NBER WORKING PAPER SERIES

DO GOVERNMENT SPENDING MULTIPLIERS DEPEND ON THE SIGN OF THE SHOCK?

Nadav Ben Zeev
Valerie A. Ramey
Sarah Zubairy

Working Paper 31015
http://www.nber.org/papers/w31015

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2023

This working paper is an extended version of a paper prepared for the American Economic Association Papers and Proceedings. We thank Regis Barnichon, Davide Debortoli, Christian Matthes, James Hamilton and participants at the AEA conference session for their comments. Valerie Ramey gratefully acknowledges financial support from National Science Foundation Grant No. 1658796. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Do Government Spending Multipliers Depend on the Sign of the Shock?
Nadav Ben Zeev, Valerie A. Ramey, and Sarah Zubairy
NBER Working Paper No. 31015
March 2023
JEL No. E62,N12

ABSTRACT

We analyze whether government spending multipliers differ by the sign of the shock. Using aggregate historical U.S. data, we apply Ben Zeev's (2020) nonlinear diagnostic tests and find evidence of nonlinearities in the impulse response functions of both government spending and GDP. We then extend Ramey and Zubairy's (2018) framework to allow for asymmetric effects as a type of state dependence to estimate multipliers. While we find differences in the impulse response functions, the resulting multipliers do not differ by sign of the shock. Thus, we find no evidence of asymmetry of government spending multipliers.

Nadav Ben Zeev
Department of Economics
Ben-Gurion University of the Negev,
Beer-Sheva, Israel
Israel
nadavbz@bgu.ac.il

Valerie A. Ramey
Department of Economics, 0508
University of California, San Diego
9500 Gilman Drive
La Jolla, CA 92093-0508
and NBER
vramey@ucsd.edu

Sarah Zubairy
Department of Economics
Texas A&M University
4228 TAMU
College Station, TX 77843
and NBER
szubairy@tamu.edu

A supplemental appendix is available at http://www.nber.org/data-appendix/w31015

1 Introduction

Much recent attention has been devoted to estimating the size of government purchases multipliers. Part of that literature has explored whether multipliers are different during recessions or when monetary policy is constrained by the zero lower bound (e.g. Auerbach and Gorodnichenko (2012), Owyang, Ramey and Zubairy (2013), Ramey and Zubairy (2018)). An older literature explored asymmetry, i.e., whether declines in government spending have larger effects on economic activity than rises in government spending. For example, Hooker and Knetter (1997) found that military cutbacks had larger effects than military buildups on state economies and Davis, Loungani and Mahidhara (1997) found asymmetries in the effects of both oil shocks and government spending shocks on regional economies.

Recently, Barnichon, Debortoli and Matthes (2022) (BDM) offered evidence for asymmetric effects of government spending at the aggregate level. Using the Functional Approximations to Impulse Responses (FAIR) method of Barnichon and Matthes (2018) on aggregate U.S. data, BDM present evidence that negative shocks to government spending result in larger multipliers than positive shocks. Depending on the sample and identification method, they estimate multipliers between 0.3 and 0.8 for positive shocks to government spending and 1.4 for negative shocks. They also show robustness checks using nonlinear local projections.

Asymmetric government spending multipliers have first-order policy implications. If multipliers on rises in government spending are indeed less than multipliers on declines in government spending, then any government spending package that is not permanent will have a net negative effect on output. That is, the positive effect on output of an increase in government spending would be dominated by the negative effect of the unwinding of government spending. This result would imply two costs of a rise in government spending: the standard cost of eventually raising taxes to finance the spending plus the amplified negative effects of the wind-down of government spending. Thus, it is important to determine whether such asymmetries exist.

In this paper, we reexamine the evidence for asymmetric government spending multipliers in aggregate data. We first apply Ben Zeev's (2020) nonlinear diagnostic tests and find evidence of nonlinearities in the impulse response functions of both government spending and GDP. Since differences in impulse response functions do not necessarily translate into differences in multipliers, we explore the issue further by extending

Ramey and Zubairy's (2018) framework to allow for asymmetric effects as a type of state dependence. While we find differences in the individual impulse response functions for positive versus negative shocks, the resulting multipliers do not differ by the sign of the shock.

We compare our results with those of BDM for our large historical sample and find that our local projection method produces more precise estimates of multipliers than their FAIR method, which is based on approximating the underlying impulse response functions. There is no evidence of difference in multipliers by sign in local projections. On balance, we conclude that the evidence for asymmetry is weak.

2 Data

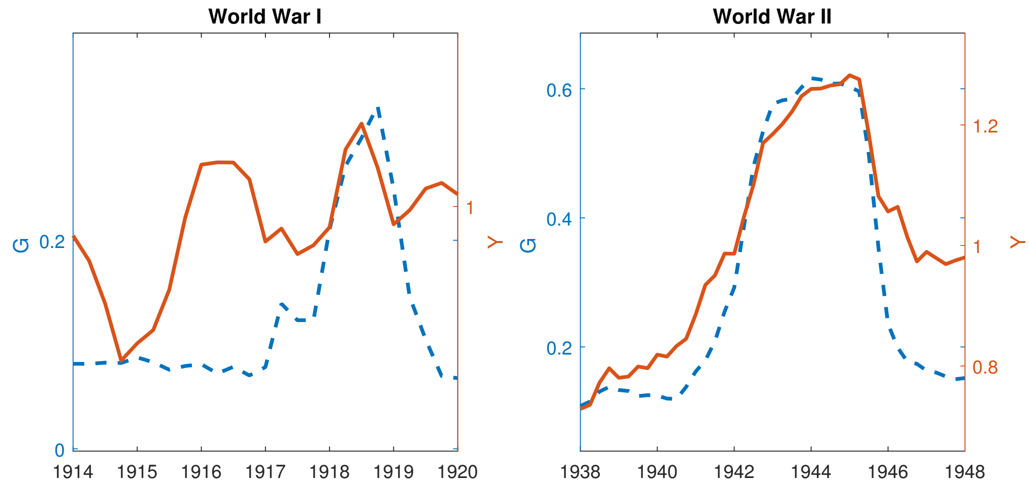
We use a subset of the data constructed by Ramey and Zubairy (2018), which consists of quarterly data from 1889 to 2015 on GDP, government purchases, a narrative military news series, and an estimate of potential GDP. Real government spending includes federal, state, and local purchases and is derived by deflating nominal expenditures with the GDP deflator in order to compare a dollar of government spending to a dollar of GDP. The military news series, which is constructed from narrative evidence, consists of changes in the expected present discounted value of the path of government purchases (Ramey (2016)).

In the baseline model, nominal government spending and GDP are first deflated by the GDP price deflator and then divided by an estimate of potential GDP, based on a sixth degree polynomial trend fit over the sample excluding the Great Depression and WWII (1930 - 1946). Military news is divided by the lagged deflator and potential GDP, since the contemporaneous deflator might be affected by military news. We also conduct a robustness check that uses an alternative transformation that does not require an estimate of potential GDP.

3 A Case Study of Two World Wars

To motivate our skepticism of asymmetric effects on multipliers, we first review events during two influential episodes in the U.S. historical data — the two World Wars.

Figure 1. Case Study of Two World Wars.



Note: Real government purchases (G, dashed) and GDP (Y, solid) divided by potential GDP. Source: Ramey and Zubairy (2018) data.

Figure 1 shows the behavior of real government purchases and real GDP, both divided by potential GDP.

In each case, government spending rose when the U.S. became involved in the war. At the end of the war, real government spending returned to its pre-war fraction of potential GDP in less than a year. GDP rose and fell along with government spending. If anything, GDP rose more with the rise in government spending than it fell with the decline in government spending at the end of the war. In both cases, the recession at the end of the war was shallow and brief.[1] Thus, there is no evidence in the raw data suggesting that declines in government spending have greater effects than rises in government spending.

4 Diagnostic Tests for Nonlinearities

We conduct some initial diagnostic tests for nonlinearity using Ben Zeev's (2020) polynomial test, augmented with Forni et al.'s (2022) method for differentiating nonlinearities due to sign versus size. Impulse response functions (IRFs) are estimated

1. In the case of WWI, a deeper recession followed in 1920-21 but is usually attributed to severe monetary tightening (e.g. Anderson and Chang (2022).

using local projections (Jordà (2005)) on a set of regressions for each horizon h, from 0 to 20 quarters, as follows:

$$(1) \qquad news_t = \delta(L)z_{t-1} + \eta_t$$

$$(2) \qquad x_{i,t+h} = \alpha_{i,h}\hat{\eta}_t + \theta_{i,h}f(\hat{\eta}_t) + \psi_{i,h}(L)z'_{t-1} + \zeta_{i,t+h}, \quad for\ i = g,y$$

In Equation 1, *news* is the military news variable and z consists of a constant term plus four lags of news, government spending, and GDP, all transformed as described in the data section. Equation 2 represents two additional equations where x is government spending (g) in one equation and GDP (y) in the other. $f(\hat{\eta}_t)$ is the nonlinear term in the shock.[2] The z' is z augmented with lags of the nonlinear shock term.

We first test the null hypothesis $\theta = 0$ against the quadratic alternative $f(\hat{\eta}_t) = \hat{\eta}_t^2$. We reject the null hypothesis at the 5% (10%) level for 19 (20) of 21 horizons for government spending and 14 (17) for GDP, suggesting nonlinearities in the IRFs for both variables. The online appendix shows the estimated IRFs.

However, the quadratic term could be picking up sign or size effects. To distinguish asymmetric effects from size nonlinearities, we conduct a test using Forni et al.'s (2022) absolute value term, $f(\hat{\eta}_t) = |\hat{\eta}_t|$, which captures only asymmetry. We find that this term is also significant at almost all horizons for both variables. Following Forni et al. (2022), we run a horse race between the quadratic and absolute value terms by including both terms in the model. The correlation between the two nonlinear terms is 0.9, so multicollinearity results in neither being individually significant in 28 of 42 cases. The absolute value term is significant in all 14 of the others cases, but the quadratic term is significant in only 2 cases. These results suggest that sign not size is the main source of nonlinearity.

However, differences in IRFs do not imply differences in multipliers since the multiplier is based on the ratio of the IRFs. To investigate whether multipliers differ by the sign of the shock, the next section develops a framework for estimating both IRFs and multipliers when there are asymmetries.

2. We use the innovation in the news equation as the shock because the polynomial method requires a mean-zero shock. In the next section, we use news itself.

5 State-Dependent Local Projections

5.1 Econometric Model

Asymmetry can easily be modeled by redefining the state in Ramey and Zubairy's (2018) (RZ) framework. A key advantage of the RZ framework is the equivalence of the three-step and one-step estimates of multipliers, which facilitates estimation of standard errors on multipliers and tests of equality of multipliers.

The impulse response functions are estimated with a set of regressions for each horizon h using the following model:

$$x_{i,t+h} = I_t^+ \left[\beta_{i,h}^+ news_t + \phi_{i,h}^+(L)z_{t-1}\right] + I_t^- \left[\beta_{i,h}^- news_t + \phi_{i,h}^-(L)z_{t-1}\right] + \varepsilon_{i,t+h}, \quad \textit{for } i = g,y \textit{ and } h = 0,\ 1,\ \ldots,\ H \tag{3}$$

Here x is either government spending (g) or GDP (y) and I^+ is a dummy variable for $news_t > 0$ and I^- is its complement. z consists of a constant term and four lags of government spending, GDP, and news. All the coefficients of the model are allowed to differ according to whether the contemporaneous shock is positive or negative; this flexibility is a necessary condition for the cumulative multiplier computed from the estimated IRFs to be equivalent to the one-step IV multiplier defined below.

The cumulative multiplier through horizon h is the ratio of the integral under the GDP IRF to the integral under the government spending IRF, i.e., the multiplier for positive shocks is $m^+ = \left(\sum_{j=0}^{h} \beta_{y,h}^+\right) / \left(\sum_{j=0}^{h} \beta_{g,h}^+\right)$ and similarly for negative shocks. This three-step method for computing multipliers produces point estimates, but obtaining standard errors and doing tests for the equality of the two multipliers is cumbersome. An easier method is the one-step local projection-instrumental variables (LP-IV) method introduced by Ramey and Zubairy (2018). This procedure involves IV estimation of a regression of the cumulative sum of GDP on the cumulative sum of government spending using the shocks as instruments. In particular, we estimate

$$\sum_{j=0}^{h} y_{t+j} = m_h^+ \left(I_t^+ \sum_{j=0}^{h} g_{t+j}\right) + m_h^- \left(I_t^- \sum_{j=0}^{h} g_{t+j}\right) + I_t^+[\gamma_h^+(L)z_{t-1}] + I_t^-[\gamma_h^-(L)z_{t-1}] + \omega_{t+h}, \quad \textit{for } h = 0,\ 1,\ \ldots,H \tag{4}$$

using $news_t^+$ ($=I_t^+ news_t$) and $news_t^-$ ($=I_t^- news_t$) as instruments for the terms in parenthesis. The cumulative multiplier through horizon h is the coefficient m_h^+ for positive shocks and m_h^- for negative shocks.[3]

5.2 Baseline Results

Figure 2 shows the estimates of the impulse responses, multipliers, and first-stage F-statistics. The impulse response estimates imply that a unit-magnitude negative shock leads to much larger responses of both government spending and GDP, though the estimates are less precise for negative shocks than positive shocks. However, the lower left panel of the figure shows that these differences do not translate into significant differences in multipliers since both the numerator and denominator of the multiplier increase roughly proportionally for negative shocks. We fail to reject equality of the multipliers for all horizons other than the first couple quarters.

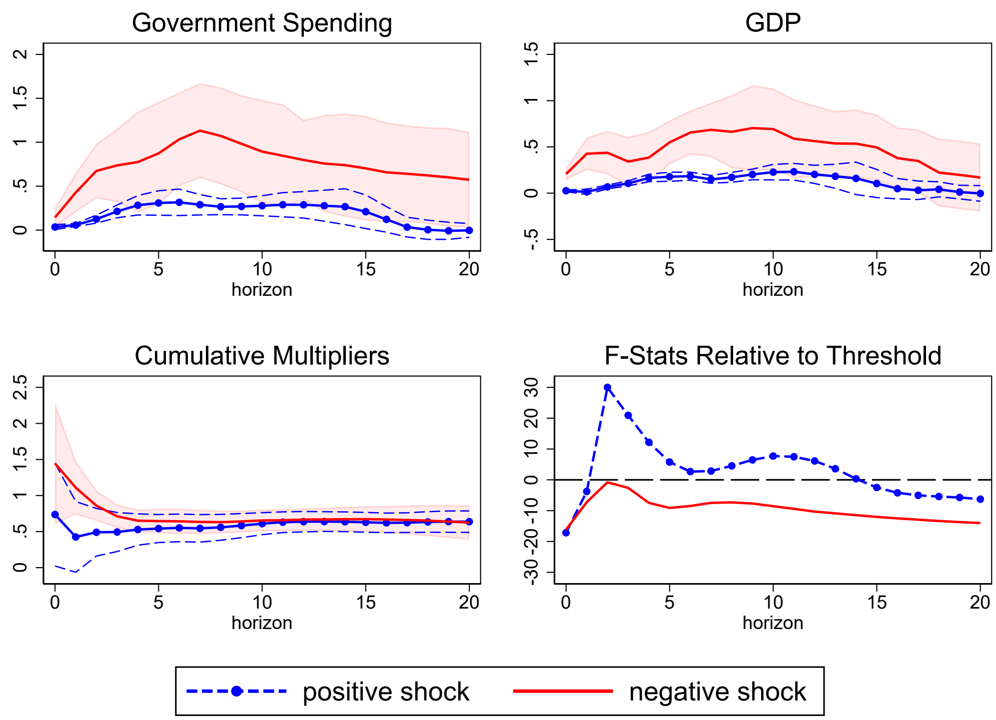
The higher multipliers in the first few quarters occur because GDP rises in response to a military news shock before government spending increases; this is known as the *anticipation effect*. Brunet (2022) argues that this effect occurs because government purchases are not recorded in the national accounts until the goods have been delivered and she presents evidence showing that annual budget authority leads annual spending. Briganti and Sellemi (2022) present higher frequency evidence in support of her hypothesis using data on military contracts and inventories. They document that the rise in GDP appears as inventory investment in the national accounts in the short run and in government purchases only after the goods have been delivered and paid for several quarters later.[4]

The lower right panel of Figure 2 shows the F-statistics for the first-stage regression of LP-IV cumulative multiplier estimation. Because the military news variable is based on changes in defense spending due to political events, it should be exogenous to the economy. The question is then of instrument relevance. Since the Staiger and Stock (1997) rule-of-thumb threshold of 10 does not apply to cases of one or two instruments or when there might be serial correlation, we instead use the Montiel Olea and Pflueger (2013) effective F-statistics and thresholds, following the approach taken in Ramey and

3. We use Newey and West (1987) corrections of the standard errors in IRF and multiplier regressions rather than lag-augmented regressions because of apparent additional sources of serial correlation. The standard errors are also robust to heteroscedasticity

4. This accounting feature also explains why Cholesky decompositions on government spending get the timing of the shock wrong, as documented by Ramey (2011)

Figure 2. Effects of Military News: Baseline Results



Note: 95% confidence bands.

Zubairy (2018). The lower right panel shows the *difference* between the first-stage effective F-statistics and the Montiel Olea and Pflueger (2013) thresholds.[5] A value above 0 means that the effective F-statistic exceeds the threshold. The F-statistics are above the threshold for most horizons for the positive shock, but are just below at all horizons for the negative shock.

In sum, the findings shown in Figure 2 are similar in spirit to Ramey and Zubairy's (2018) finding for slack states: news shocks during both slack and negative shock states generate bigger changes in government spending. However, GDP rises proportionally, so there is no difference in multipliers.

5. We use the threshold for the 5 percent critical value for testing the null hypothesis that the TSLS bias exceeds 10 percent of the OLS bias. For one instrument, this threshold is always 23.1. The threshold is 19.7 percent for the ten percent critical value.

5.3 Robustness Checks

We explored a number of alternative specifications to check robustness of our results. These alternative specifications include using Ben Zeev's (2020) quadratic approximation model, using innovations to military news rather than news itself, and including additional controls such as taxes. The results, which are reported in the online appendix, indicate no difference between multipliers according to the sign of the shock.

We show the results of one additional robustness check here and another in the next section that compares our findings to those of Barnichon, Debortoli and Matthes (2022). The robustness check we show here is for an alternative transformation of the three key variable, government spending, GDP, and military news.

When estimating multipliers, one should not use logarithms of variables since they can lead to biased multiplier estimates (Owyang, Ramey and Zubairy (2013)). However, since GDP and government spending have exponential trends, some kind of transformation is required. Our baseline specification follows Ramey and Zubairy (2018) in using a Gordon and Krenn (2010) transformation, which normalizes GDP, government spending, and military news by an estimate of potential GDP. However, that normalization implicitly does not allow government spending to affect potential GDP since potential GDP is estimated using a sixth-order polynomial deterministic trend. To avoid this assumption, as well as the use of any deterministic trends, we re-estimate our model using the transformation employed by Hall (2009) and Barro and Redlick (2011). This transformation uses cumulative differences of real government spending and GDP, and divides both by lagged real GDP. We also divide the military news variable by lagged nominal GDP (since military news is nominal). In particular, the local projections for estimating the IRFs are:

$$\frac{x_{i,t+h} - x_{i,t-1}}{y_{t-1}} = I_t^+ \left[\beta_{i,h}^+ news_t + \phi_{i,h}^+(L) z_{t-1}\right] + I_t^- \left[\beta_{i,h}^- news_t + \phi_{i,h}^-(L) z_{t-1}\right] + \varepsilon_{i,t+h}, \tag{5}$$

$$\textit{for } i = g, y \textit{ and } h = 0,\ 1,\ \ldots,\ H$$

Here the x is either real government spending (g) or real GDP (y), with no potential GDP normalization as in the baseline. z consists of a constant term and lagged values of military news (divided by lagged nominal GDP), as well as lags of first differences of real government spending and GDP, both divided by lagged real GDP. The equation for the one-step LP-IV estimate of the cumulative multipliers using the Hall-Barro-Redlick transformation is:

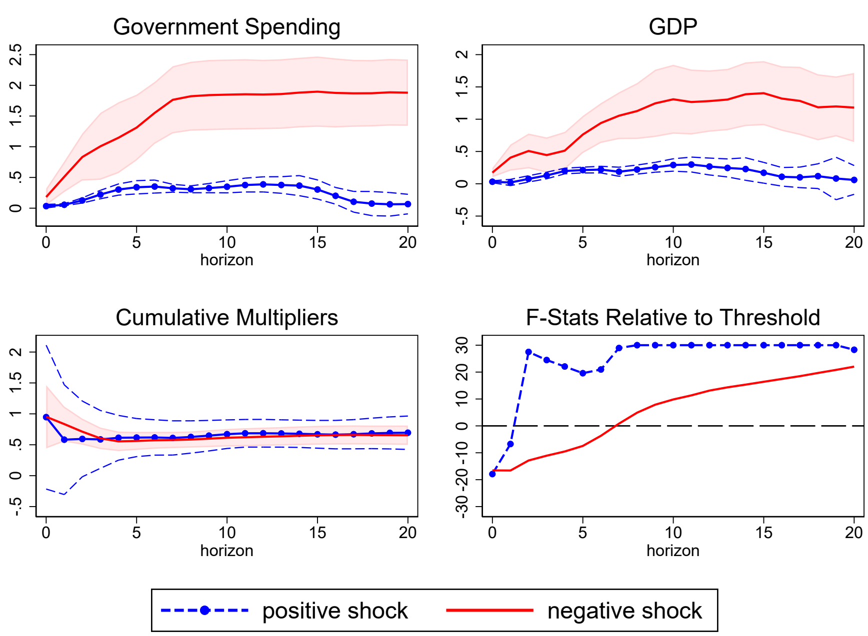
$$\sum_{j=0}^{h} \frac{y_{t+j} - y_{t-1}}{y_{t-1}} = m_h^{+} \left(I_t^{+} \sum_{j=0}^{h} \frac{g_{t+j} - g_{t-1}}{y_{t-1}} \right) + m_h^{-} \left(I_t^{-} \sum_{j=0}^{h} \frac{g_{t+j} - g_{t-1}}{y_{t-1}} \right) \tag{6}$$

$$+ I_t^{+}[\gamma_h^{+}(L) z_{t-1}] + I_t^{-}[\gamma_h^{-}(L) z_{t-1}] + \omega_{t+h}, \;\; \textit{for } h = 0, 1, \ldots, H$$

using $news_t^{+}$ and $news_t^{-}$ as instruments for the terms in parenthesis.

Figure 3 shows the results. The impulse responses of government spending and GDP to the negative shocks suggest very persistent effects on both series, whereas the positive shocks display more mean reversion. However, as the lower left panel shows, both imply the same values of cumulative multipliers by sign of shock, similar to our baseline.

Figure 3. Results using Hall-Barro-Redlick Transformation



Note: Responses to military news using the Hall-Barro-Redlick transformation of variables in our baseline specification.

Interestingly, the first-stage F-statistics shown in the lower right panel are significantly higher for both positive and negative shocks at most horizons relative to our baseline specification. Thus, the evidence for instrument relevance is stronger with the Hall-Barro-Redlick transformation than with the Gordon-Krenn transformation in this application.

6 Comparison to BDM Results

Our results contrast with Barnichon, Debortoli and Matthes's (2022) (BDM) results, which estimate higher multipliers for negative shocks. In this section we report the results of our investigation of possible sources for the differences. We limited our comparisons to their specifications that use Ramey and Zubairy's (2018) data, historical sample, and narrative military news shock. More details of the results are reported in our online appendix.

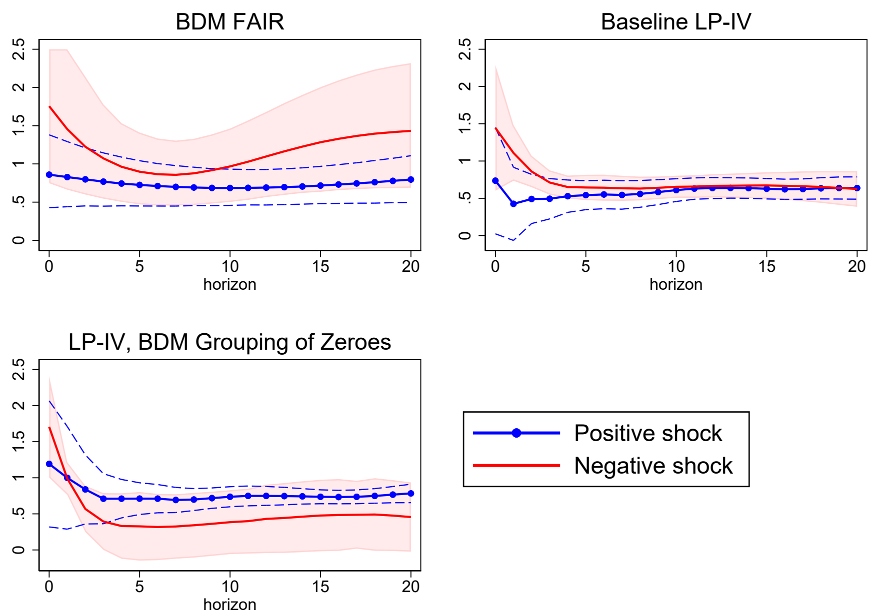
BDM use functional approximations to impulse responses (FAIR) to estimate impulse responses and then construct multipliers from those. As BDM note, the FAIR method likely induces bias, but they advocate its use based on efficiency gains.

The upper left panel of Figure 4 shows our replication of BDM's FAIR results. The results suggest a multiplier in response to a negative shock that starts higher initially, converges toward the positive shock multiplier during the second year, and then sweeps back up again as the horizon grows. The upper right panel reproduces our baseline multiplier estimates for comparison. Our estimates also show that the negative multiplier is initially higher, due to the anticipation effect discussed above, but it quickly falls to the level of the positive multiplier after the first few horizons.

A comparison of the 95% confidence bands for our baseline LP-IV with the 95% posterior probabilities for Barnichon, Debortoli and Matthes's (2022) FAIR estimates suggests that our LP-IV estimates are much more precise. Thus, despite using approximations that are parsimonious in parameters, the FAIR method does not improve upon the LP-IV precision.

However, part of the difference in estimates and confidence bands across the two graphs is how negative and positive shocks are defined. 78 percent of the military news observations are zeroes, so it matters how they are grouped. Based on the paucity of strictly negative shocks (only 6.5 percent of the sample) and the low F-statistics for the first stage for strictly negative shocks, we grouped the zeroes with the negative

Figure 4. A Comparison of Multiplier Estimation Precision



Note: 95% bands. Baseline LP-IV includes zeroes with the negative shock. The other two specifications includes the zeroes with the positive shock.

shocks in our baseline model. In contrast, BDM grouped the zeroes with the positive shocks. The bottom right panel of Figure 4 shows our LP-IV model estimates but now using BDM's grouping of the zeroes with the positive news values. A comparison with our baseline model shows that BDM's grouping widens the confidence bands for the negative shocks but does not narrow them for the positive shocks. Nevertheless, these alternative LP-IV estimates are still more precise than the FAIR estimates on average. Our online appendix shows the average standard errors of the estimates across horizons for several specifications.

As shown by Plagborg-Møller and Wolf (2021), VARs and LPs estimate the same impulse response functions in population, so with our controls similar to a VAR equation, we likely have efficiency gains over the moving-average representation of BDM because we have such a large sample (500 observations). Also, the cumulative multipliers are complicated nonlinear functions of IRFs, so any precision benefits of the FAIR approximation to the IRFs may not translate to the cumulative multipliers.

BDM check for possible bias in their FAIR IRFs estimates by estimating a local projections model and again find differences in point estimates of multipliers by sign. Our examination of their replication programs revealed some issues in implementation. For example, they were unable to formally test for equality of multipliers by sign of shock because they did not set up their local projections so that the one-step LP-IV would be equivalent to the three-step procedure. Second, our investigations found that the rise in multipliers at the later horizons on which they focused owed to their inclusion of a quartic trend in a system that had already normalized government spending and output by an estimate of potential GDP based on a sixth-order polynomial. This concern about deterministic trends possibly distorting the results is what led us to check robustness in the last section using the Hall-Barro-Redlick transformation, since it involved no normalization or deterministic trends. That analysis revealed persistence in the responses of both government spending and GDP to the negative shocks. However, those differences in IRFs did not translate into differences in multipliers. We discuss some of these issues in more detail in our online appendix.

7 Conclusion

In this paper, we have reexamined the question of whether government spending multipliers depend on the sign of the shock. Our tests did find evidence of sign-dependence of impulse response functions. However, when we estimated impulse response functions and multipliers in a state-dependent local projection model, we found no significant differences in multipliers by sign of shock, either quantitatively or statistically. Furthermore, we found that the multipliers were estimated more precisely by our state-dependent local projections model than by the FAIR method. In sum, our estimates suggest no asymmetry in government spending multipliers.

References

Anderson, Haelim, and Jin-Wook Chang. 2022. "Labor Market Tightness during WWI and the Postwar Recession of 1920-1921." *FEDS Working Paper*.

Auerbach, Alan J, and Yuriy Gorodnichenko. 2012. "Measuring the output responses to fiscal policy." *American Economic Journal: Economic Policy*, 4(2): 1–27.

Barnichon, Regis, and Christian Matthes. 2018. "Functional approximation of impulse responses." *Journal of Monetary Economics*, 99: 41–55.

Barnichon, Regis, Davide Debortoli, and Christian Matthes. 2022. "Understanding the size of the government spending multiplier: It's in the sign." *The Review of Economic Studies*, 89(1): 87–117.

Barro, Robert, and Charles Redlick. 2011. "Macroeconomic Effects From Government Purchases and Taxes." *The Quarterly Journal of Economics*, 126(1): 51–102.

Ben Zeev, Nadav. 2020. "Identification of Sign Dependency of Impulse Responses." *Ben-Gurion University of Negev working paper*.

Briganti, Edoardo, and Victor Sellemi. 2022. "Why Does GDP Move Before G? It's All in the Measurement." *UCSD working paper*.

Brunet, Gillian. 2022. "When Does Government Spending Matter? Evidence from a New Measure of Spending." *Working Paper*.

Davis, Steven J, Prakash Loungani, and Ramamohan Mahidhara. 1997. "Regional labor fluctuations: oil shocks, military spending, and other driving forces." *Board of Governors of the Federal Reserve System International Finance Disc. Papers*, 578.

Forni, Mario, Luca Gambetti, Nicolò Maffei-Faccioli, and Luca Sala. 2022. "The Nonlinear Transmission of Financial Shocks: Some Evidence." *Norges Bank working paper*.

Gordon, Robert J., and Robert Krenn. 2010. "The End of the Great Depression: VAR Insight on the Roles of Monetary and Fiscal Policy." NBER Working paper 16380.

Hall, Robert E. 2009. "By How Much Does GDP Rise If the Government Buys More Output?" *Brookings Papers on Economic Activity*, 40(2 (Fall)): 183–249.

Hooker, Mark A, and Michael M Knetter. 1997. "The effects of military spending on economic activity: Evidence from state procurement spending." *Journal of Money, Credit, and Banking*, 400–421.

Jordà, Òscar. 2005. "Estimation and Inference of Impulse Responses by Local Projections." *American Economic Review*, 95(1): 161–182.

Montiel Olea, Jose Luis, and Carolin Pflueger. 2013. "A Robust Test for Weak Instruments." *Journal of Business & Economic Statistics*, 31(3): 358–369.

Newey, Whitney K, and Kenneth D West. 1987. "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica*, 55(3): 703–08.

Owyang, Michael T, Valerie A Ramey, and Sarah Zubairy. 2013. "Are government spending multipliers greater during periods of slack? Evidence from twentieth-century historical data." *American Economic Review*, 103(3): 129–34.

Plagborg-Møller, Mikkel, and Christian K. Wolf. 2021. "Local Projections and VARs Estimate the Same Impulse Responses." *Econometrica*, 89(2): 955–980.

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

Ramey, Valerie A. 2016. "Defense News Shocks, 1889–2015: Estimates Based on News Sources." University of California, San Diego Unpublished paper.

Ramey, Valerie A., and Sarah Zubairy. 2018. "Government spending multipliers in good times and in bad: evidence from US historical data." *Journal of Political Economy*, 126(2): 850–901.

Staiger, Douglas, and James H. Stock. 1997. "Instrumental Variables Regression with Weak Instruments." *Econometrica*, 65(3): 557–586.