one quite well, with correlation coefficients of 0.88 (BP-VAR) and 0.86 (FE), respectively. The variations in the two recovered shock sequences turn out to be similar, with a mutual correlation coefficient of 0.89. The high correlation between the two shock sequences may result from the similarity in the information sets used to construct them. Due to the lagged structure in VARs, past information about the endogenous variables is automatically controlled for when identifying the BP shock. This methodology, as mentioned earlier, is similar to the concept of establishing the FE shock sequence, where the forecast error is regressed against the current and lagged values of the endogenous variables.

LP-IV with BP and FE identifications. We also present results from LP-IV models using BP shock, FE shock, and both shocks as instrumental variables. When adopting the LP-IV method, it is necessary to ensure that the identified shocks are relevant instruments to avoid the weak instrument problem. Olea and Pflueger (2013) show that in the presence of serially correlated errors, the F-statistic should exceed the threshold value to ensure the adopted shock sequence is not affected by the weak instrument problem. Accordingly, we plot the effective F-statistics using (Olea and Pflueger, 2013) with a threshold value of 37 for the 5% significance level in Fig. 5. While the FE shock may have a potential relevance problem in the medium to long run, the F-statistics for the BP shock exceed the threshold value over the entire horizon. This implies that the empirical procedure is unlikely to be affected by the weak IV problem when using the BP shock, as pointed out by Ramey and Zubairy (2018). Note that the weak IV problem for the FE shock in the medium to long run can be addressed by adding the BP shock, which has high relevance across all horizons. Accordingly, as Fig. 5 shows, the effective F-statistics are above the thresholds when using both shocks as instruments.[12]

2.4. Variable transformations: Log vs. Gordon-Krenn

Throughout our empirical analyses, we use two methods for transforming NIPA variables: (1) a log transformation and (2) the GK transformation as in Ramey and Zubairy (2018).

Log Transformation. Government spending multipliers derived from a log-transformed model can be expressed as follows:

$$\text{Government spending multiplier } (k) = \frac{\Delta \log(Y_k)}{\Delta \log(G_k)}\frac{\bar{Y}}{\bar{G}} = \frac{\Delta Y_k / Y_k}{\Delta G_k / G_k}\frac{\bar{Y}}{\bar{G}} \approx \frac{\Delta Y_k}{\Delta G_k}, \tag{7}$$

where G and Y represent government spending and output, respectively, in level terms. $\bar{G}$ and $\bar{Y}$ denote the means of government spending and output over the sample period. As Eq. (7) makes explicit, log-transformations of G and Y necessitate $\bar{Y}/\bar{G}$, commonly

[11] We also conduct subsequent analyses using the SPF data and find that they are less successful in recovering the DSGE multipliers compared to those using model-based expectations. The results associated with the SPF dataset are available upon request.

[12] Ramey and Zubairy (2018) also address the weak IV problem of the military news shock by adding the BP shock as an additional instrument. Since the identified news shock is anticipated several quarters before government spending is implemented, it encounters a relevance problem in the short-run horizons.

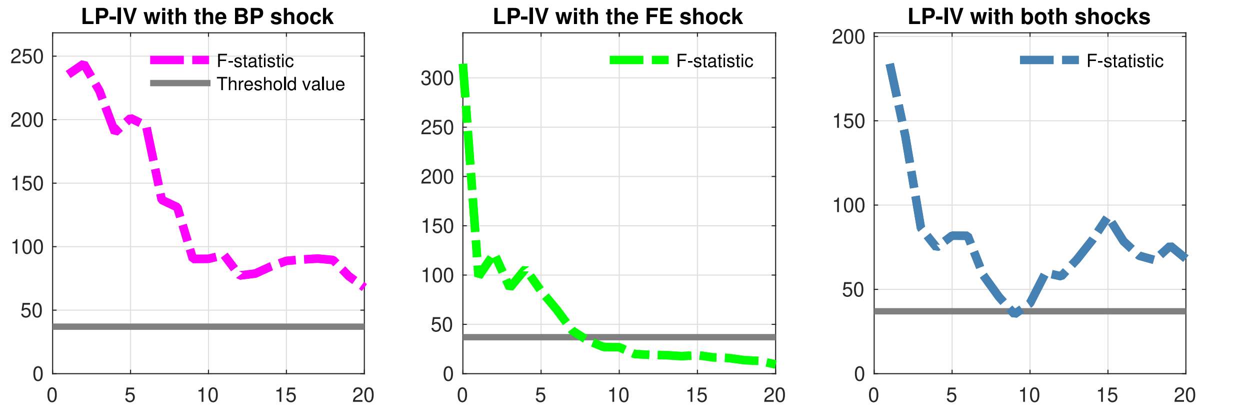

Fig. 5. The F-statistics in Olea and Pflueger (2013) for checking the relevance condition of the instrumental variables in the LP-IV models. In each panel, the horizontal line indicates the Olea and Pflueger's threshold vale. The x-axis measures quarters.

referred to as an ex-post conversion factor. This factor is used to transform the ratio between the impulse responses of output and government spending, $\Delta\log(Y_k)/\Delta\log(G_k)$, into dollar-valued multipliers. Eq. (7) represents the prototypical method for measuring government spending multipliers in previous studies employing log-transformations, as seen in works such as Blanchard and Perotti (2002) and Mountford and Uhlig (2009).

Gordon–Krenn Transformation. Ramey and Zubairy (2018) suggest using the variable transformation method proposed in Gordon and Krenn (2010). The key idea behind the Gordon–Krenn transformation is to divide each real National Income and Product Accounts (NIPA) series by an estimate of potential real GDP. For instance, Ramey and Zubairy's (2018) GK-transformation procedure consists of two steps. In the first step, they regress logged real GDP on a 6th-degree polynomial. Then, by taking the exponential of the fitted values from the regression, the GK-transformed variables in levels are obtained. With the variables now in levels, rather than a log–log form, there is no need to use the ex-post conversion factor in calculating government spending multipliers.

3. Results

This section assesses the accuracy of the econometrically estimated multipliers by comparing them with the model-implied government spending multiplier obtained in Section 2.1. The precision of each econometric methodology in estimating government spending multipliers is evaluated on the basis of how closely the econometrically obtained multipliers align with their DSGE counterparts.

In doing so, we start from the scenario in which economists have full access to the true government spending shock sequence. This assumption, of course, may not hold in reality where the true government spending shock sequence is unavailable and must instead be inferred using sound econometric approaches. However, this scenario serves as our reference point. We then consider more realistic scenarios where the true government spending shock sequence is unknown and can only be recovered using an econometric model, either VARs or LPs.

3.1. True shock is known: LPs with the true government spending shock sequence

We begin by comparing the government spending multipliers that emerged from the estimated DSGE model with those from the LP model in Eq. (4), assuming that the true government spending shock sequence is accessible to econometricians. A salient feature of employing a DSGE model is to obtain model-implied sequences for exogenous shocks, with our primary interest being the government spending shock sequence as depicted in Fig. 2.

Fig. 6 presents the DSGE-based government spending multipliers alongside those recovered from the LP model using log and GK6 transformations, respectively. First, focusing on the government spending multipliers obtained from the LP model with log transformation(LP-log), they tend to be underestimated initially and then overestimated in both the short and medium runs. However, this tendency reverses about four years after a shock, as the LP-log multipliers begin to show some degree of underestimation.

As plotted with the solid line with squares in the figure, the results from the GK-transformed LP model (LP-GK) provide much more accurate estimates for the multipliers compared to the log-transformed model. Although they also display underestimation at impact and overestimation in the medium run, it is notable that the GK-transformed LP multipliers closely recover the model-implied ones, particularly for horizons from two to seven quarters after a shock.

These empirical results highlight two important findings. First, the medium- to longer-run underestimation of government spending multipliers associated with the LP method tends not to vanish even when the true shock sequence for government spending is utilized. This tendency is observed regardless of the variable transformation method employed, occurring with both log and GK transformations. Second, the finding echoes (Owyang et al., 2013; Ramey and Zubairy, 2018) that the GK transformation is likely to

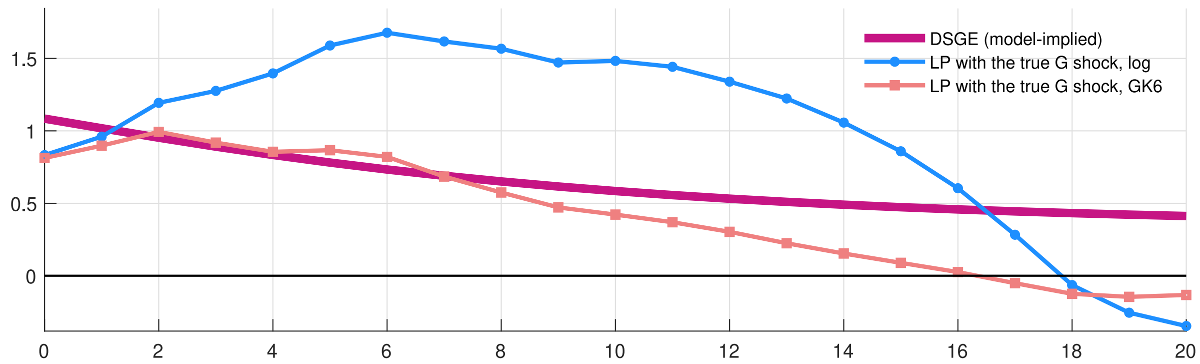

Fig. 6. DSGE- and LP-based government spending multipliers for output. The solid line indicates the cumulative multipliers evaluated at the median of the posterior parameter estimates from the DSGE model, while the solid lines with circles and squares represent those derived from the LP models in log and GK6 transformation, respectively, when the true shock sequence is available. The x-axis measures quarters.

be a better approach when the size of the ex-post conversion factor is considerable. The ex-post conversion factor is approximately 8 in Owyang et al. (2013), and it turns out that the value for our sample is quite comparable to theirs. The short- to medium-run overestimation under the log transformation likely stems from this factor.[13]

3.2. Performance under unknown shocks

In practice, econometricians rarely have access to the true shock sequence that we used as our benchmark in the previous section. When the true government spending shock sequence is unknown, researchers must rely on econometric identification strategies to recover shocks from the data. In this section, we compare the performance of these alternative identification strategies using our simulated data, evaluating how closely they can approximate the model-implied government spending multipliers when the true shock sequence is unavailable.

3.2.1. Recursive/BP shocks

Variable Transformation Effects. We first examine how the choice of variable transformation affects multiplier estimates in both recursive VARs and LPs with BP shocks. Fig. 7 presents the performance of both methodologies under log and GK6 transformations, demonstrating that the issue of overestimation associated with the ex-post conversion factor persists in log-transformed models and is evident in both VAR and LP frameworks. The use of log transformation consistently overestimates government spending multipliers across most horizons considered, regardless of the econometric method employed. In sharp contrast, both GK-transformed VAR and LP-BP models produce multipliers that are considerably more aligned with the DSGE benchmark.

Model Comparison with GK6 Transformation. Having established the superiority of GK transformation, we now focus on comparing the performance of different econometric approaches when using GK6-transformed variables. Fig. 8 presents the government spending multipliers from recursive VARs, LP with true shocks, and LP with BP shocks, all using GK6 transformation.

Fig. 8 reveals that while LP with true shocks (solid purple line) performs best in the very short run (horizons 1–4), the recursive VAR with GK6 transformation (cyan line with squares) demonstrates superior performance in the medium to long run. Notably, when the shock identification changes from the true sequence to BP identification, the LP method loses its short-run advantage and not only fails to capture the accuracy of LP with true shocks but also underestimates multipliers more severely than the recursive VAR in the long run.

Table 1 presents the root mean squared errors (RMSE) for different econometric specifications with GK transformation across various time horizons. While LP with true shocks achieves the lowest RMSE in the short run (1–8 quarters) with an RMSE of 0.20, and demonstrates the best overall performance (1–20 quarters) with an RMSE of 1.37, recursive VAR with GK6 transformation dominates in the medium to long run (9–20 quarters) with an RMSE of 0.75, significantly outperforming both LP with true shocks (RMSE = 1.36) and LP-BP-GK6 (RMSE = 1.51) in this horizon.

3.2.2. FE and LP-IV: Performance with less accurate shock identification

Based on our DSGE model, Section 2.3.2 showed that both BP and FE identification strategies recover shock sequences with correlations to the true DSGE-implied shocks of 0.88 and 0.86, respectively. Although FE identification is a widely used approach in the literature, it produces somewhat less accurate shock recovery in our DGP compared to the BP method. We exploit this property to examine how econometric methods perform under less precise shock identification by investigating the performance of FE shocks in both VAR and LP specifications, as well as LP-IV models.

[13] The differences between log and GK6 transformations are evident across most VAR and LP specifications in our paper, as will be illustrated in the additional figure presented in Section 3.2.3.

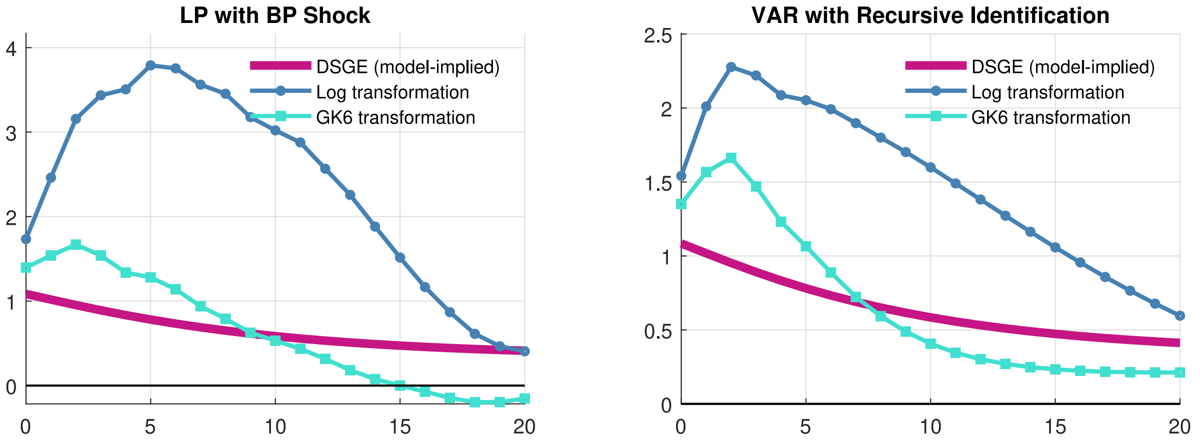

Fig. 7. Variable transformation effects in recursive VAR and LP-BP models. [Left panel] DSGE- and LP-BP-based government spending multipliers for output. The solid line indicates the cumulative multipliers from the DSGE model, while the lines with circles and squares represent those from LP models with BP shocks in log and GK6 transformation, respectively. [Right panel] DSGE- and VAR-based government spending multipliers for output. The solid line indicates the cumulative multipliers from the DSGE model, while the dashed lines with circles and squares represent those from recursive VAR models in log and GK6 transformation, respectively. The x-axis measures quarters.

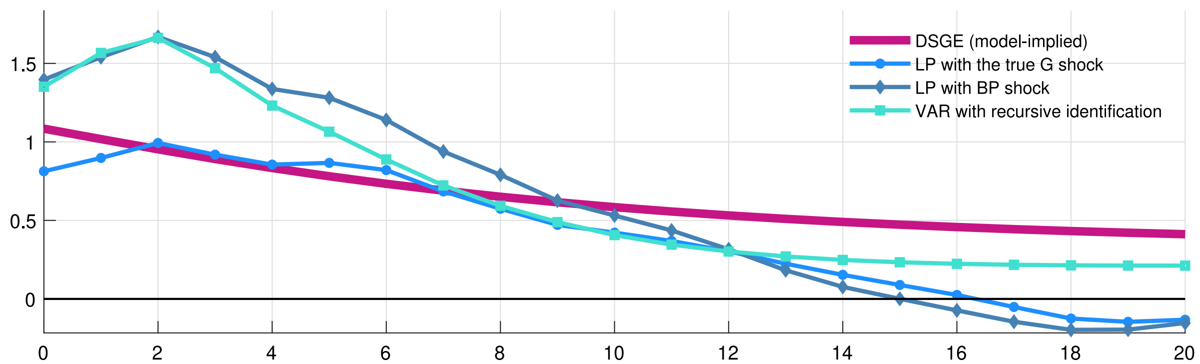

Fig. 8. Comparison of econometric methods with GK6 transformation. The solid purple line indicates the cumulative multipliers from the DSGE model. The lines with circles, diamonds, and squares represent the multipliers associated with LP using true government spending shocks with GK6 transformation, LP with BP shocks with GK6 transformation, and recursive VAR with GK6 transformation, respectively. The x-axis measures quarters. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 1
Root mean squared errors (RMSE) comparison for key econometric methodologies with GK transformation. Rankings in terms of the RMSE statistics are reported in parentheses for the top two performers in each horizon category.

Specification	1–20 quarters	1–8 quarters	9–20 quarters
LP with the true G shock	1.37 (1)	0.20 (1)	1.36 (2)
Recursive VAR	1.40 (2)	1.19 (2)	0.75 (1)
LP with the BP shock	2.06	1.40	1.51

Fig. 9 presents a comprehensive comparison of government spending multipliers obtained using FE identification and LP-IV methods, both representing scenarios where shock identification is less closely aligned with the true DSGE-implied sequence. The results in the left panel of Fig. 9 reveal important insights about how shock identification affects different econometric methods. When using FE identification instead of BP identification, LP methods show deteriorated performance primarily in the short run, with LP-FE exhibiting significant overestimation in the short to medium run compared to LP-BP. However, the long-run performance of LP methods remains relatively similar regardless of the shock identification strategy employed, as both LP-BP and LP-FE exhibit underestimation. In contrast, the VAR demonstrates sensitivity to shock identification across all horizons, showing that shock identification affects VAR performance throughout the entire time horizon.

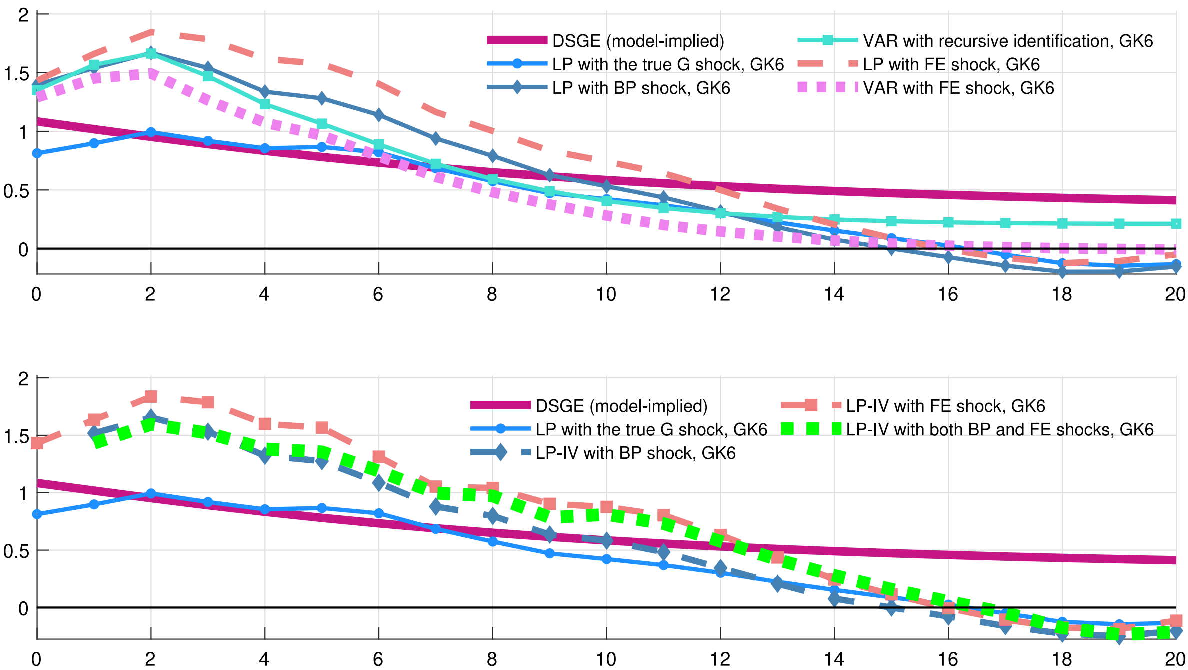

Fig. 9. Government spending multipliers with less accurate shock identification, including forecast error shocks. [Upper panel] The solid line indicates the cumulative multipliers from the DSGE model. The lines with circles, diamonds, and squares represent the multipliers associated with LP using true government spending shocks with GK6 transformation, LP with BP shocks with GK6 transformation, and recursive VAR with GK6 transformation, respectively. Additionally, the dashed and dotted lines indicate the multipliers associated with LP using FE shocks with GK6 transformation and VAR using FE shocks with GK6 transformation, respectively. The x-axis measures quarters. [Lower panel] The solid line indicates the cumulative multipliers from the DSGE model, while the line with circles represents the multipliers associated with LP using true government spending shocks with GK6 transformation. The dashed line with diamonds, the dashed line with squares, and the dotted line indicate the multipliers associated with LP-IV using BP shocks with GK6 transformation, LP-IV using FE shocks with GK6 transformation, and LP-IV using both BP and FE shocks with GK6 transformation, respectively. The x-axis measures quarters.

The right panel of Fig. 9 shows LP-IV results using different instrumental variable combinations. Consistent with the IV F-statistics results in Section 2.3.2, FE shocks exhibit weaker instrumental relevance compared to BP shocks. The LP-IV results exhibit a similar pattern to the standard LP methods: the choice of instrumental variables primarily affects short-run performance, with LP-IV using FE shocks showing larger estimation errors in the short run compared to LP-IV using BP shocks. However, in the long run, the differences between different LP-IV specifications become minimal, suggesting that the tendency of the LP methodology to underestimate multipliers in the long run dominates over the quality of instrumental variables at longer horizons.

3.2.3. Variable transformation vs. econometric method choice

An important finding from our analysis is that the choice of variable transformation can have a more substantial impact on multiplier estimates than the choice of econometric methods. Fig. 10 illustrates this point by comparing the magnitude of differences arising from variable transformation choices versus econometric method choices.

As shown in the left panel of Fig. 10, the differences between log and GK6 transformations are substantial and systematic across all horizons, with log transformations consistently overestimating multipliers by large margins. In contrast, the right panel shows that when using the preferred GK6 transformation, the differences between VAR and LP methods (excluding LP with true shocks) are much smaller and more nuanced, varying primarily by time horizon rather than showing systematic bias.

Regarding the transformation methods for the model's endogenous variables, we support the finding by Owyang et al. (2013) that the GK transformation approach yields more precise government spending multipliers than the logarithmic transformation method in both VAR and LP frameworks. The ex-post conversion factor turns out to be sizable in our model, and the logarithmic transformation tends to introduce an upward bias in the multiplier estimates, regardless of the econometric specifications. It is notable that this upward bias is not limited to a VAR framework; it hardly disappears even when the true shock sequence is used in an LP framework.

3.3. Overall summary

The key takeaways are as follows: the accuracy of the LP method and recursive VARs in estimating government spending multipliers tends to vary across different time horizons. The importance of variable transformation is evident throughout our

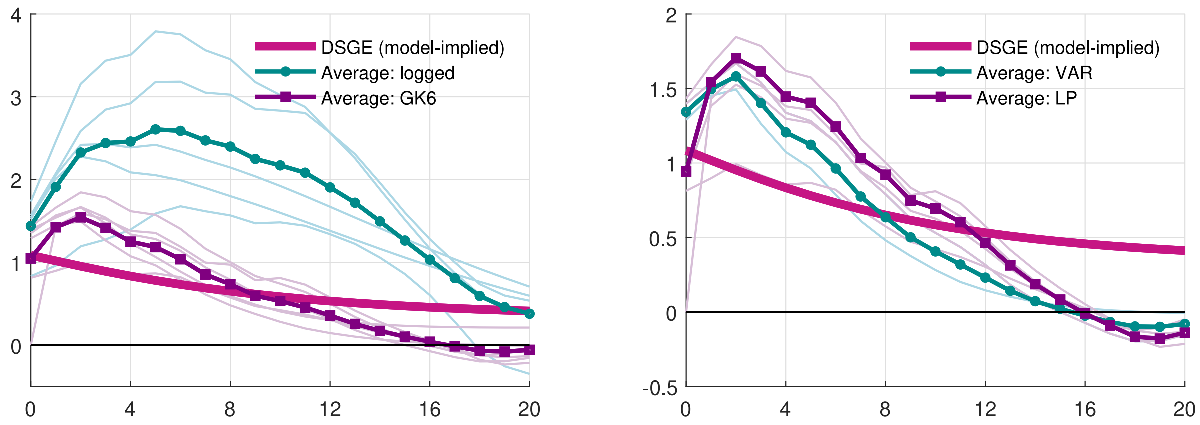

Fig. 10. [Left panel] Combined overview showing the differences between log (solid line with circles) and GK6 (solid line with squares) transformations across all econometric specifications. [Right panel] Comparison of VAR (solid line with circles) and LP (solid line with squares) average performance using GK6 transformation, excluding LP with true shocks. In each panel, the solid line indicates the cumulative multipliers from the DSGE model, and the x-axis measures quarters.

Table 2
Root mean squared errors (RMSE) associated with each econometric methodology. Rankings in terms of the RMSE statistics are reported in parentheses.

Specification	1–20 quarters	1–8 quarters	9–20 quarters
LP with the true G shock, log	3.01	1.94	2.30
LP with the true G shock, GK6	1.37 (1)	0.20 (1)	1.36 (3)
Recursive VAR, log	4.23	3.47	2.41
Recursive VAR, GK6	1.40 (2)	1.19 (2)	0.75 (1)
LP with the BP shock, GK6	2.06	1.40	1.51
LP with the FE shock, GK6	2.38	2.01	1.28 (2)
LP-IV with the BP shock, GK6	2.06	1.35	1.56
LP-IV with the FE shock, GK6	2.40	1.95	1.40
LP-IV with both the BP and FE shocks, GK6	1.97	1.42	1.36 (3)

analysis. Log-transformed models consistently overestimate government spending multipliers across most horizons, regardless of the econometric method employed, due to issues associated with the ex-post conversion factor. In contrast, GK6 transformation produces multipliers that are considerably more aligned with the DSGE benchmark across all methodologies examined.

Table 2 provides a comprehensive comparison of all econometric methodologies examined in this study, including both FE identification and LP-IV specifications with different instrumental variable combinations. As shown in the second column of Table 2, LP with true shocks using GK6 transformation achieves the best overall performance (1–20 quarters) with an RMSE of 1.37, followed closely by recursive VAR with GK6 transformation (RMSE = 1.40). In the short run (1–8 quarters), LP with true shocks and GK6 transformation dominates with an RMSE of 0.20, significantly outperforming all other methods. For the medium to long run (9–20 quarters), recursive VAR with GK6 transformation demonstrates superior performance with an RMSE of 0.75, followed by LP with FE shock (RMSE = 1.28). As Figs. 8 and 9 illustrate, the econometrically produced multipliers show varying degrees of accuracy across different methodologies and time horizons.

The RMSE analysis confirms several findings. First, based on our DSGE model, FE-based identification generally produces larger estimation errors compared to BP identification, with LP-FE showing an RMSE of 2.38 compared to 2.06 for LP-BP in the overall performance (1–20 quarters). This deterioration in performance reflects the lower correlation between FE shocks and the true shock sequence in our particular model specification. Second, the RMSE associated with the LP-IV specification using both shocks (1.97) is lower than that of the one-shock-only specifications, which are 2.06 and 2.40, respectively.

Nevertheless, the LP with true shocks using GK6 transformation demonstrates the best overall performance across all horizons (RMSE = 1.37), with recursive VAR with GK6 transformation showing competitive performance (RMSE = 1.40) and superior results in the medium to long run, reinforcing our main finding about the importance of both shock quality and variable transformation when the true shock sequence is unavailable.[14]

[14] While BP identification may perform less effectively under certain conditions — such as when government spending responds to current output or when shocks are anticipated — which we examine in Section 4, recursive VAR consistently demonstrates strong performance across various identification strategies and model specifications.

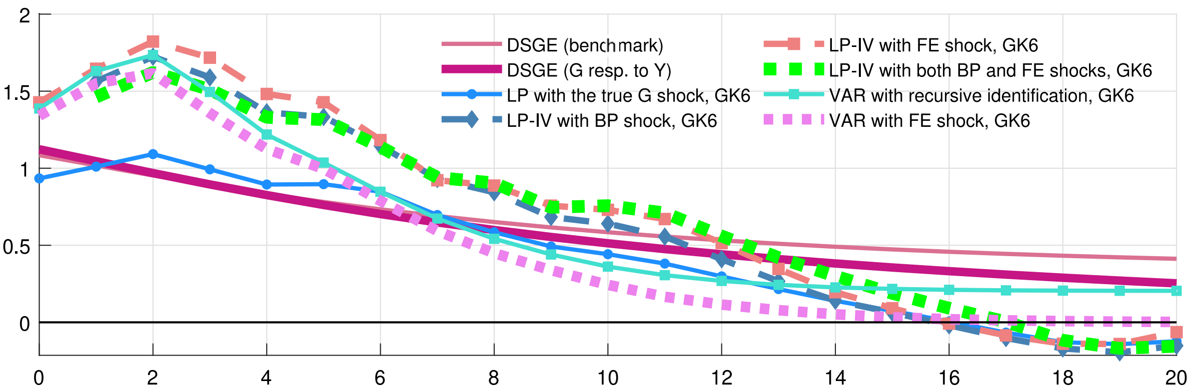

Fig. 11. [Robustness] Cumulative government spending multipliers: DSGE model with endogenous government response to contemporaneous output. The thin and thick solid lines indicate the cumulative multipliers from the baseline DSGE model and the one associated with the endogenous government response to contemporaneous output, respectively. The remaining lines emerge from the alternative DSGE model specification. The lines with circles and cyan line with squares represent the multipliers associated with LP using true government spending shocks with GK6 transformation and recursive VAR with GK6 transformation, respectively. The dashed line with diamonds, the dashed line with squares, and the thick dotted line indicate the multipliers associated with LP-IV using BP shocks with GK6 transformation, LP-IV using FE shocks with GK6 transformation, and LP-IV using both BP and FE shocks with GK6 transformation, respectively. The thin dotted line indicates the multipliers associated with VAR using FE shocks with GK6 transformation. The x-axis measures quarters. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

4. Robustness

Our benchmark results favor using the BP-identified VAR model or an LP that utilizes the shock sequence from a BP-identified VAR model for recovering government spending multipliers using econometric models. However, given the nature of the BP identification methodology, it is worth investigating at least two dimensions to determine whether the conclusions stem directly from the specific modeling choice in the DSGE model, as described in Section 2.1. The first dimension concerns the government spending process in the DSGE model, as detailed in Eq. (1). In this specification, government spending is assumed to be unresponsive to contemporaneous changes in output ($\hat{Y}_t$), which aligns precisely with the assumptions of the BP identification methodology. The second dimension involves the absence of anticipated effects of government spending in the model, commonly referred to as unanticipated government spending shocks. The BP identification methodology is designed to capture the effects of unanticipated changes in government spending. However, in reality, these changes are often anticipated by private agents due to legislation and implementation lags.[15]

4.1. Allowing responses of government spending to changes in current output

Previous studies, such as Feve et al. (2013), suggest that government spending is responsive to changes in output during the current period. These studies, in particular, assume that government spending tends to be countercyclical, as one of the key objectives of fiscal policy is to maintain macroeconomic stability. Consequently, the government spending process in Eq. (1) can be modified as follows:

$$\hat{G}_t = \rho_G \hat{G}_{t-1} - (1-\rho_G)\left(\psi_Y \hat{Y}_t + \gamma_G \hat{s}^b_{t-1}\right) + \epsilon^G_t, \tag{8}$$

where the parameter ψ_Y captures the degree of government spending responsiveness to fluctuations in contemporaneous output.

We re-estimate the DSGE model using the government spending process specified in Eq. (8) instead of the original specification in Eq. (1). In doing so, we choose a gamma prior with a mean of 0.5 and a standard deviation of 0.25 for the parameter ψ_Y. The Bayesian estimation indicates that the posterior distribution for the parameter has a median of 0.49, with a 90% interval ranging from 0.19 to 1.02.[16] Nevertheless, the similarity in the prior and posterior distributions for the government spending response to current output suggests that this parameter is unlikely to be well-identified.

Fig. 11 compares the output multipliers from this alternative DSGE specification with those estimated using VAR and LP methods. As shown in Figure D3, allowing government spending to respond to contemporaneous output changes hardly affects the model-implied multipliers, particularly in the short run. In this case, the VAR with recursive identification (cyan with squares) continues to track the model-implied responses reasonably well, especially in the long run.

[15] We are grateful to the referees for suggesting this extension. The following discussion has been added in response to their helpful comment.

[16] The complete sets of selected parameter estimates associated with the alternative specifications are provided in the companion appendix in Section D.2.

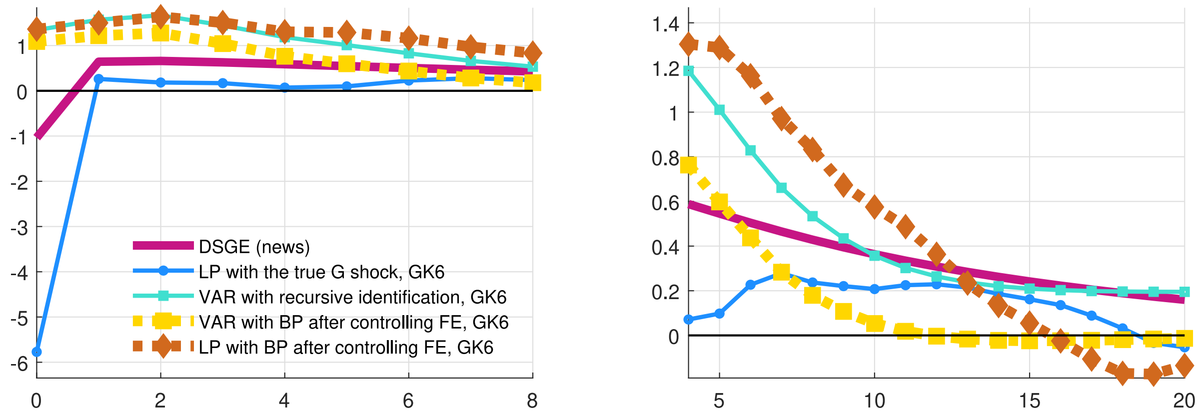

Fig. 12. [Robustness] Cumulative government spending multipliers: DSGE model with an anticipated component of government spending. [Left panel: in the short run] The solid line indicates the short-run (from impact to the horizon of 8 quarters) cumulative multipliers from the DSGE model with the anticipated component of government spending. The lines with circles and squares represent the multipliers associated with LP using true government spending shocks with GK6 transformation and recursive VAR with GK6 transformation, respectively. The dotted line with squares and the dotted line with diamonds indicate the multipliers associated with VARs using BP and FE shocks, respectively, with GK6 transformation, after controlling for forecast error for government spending. The x-axis measures quarters. [Right panel: in the longer runs] The identical multipliers are plotted for the longer runs, from the horizon of 4 to 20 quarters. The x-axis measures quarters.

4.2. *Allowing anticipated component of government spending*

The second alternative specification involves examining whether our main conclusion remains robust when the model is augmented with an anticipated component of government spending. To accomplish this, we extend the government spending process in Eq. (1) to account for both unanticipated and anticipated changes in government spending. Following Walker and Leeper (2011), both components are included as follows:

$$\hat{G}_t = \rho_G \hat{G}_{t-1} - (1-\rho_G)\gamma_G \hat{s}^b_{t-1} + \theta_0 \epsilon^G_t + \theta_1 \epsilon^G_{t-1}, \tag{9}$$

where the parameters θ_0 and θ_1 determine the relative importance of unanticipated changes and 1-period anticipated changes in government spending, respectively. These parameters are subject to the constraints $\theta_0 + \theta_1 = 1$, with $\theta_0 \geq 0$ and $\theta_1 \geq 0$. As in the previous robustness check, the DSGE model is reestimated by substituting the government spending process in Eq. (1) with that in Eq. (9). We adopt a beta prior with a mean of 0.75 and a standard deviation of 0.1 for the parameter θ_1, thereby giving relatively more weight to the anticipated component *a priori*. The posterior distribution for this parameter turns out to have a median of 0.92 with a 90% interval of [0.86, 0.97]. These estimates suggest that the data assign relatively more weight to the anticipated component of changes in government spending than to the unanticipated component.

Given the high estimated weight on the anticipated component ($\theta_1 = 0.92$), it becomes crucial to isolate the surprise element in government spending when making comparisons with econometric models. To address this concern, we employ the fixed effects (FE) controlled specification in our LP and VAR analyses, which helps to filter out predictable components and better capture the unanticipated portion of government spending shocks that is the focus of our identification strategy.[17]

Fig. 12 presents the comparison between the DSGE model with news shocks and the VAR/LP estimates for both short-run and long-run horizons. The inclusion of anticipated government spending shocks in the DSGE model creates markedly different dynamics from the baseline model, especially in the short-run. In the short-run (left panel), the model exhibits the characteristic delayed response pattern, with the true multiplier starting negative and gradually rising. Most econometric methods, except for LP with true G shocks, fail to recover this initial pattern of the model-implied multiplier. However, in the long-run (right panel), the recursive VAR demonstrates improved performance and successfully recovers the model-implied multiplier pattern, contrasting with its initial overestimation in the short-run.

4.3. *Additional robustness checks*

The empirical methodologies may vary in how effectively they recover the government spending multiplier implied by the model and the corresponding model-implied impulse response functions (IRFs). In particular, it is possible for some methodologies to yield accurate government spending multiplier estimates while showing discrepancies in their recovery of IRFs, as noted by Ben Zeev

[17] When directly using FE identification instead of FE-controlled specifications, the estimation errors become even larger compared to the true DSGE model (available upon request).

Table 3
Ratio of the impulse response functions (IRF) between government spending and output (non-cumulative). Rankings in terms of proximity to DSGE model-implied values are reported in parentheses.

Specification	$h = 0$	$h = 4$	$h = 8$	$h = 16$
DSGE (model-implied)	1.08	0.61	0.31	0.14
LP with the true G shock, log	0.09	0.19	0.11	−0.69
LP with the true G shock, GK6	0.81(1)	0.66(1)	−0.05	−0.71
Recursive VAR, log	0.16	0.17	0.11	−0.07
Recursive VAR, GK6	1.35	0.23	−0.49	0.06(1)
LP with the BP shock, log	0.17	0.48	0.33(1)	−0.45
LP with the BP shock, GK6	1.25	0.72	−0.10	−1.11
LP with the FE shock, log	0.16	0.34	0.29	−0.60
LP with the FE shock, GK6	1.43	1.55	0.02	−1.41

et al. (2023). To examine this, we calculate the non-cumulative ratio of government spending to output IRFs at each horizon under different methodologies and assess how closely they align with the DSGE model-implied ratios.

Table 3 indicates that the accuracy of government spending multiplier (GSM) estimation closely mirrors the patterns observed in IRF ratios. Specifically, the GK-transformed LP method recovers both the GSM and IRF ratios most accurately in the short run when the true shock sequence is known ($h = 0$ and $h = 4$). For instance, at $h = 0$, the IRF ratio for the GK-transformed LP method with true shocks is 0.81, which closely aligns with the DSGE-implied value of 1.08. Similarly, at $h = 4$, the IRF ratio is 0.66, compared to the DSGE value of 0.61. These findings are consistent with the GSM estimation results, where the same method achieves the lowest RMSE values, indicating its robustness in both multiplier and IRF ratio recovery.

In contrast, recursive VARs struggle to recover the DSGE-implied IRF ratios in the short run, showing significant deviations such as 1.35 at $h = 0$ and 0.23 at $h = 4$. However, their performance improves notably in the long run ($h = 16$), where the recursive VAR with the GK6 transformation produces an IRF ratio of 0.06, closely approximating the DSGE-implied value of 0.14. This pattern also parallels the results of Table 2, where recursive VARs perform relatively better over longer horizons compared to other methods.

5. Conclusion

This paper uses a DSGE model as the benchmark to evaluate government spending multipliers estimated through various econometric specifications. Although these models are fitted to the same DSGE-model-implied simulated data, estimates of government spending multipliers over time critically depend on three key modeling choices: (1) variable transformations, (2) the use of VARs versus Local Projections (LPs), and (3) the identification strategy for government spending shocks.

Our findings demonstrate that the GK transformation substantially reduces bias relative to log transformations in both VAR and LP frameworks, bringing estimates closer to the DSGE benchmark. The choice between VARs and LPs should be informed by the quality and exogeneity of the identified government spending shocks. When robust instrumental variables or credible identification strategies are available, LPs provide more reliable estimates of government spending multipliers, particularly in the short run. However, LPs are highly sensitive to shock identification quality, as evidenced by the substantial variation in multiplier estimates across different identification schemes.

The framework developed in this paper can be applied to assess the reliability of LPs in estimating the state-dependent effects of fiscal policy. We leave the question of whether an LP framework is suitable for estimating the expansionary effects of government spending under time-varying monetary policy behavior for future research.

Disclaimer

The views expressed herein are those of the authors and do not necessarily represent those of the Bank of Korea.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We would like to thank Neville Francis (the editor) and two anonymous referees whose comments were particularly helpful. We also thank Jaeho Kim, Sungwon Lee, and seminar participants at Korea University, the Bank of Korea, and the Korea Development Institute for their helpful comments. Any remaining errors are the responsibility of the authors.

Appendix A. Related literature

A.1. General applications: VARs vs. LPs

There is a plethora of literature comparing VARs and LPs under various environments. Within this body of literature, the first strand focuses on the conditions under which VARs and LPs become identical. In a population with a sufficiently large lag length, Plagborg-Møller and Wolf (2021) show that LPs and VARs yield identical estimates of impulse responses, regardless of the identification scheme or underlying data generating process. For a finite sample setting, Jordà (2005) demonstrates that VAR impulse responses can be obtained consistently via LPs under the assumption of a finite order. However, Kilian and Lütkepohl (2017) argue against the equivalence between VARs and LPs, illustrating that finite-order LPs are generally more robust to model misspecification than VARs with a finite lag length.

The second type of research compares the relative accuracy of VARs and LPs by employing specific DGPs. Meier (2005) examines the performance of VARs and LPs using the estimated DSGE model by Smets and Wouters (2003) as the underlying DGP for the Euro area. The results indicate that, for the three exogenous shocks — monetary policy, technology, and labor supply — LPs do not outperform standard VAR methods and may even exhibit higher bias and variance. In contrast, Brugnolini (2018) conditions on the application of the VAR(12) data-generating process utilized in Christiano et al. (1999) to identify US monetary policy shocks. They report that the VAR model is generally a more preferred choice, while LPs offer a competitive alternative in situations of small sample size and misspecified lag length.

Another line of research compares VAR and LP methodologies, highlighting their strengths in terms of inference horizon. VAR and LP use different period of sample autocovariance for inference: the LP method forecasts future outcomes based on current covariates, while a VAR(p) model predicts longer-term impulse responses from the initial p sample autocovariances. Consequently, estimation results diverge between VAR and LP models, especially beyond the lag length horizon.

Regarding inference horizon, Stock and Watson (2018) emphasize that while VAR models capture immediate responses effectively due to their structured dynamic framework, they often struggle with accurately modeling long-term effects under conditions of model misspecification. Conversely, LP models offer greater flexibility and robustness for long-term inference, as demonstrated by Jordà (2005). According to Jordà (2005), this flexibility allows LPs to trace the evolution of economic impacts over time more effectively. Through the application of thousands of empirically relevant DGPs, Li et al. (2024) identified a trade-off between bias and variance in finite sample sizes for LP and VAR estimates. Specifically, they demonstrate that LP estimators tend to have lower bias compared to VARs. However, this advantage is counterbalanced by a significant increase in variance at intermediate and long horizons. The heightened variance in LP estimators stems from their increased sensitivity to data fluctuations at these horizons.

A.2. Size of government spending multipliers: VAR- and LP-based evidence

Similar to the categorization of dynamic causal inference steps by Nakamura and Steinsson (2018), the estimation of government spending multipliers involves three stages: identifying shocks, computing impulse responses to the identified shocks, and calculating multipliers from the impulse responses. Therefore, even after adopting a specific empirical methodology, such as VARs and LPs, factors such as the dataset, time period, or nature of the identified shocks can influence the size of the estimated government spending multipliers.

Given this stylized fact, several previous studies either survey the estimates of government spending multiplier in the existing literature or conduct meta-analyses using empirical strategies within it. Ramey (2016, 2019) and Gechert and Rannenberg (2018) are examples of studies within the VAR literature that have addressed this issue. In a summary, Ramey (2019) indicates that VAR-based estimates of government spending multipliers for the US typically fall within the range of 0.5 to 2. Ramey (2016) makes it explicit that government spending multiplier estimates for the US depend on the identification scheme and the method used to calculate multipliers from the estimated impulse responses, even after conditioning on a specific dataset. The methodologies employed in the paper include using defense spending or government spending residuals from defense spending as instrumental variables (IV), employing SVAR with Cholesky decomposition, or utilizing SVAR with news variables such as forecast error. The resulting estimates of government spending multipliers from the meta-analysis fall within the range of 0.6 to 1.5. Gechert and Rannenberg (2018) extend the analysis to countries beyond the US and find that government spending multiplier estimates using VARs exhibit a broader range, ranging from 0.85 to 5.05, across various countries.

Unlike VARs, to the best of our knowledge, there is a scarcity of surveys or meta-analyses on LPs in the context of government spending multipliers. A notable exception is found in Ramey (2016), who utilizes LPs to compare government spending multiplier estimates by employing three distinct government spending shock sequences recovered through the identification methodologies of Blanchard and Perotti (2002), Ramey (2011), and Ben Zeev and Pappa (2017) as IVs. The paper concludes that the three different shock sequences result in distinct impulse responses and government spending multipliers.

A.3. Size of government spending multipliers: How to treat variables matters

There is well-established evidence in the existing literature that the size of VAR- and LP-based government spending multipliers critically depends on how endogenous variables are treated in the model. In particular, Owyang et al. (2013) demonstrate that log-transformations of the model's endogenous variables when calculating government spending multipliers tend to produce an upward bias, irrespective of the econometric models used. To illustrate this issue, we begin by noting that government spending multipliers derived from a log-transformed model can be expressed as follows:

$$\text{Government spending multiplier } (k) = \frac{\Delta \log(Y_k)}{\Delta \log(G_k)} \frac{\bar{Y}}{\bar{G}} = \frac{\Delta Y_k / Y_k}{\Delta G_k / G_k} \frac{\bar{Y}}{\bar{G}} \approx \frac{\Delta Y_k}{\Delta G_k}, \tag{10}$$

where G and Y represent government spending and output, respectively, in level terms. $\bar{G}$ and $\bar{Y}$ denote the means of government spending and output over the sample period. As Eq. (10) makes explicit, log-transformations of G and Y necessitate $\bar{Y}/\bar{G}$, commonly referred to as an ex-post conversion factor. This factor is used to transform the ratio between the impulse responses of output and government spending, $\Delta \log(Y_k)/\Delta \log(G_k)$, into dollar-valued multipliers. Eq. (10) represents the prototypical method for measuring government spending multipliers in previous studies that employ log-transformations, as seen in works such as Blanchard and Perotti (2002) and Mountford and Uhlig (2009).

Owyang et al.'s (2013) main point is that a log-transformation can lead to an overestimation of government spending multipliers, especially when the ex-post conversion factor is high. In addition, Ramey and Zubairy (2018) highlight that such multiplier calculations can be inaccurate, particularly for a long sample period where the ex-post conversion factor varies widely.[18]

As an alternative approach, Ramey and Zubairy (2018) suggest using the variable transformation method proposed in Gordon and Krenn (2010). The key idea behind the Gordon–Krenn transformation is to divide each real National Income and Product Accounts (NIPA) series by an estimate of potential real GDP. For instance, Ramey and Zubairy's (2018) GK-transformation procedure consists of two steps. In the first step, they regress logged real GDP on a 6th-degree polynomial. Then, by taking the exponential of the fitted values from the regression, the GK-transformed variables in levels are obtained. With the variables now in levels, rather than a log–log form, there is no need to use the ex-post conversion factor in calculating government spending multipliers.

Appendix B. Data

All components of the US national income account are obtained from the National Income and Product Accounts (NIPA) tables available at Bureau of Economic Analysis website. The interest rate data is downloaded from the Federal Reserve Economic Data (FRED) website of the Federal Reserve Board of St.Louis. The population measure is from US Bureau of Labor Statistics.

All the components of national income are in real per capita terms and are transformed from the nominal values by dividing them by corresponding price indexes and population measure. Each data series is extracted as follows from the table and row numbers as per each organization's relevant data file as follows:

- Consumption: 'Personal Consumption Expenditures' (NIPA Table 1.1.5, Line 2).
- Investment: 'Gross Private Domestic Investment' (NIPA Table 1.1.5, Line 7).
- Consumption Tax Revenues: 'Federal Government Excise Taxes' (NIPA Table 3.2, Line 5) and 'Federal Government Customs Duties' (NIPA Table 3.2, Line 6).
- Capital and Labor Tax Revenues: As in Traum and Yang (2015), capital and labor tax revenues are constructed as follows. First, the average personal income tax rate, τ^p, is computed as

  $$\tau^p = \frac{IT}{W + PRI/2 + CI},$$

  where IT is 'Federal Government Personal Current Taxes' (NIPA Table 3.2, Line 3), W is 'Wage and Salary Accruals' (NIPA Table 1.12, Line 3), PRI is 'Proprietors' Income' (NIPA Table 1.12, Line 9), and CI is capital income which is the summation of 'Rental Income of Persons' (NIPA Table 1.12, Line 12), 'Corporate Profits' (NIPA Table 1.12, Line 13), 'Net Interest and Miscellaneous Payments' (NIPA Table 1.12, Line 18), and $PRI/2$.
  The labor income tax revenue is calculated as

  $$\tau^p(W + PRI/2) + CSI$$

  where CSI is 'Contributions for Government Social Insurance' (NIPA Table 3.2, Line 11).
  The capital income tax revenue is calculated as

  $$\tau^p CI + CT$$

  where CT is 'Taxes on Corporate Income' (NIPA Table 3.2, Line 7).
- Total Tax Revenues: Total tax revenues include consumption, capital, and labor tax revenues.

[18] In their study, the ex-post conversion factor is found to range from 2 to 24.

- Government Spending: ‘Federal Government Current Expenditures’ (NIPA Table 3.2, Line 20), ‘Federal Government Gross Investment on National Defense’ (NIPA Table 3.9.5, Line 13), ‘Federal Government Net Purchases of Non-produced Assets’ (NIPA Table 3.2, Line 44), and ‘Federal Government Gross Investment on Non-defense’ (NIPA Table 3.2, Line 18), minus ‘Federal Government Consumption of Fixed Capital’ (NIPA Table 3.2, Line 45).
- Government Debt: Government debt, B, is the market value of privately held gross federal debt, obtained from the Federal Reserve Bank of Dallas. The quarterly values are constructed from the monthly values at the beginning of each quarter.
- Hours Worked: Hours worked are constructed from the following variables:
  - H: the index for nonfarm business, all persons, average weekly hours duration, 2009 = 100, seasonally adjusted (from the Department of Labor).
  - Emp: civilian employment for sixteen years and over, measured in thousands, seasonally adjusted (from the Department of Labor, Bureau of Labor Statistics, CE16OV). The series is transformed into an index where 2009:Q3 = 100.
    Hours worked are then defined as
$$H = \frac{H \times Emp}{100}.$$
- Wage Rate: The wage rate is defined as the index for hourly compensation for nonfarm business, all persons, 2009 = 100, seasonally adjusted (from the U.S. Department of Labor)
- Inflation: The gross inflation rate is defined using the GDP deflator (NIPA Table 1.1.4, line 1).
- Federal Funds Rate: ‘Effective Federal Funds Rate’ (FRED Series ID: FEDFUNDS).
- Population: ‘Civilian Noninstitutional Population in Thousands, Ages 16 Years and Over, Seasonally Adjusted’ (BLS Series ID: LNS10000000).

Appendix C. DSGE model details

C.1. Summary of the model in Leeper et al. (2017)

The economy consists of a continuum of households indexed by $j \in [0,1]$. Among them, a fraction $1-\mu$ of households are savers (denoted by the superscript S) and a fraction μ are non-savers (denoted by the superscript N) as in Galí et al. (2007).

Savers maximize their expected lifetime utility, given by

$$E_0 \sum_{t=0}^{\infty} \beta^t u_t^b \left[\ln\left(C_t^{*S}(j) - \theta \tilde{C}_{t-1}^{*S}\right) - \left(L_t^S(j)^{1+\xi}\right)/(1+\xi) \right], \tag{11}$$

where β is the subjective discount factor, ξ is the inverse of Frisch elasticity of labor supply, and $L_t^S(j)$ is labor hours at time t by agent $j \in [0, 1-\mu]$. $C_t^{*S}(j)$ is agent j's composite consumption defined as $C_t^{*S}(j) = C_t^S(j) + \alpha_G G_t$, where $C_t^S(j)$ and G_t are private and public consumption goods, respectively. The parameter α_G determines the degree of substitutability of the consumption goods: when $\alpha_G < 0$, private and public consumption goods are complements; when $\alpha_G > 0$, the goods are substitutes. θ is the degree of consumption habit formation. Lastly, u_t^b is an *i.i.d.* normal preference shock with mean 0 and standard deviation σ_b.

Meanwhile, non-savers are assumed to have the same preferences as savers. They, however, behave in a “hand-to-mouth” fashion, in that they consume their entire per-period disposable income.

The government's behavior is summarized as follows: government spending, G_t, and transfers to households, Z_t, are financed by proportional taxes levied against consumption, C_t, labor income, L_t, and capital returns, K_t, and by issuing one-period nominal debt, B_t. Hence, the nominal flow government budget constraint is

$$P_t^B B_t + \tau_t^K R_t^K K_t + \tau_t^L W_t L_t + P_t \tau_t^C C_t = \left(1 + \rho P_t^B\right) B_{t-1} + P_t G_t + P_t Z_t, \tag{12}$$

where B_t is a portfolio of long-term nominal government bonds, and P_t^B denotes their price. The maturity of the bonds is assumed to decay at a constant rate $\rho \in [0,1]$, resulting in a duration of $(1-\beta\rho)^{-1}$. P_t represents the price of final goods, while R_t^K and W_t denote the rental rate on effective capital and nominal wage, respectively. Additionally, τ_t^K, τ_t^L, and τ_t^C denote the capital, labor, and consumption tax rates, respectively.

Fiscal variables are governed by the following processes:

$$\hat{G}_t = \rho_G \hat{G}_{t-1} - (1-\rho_G)\gamma_G \hat{s}_{t-1}^b + \epsilon_t^G, \tag{13}$$

$$\hat{\tau}_t^K = \rho_K \hat{\tau}_{t-1}^K + (1-\rho_K)\gamma_K \hat{s}_{t-1}^b + \epsilon_t^K, \tag{14}$$

$$\hat{\tau}_t^L = \rho_L \hat{\tau}_{t-1}^L + (1-\rho_L)\gamma_L \hat{s}_{t-1}^b + \epsilon_t^L, \tag{15}$$

$$\hat{\tau}_t^C = \rho_C \hat{\tau}_{t-1}^C + \epsilon_t^C, \tag{16}$$

$$\hat{Z}_t = \rho_Z \hat{Z}_{t-1} - (1-\rho_Z)\gamma_Z \hat{s}_{t-1}^b + \epsilon_t^Z, \tag{17}$$

where a hat (^) denotes the log deviation from the steady state. $s_{t-1}^b \equiv \left(P_{t-1}^B B_{t-1}\right) / \left(P_{t-1} Y_{t-1}\right)$ and $\epsilon_t^X \sim i.i.d.\ N(0, \sigma_X)$ where $X = \{G, K, L, C, Z\}$.

Table C.1
Prior and posterior distributions for the selected model parameters.

Parameter		Prior			Posterior	
		Fnc.	Mean	Std.	Median	[5%, 95%]
Preference						
θ	Habit formation in consumption	B	0.5	0.2	0.96	[0.84, 0.98]
μ	Fraction of non-savers	B	0.3	0.1	0.02	[0.01, 0.04]
α_G	Substitutability of private/public consumption	U	0	1.01	−0.11	[−0.42, 0.23]
Fiscal policy						
γ_K	Capital tax response to debt/GDP	N	0.15	0.1	0.19	[0.04, 0.33]
γ_L	Labor tax response to debt/GDP	N	0.15	0.1	0.16	[0.05, 0.27]
γ_G	Government spending response to debt/GDP	N	0.15	0.1	0.26	[0.09, 0.41]
γ_Z	Transfers response to debt/GDP	N	0.15	0.1	0.09	[−0.04, 0.23]
Monetary policy						
ρ_R	Lagged interest rate response	B	0.5	0.2	0.81	[0.74, 0.86]
ϕ_π	Interest rate response to inflation	N	1.5	0.2	1.40	[1.07, 1.73]
ϕ_y	Interest rate response to output	N	0.125	0.05	0.16	[0.11, 0.22]

The monetary authority sets interest rate policy according to the following Taylor-type rule

$$\hat{R}_t = \rho_R \hat{R}_{t-1} + (1-\rho_R)\left(\phi_\pi \hat{\pi}_t + \phi_y \hat{Y}_t\right) + \epsilon_t^R, \quad \epsilon_t^R \sim i.i.d.\ N(0, \sigma_R), \tag{18}$$

so that the nominal interest rate is adjusted in response to fluctuation in both output and inflation.

C.2. Parameter estimates

We estimate the DSGE model using Bayesian inference methods to construct the posterior distribution of the parameters, which combines the likelihood function and prior information. The model is estimated using 11 observables, representing US quarterly time series data from 1983:Q3 to 2007:Q4.

The prior distributions for the estimated parameters are set to align with Smets and Wouters (2007) and Leeper et al. (2017). Other parameters are fixed at well-established values; for instance, $\beta = 0.99$ (discount factor), $\alpha = 0.4$ (share of capital in the production function), and $\delta = 0.025$ (depreciation rate). Subsequently, the posterior distribution is sampled using the random walk Metropolis–Hastings (MH) algorithm. A million draws are simulated, with the first 400,000 used as a burn-in period and every 20th draw thinned, resulting in a sample size of 30,000.

Notice that the posterior parameters are crucial for understanding the impacts of government spending on output—often referred to as government spending multipliers. In particular, Leeper et al. (2017) demonstrate that the parameters governing the degree of habit formation in consumption, the fraction of non-savers, and the substitutability between private and public consumption are critical for determining the size of government spending multipliers in the short run. Meanwhile, the importance of identifying which fiscal instrument is at work in stabilizing government debt, raised by expansionary fiscal policy, or understanding how monetary policy behaves in times of fiscal expansion, tends to be significant in shaping the efficacy of government spending in the medium to longer runs, as underscored by Leeper et al. (2010) and Leeper et al. (2017).

The last two columns of Table C.1 display the median and 90th percentile estimates of the posterior distributions. The habit parameter estimates, denoted as θ, are higher than those reported in Smets and Wouters (2007), whose mean and 90th percentile are 0.71 and [0.64, 0.78], respectively, using the sample from 1966 to 2004. These values are also larger than the point estimate of 0.65 in Christiano et al. (2005). The posterior estimates for the fraction of non-savers, denoted as μ, are significantly low with median and 90th percentile estimates of 0.02 and [0.01, 0.04], respectively. These values are smaller than the estimates in Traum and Yang (2015), which are 0.18 and [0.10, 0.27], as well as the commonly used calibrated value of 0.5 as in Galí et al. (2007). The degree of substitutability between private and public consumption is estimated to be higher than the values in Leeper et al. (2017), whose mean estimate is −0.24. Our median estimates of −0.11 suggest that private and public consumption goods tend to be complementary.

Some of the fiscal policy parameters associated with government debt are rather weakly identified, as the posterior distributions for γ_L, γ_K, and γ_Z are close to their priors. Nevertheless, the response of government spending to the debt-to-GDP ratio is likely to be an exception, as its posterior distribution is higher than the prior. This finding suggests that initially deficit-financed government spending expansions tend to be financed through future reductions in government spending, as echoed by Corsetti et al. (2012). Finally, our estimated responses of the interest rate to inflation and output are 1.40 and 0.16, respectively, based on the median values.

Appendix D. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jmacro.2025.103721.

References

Auerbach, A.J., Gorodnichenko, Y., 2012. Measuring the output responses to fiscal policy. Am. Econ. J.: Econ. Policy 4 (2), 1–27. http://dx.doi.org/10.1257/pol.4.2.1.

Auerbach, A.J., Gorodnichenko, Y., 2013. Output spillovers from fiscal policy. Am. Econ. Rev. 103 (3), 141–146. http://dx.doi.org/10.1257/aer.103.3.141.

Barro, R.J., Redlick, C.J., 2011. Macroeconomic effects from government purchases and taxes. Q. J. Econ. 126 (1), 51–102. http://dx.doi.org/10.1093/qje/qjq002, URL: http://qje.oxfordjournals.org/cgi/doi/10.1093/qje/qjq002.

Ben Zeev, N., Pappa, E., 2017. Chronicle of a war foretold: The macroeconomic effects of anticipated defence spending shocks. Econ. J. 127 (603), 1568–1597. http://dx.doi.org/10.1111/ecoj.12349.

Ben Zeev, N., Ramey, V.A., Zubairy, S., 2023. Do government spending multipliers depend on the sign of the shock? AEA Pap. Proc. 113, 382–387. http://dx.doi.org/10.1257/pandp.20231071, URL: https://www.aeaweb.org/articles?id=10.1257/pandp.20231071.

Bianchi, F., 2013. Regime switches, agents' beliefs, and post-world war II U.S. macroeconomic dynamics. Rev. Econ. Stud. 80 (2), 463–490. http://dx.doi.org/10.1093/restud/rds032.

Blanchard, O., Perotti, R., 2002. An empirical characterization of the dynamic effects of changes in government spending and taxes on output. Q. J. Econ. 117 (4), 1329–1368. http://dx.doi.org/10.1162/003355302320935043, URL: http://qje.oxfordjournals.org/content/117/4/1329.abstract.

Brugnolini, L., 2018. About Local Projection Impulse Response Function Reliability. CEIS Working Paper, Vol. 440, pp. 1–38. http://dx.doi.org/10.2139/ssrn.3137229.

Caldara, D., Kamps, C., 2008. What are the Effects of Fiscal Policy Shocks? A VAR-Based Comparative Analysis. ECB Working Paper Series No. 2008-877.

Christiano, L.J., Eichenbaum, M., Evans, C.L., 1999. Monetary policy shocks: What have we learned and to what end? In: Handbook of Macroeconomics. Vol. 1, Elsevier, pp. 65–148.

Christiano, L.J., Eichenbaum, M., Evans, C.L., 2005. Nominal rigidities and the dynamic effects of a shock to monetary policy. J. Political Econ. 113 (1), 1–45. http://dx.doi.org/10.1086/426038.

Clarida, R., Galí, J., Gertler, M., 2000. Monetary policy rules and macroeconomic stability: Evidence and some theory. Q. J. Econ. 115 (1), 147–180. http://dx.doi.org/10.1162/003355300554692.

Corsetti, G., Meier, A., Müller, G.J., 2012. Fiscal stimulus with spending reversals. Rev. Econ. Stat. 94 (4), 878–895. http://dx.doi.org/10.1162/REST_a_00223.

Dupor, B., Li, R., 2015. The expected inflation channel of government spending in the postwar US. Eur. Econ. Rev. 74, 36–56. http://dx.doi.org/10.1016/j.euroecorev.2014.11.001.

Fatás, A., Mihov, I., 2001. The Effects of Fiscal Policy on Consumption and Employment: Theory and Evidence. INSEAD and CEPR Discussion Paper 2760.

Feve, P., Matheron, J., Sahuc, J.-G., 2013. A pitfall with estimated DSGE-based government spending multipliers. Am. Econ. J.: Macroecon. 5 (4), 141–178.

Galí, J., López-Salido, J.D., Vallés, J., 2007. Understanding the effects of government spending on consumption. J. Eur. Econ. Assoc. 5 (1), 227–270. http://dx.doi.org/10.1162/JEEA.2007.5.1.227.

Gechert, S., Rannenberg, A., 2018. Which fiscal multipliers are regime-dependent? A meta-regression analysis. J. Econ. Surv. 32 (4), 1160–1182. http://dx.doi.org/10.1111/joes.12241.

Gordon, R.J., Krenn, R., 2010. The End of the Great Depression: VAR Insight on the Roles of Monetary and Fiscal Policy. NBER Working Paper No. 16380, pp. 183–231. http://dx.doi.org/10.3386/w16380.

Jordà, Ò., 2005. Estimation and inference of impulse responses by local projections. Am. Econ. Rev. 95 (1), 161–182. http://dx.doi.org/10.1257/0002828053828518.

Kilian, L., Lütkepohl, H., 2017. Structural Vector Autoregressive Analysis. Cambridge University Press.

Leeper, E.M., Traum, N., Walker, T.B., 2017. Clearing up the fiscal multiplier morass. Am. Econ. Rev. 107 (8), 2409–2454. http://dx.doi.org/10.1257/aer.20130330.

Leeper, E.M., Walker, T.B., Yang, S.-C.S., 2010. Government investment and fiscal stimulus. J. Monet. Econ. 57 (8), 1000–1012. http://dx.doi.org/10.1016/j.jmoneco.2010.09.002.

Li, D., Plagborg-Møller, M., Wolf, C.K., 2024. Local projections vs. VARs: Lessons from thousands of DGPs. J. Econometrics 105722. http://dx.doi.org/10.1016/j.jeconom.2023.105722.

Lubik, T.A., Schorfheide, F., 2004. Testing for indeterminacy: An application to US monetary policy. Am. Econ. Rev. 94 (1), 190–217. http://dx.doi.org/10.1257/000282804322970728.

Meier, A., 2005. How Big is the Bias in Estimated Impulse Responses? A Horse Race between VAR and Local Projection Methods. European University Institute, Manuscript.

Mertens, K., Ravn, M.O., 2013. A reconciliation of SVAR and narrative estimates of tax multipliers. J. Monet. Econ. http://dx.doi.org/10.1016/j.jmoneco.2013.04.004.

Mountford, A., Uhlig, H., 2009. What are the effects of fiscal policy shocks? J. Appl. Econometrics 24 (6), 960–992. http://dx.doi.org/10.1002/jae.1079.

Nakamura, E., Steinsson, J., 2018. Identification in macroeconomics. J. Econ. Perspect. 32 (3), 59–86. http://dx.doi.org/10.3386/w23968.

Olea, J.L.M., Pflueger, C., 2013. A robust test for weak instruments. J. Bus. Econom. Statist. 31 (3), 358–369. http://dx.doi.org/10.1080/07350015.2013.818890.

Owyang, M.T., Ramey, V.A., Zubairy, S., 2013. Are government spending multipliers greater during periods of slack? Evidence from twentieth-century historical data. Am. Econ. Rev. 103 (3), 129–134. http://dx.doi.org/10.1257/aer.103.3.129.

Perotti, R., 2005. Estimating the Effects of Fiscal Policy in OECD Countries. Federal Reserve Bank of San Francisco, Proceedings.

Plagborg-Møller, M., Wolf, C.K., 2021. Local projections and VARs estimate the same impulse responses. Econometrica 89, 955–980. http://dx.doi.org/10.3982/ECTA17813, URL: https://www.econometricsociety.org/doi/10.3982/ECTA17813.

Ramey, V.A., 2011. Identifying government spending shocks: It's all in the timing. Q. J. Econ. 126 (1), 1–50. http://dx.doi.org/10.1093/qje/qjq008.

Ramey, V.A., 2016. Defense News Shocks, 1889–2015: Estimates Based on News Sources. University of California, San Diego, Manuscript.

Ramey, V.A., 2019. Ten years after the financial crisis: What have we learned from the renaissance in fiscal research? J. Econ. Perspect. 33 (2), 89–114. http://dx.doi.org/10.1257/jep.33.2.89.

Ramey, V.A., Zubairy, S., 2018. Government spending multipliers in good times and in bad: evidence from US historical data. J. Political Econ. 126 (2), 850–901. http://dx.doi.org/10.1086/695404.

Rossi, B., Zubairy, S., 2011. What is the importance of monetary and fiscal shocks in explaining U.S. macroeconomic fluctuations? J. Money Credit. Bank. 43 (6), 1247–1270. http://dx.doi.org/10.1111/j.1538-4616.2011.00424.x.

Smets, F., Wouters, R., 2003. An estimated dynamic stochastic general equilibrium model of the Euro area. J. Eur. Econ. Assoc. 1 (5), 1123–1175. http://dx.doi.org/10.1162/154247603770383415.

Smets, F., Wouters, R., 2007. Shocks and frictions in US business cycles: A Bayesian DSGE approach. Am. Econ. Rev. 97 (3), 586–606. http://dx.doi.org/10.1257/aer.97.3.586.

Stock, J.H., Watson, M.W., 2018. Identification and estimation of dynamic causal effects in macroeconomics using external instruments. Econ. J. 128 (610), 917–948. http://dx.doi.org/10.1111/ecoj.12593.

Traum, N., Yang, S.-C.S., 2015. When does government debt crowd out investment? J. Appl. Econometrics 30 (1), 24–45. http://dx.doi.org/10.1002/jae.2346.

Walker, T.B., Leeper, E.M., 2011. Information flows and news driven business cycles. Rev. Econ. Dyn. 14 (1), 55–71. http://dx.doi.org/10.1016/j.red.2010.08.003.