INSTITUTE OF ECONOMICS

Scuola Superiore Sant'Anna

LEM | Laboratory of Economics and Management

Institute of Economics
Scuola Superiore Sant'Anna

Piazza Martiri della Libertà, 33 - 56127 Pisa, Italy
ph. +39 050 88.33.43
institute.economics@sssup.it

LEM

Working Paper Series

Reassessing the Cross-Sectional Fiscal Multiplier: Evidence from U.S. Defense Procurement, 1966–2019

Gianluca Pallante [a]

[a] Institute of Economics and L'EMbeDS, Scuola Superiore Sant'Anna, Pisa, Italy

2025/18 April 2025

ISSN(ONLINE): 2284-0400
DOI: 10.57838/sssa/3nns-1t43

Reassessing the Cross-Sectional Fiscal Multiplier: Evidence from U.S. Defense Procurement, 1966–2019

Gianluca Pallante *[1]

[1]Institute of Economics and l'EMbeDS, Sant'Anna School of Advanced Studies, Pisa, Italy

Abstract

This paper revisits the empirical analysis of Nakamura and Steinsson (2014). I reconstruct and extend the original dataset to cover the period 1966-2019, harmonizing two major sources of data: the Defense Contract Action Data System (DCADS) and USAspending.gov. I discuss how to aggregate these contract-level data to better capture spending more directly tied to domestic stimulus. Estimated multipliers are slightly lower in narrow replications but increase when incorporating later fiscal episodes. I also assess the validity and the stability of cross-sectional estimates. While some heterogeneity exists, dispersion in state-level responses remains within reasonable boundaries, especially when accounting for dynamic persistence.

Keywords: Fiscal Policy; Government expenditures; Cross-sectional multipliers; Defense Procurement
JEL classification: E62 · H57 · C18 · R12

*Corresponding author: gianluca.pallante@santannapisa.it - Institute of Economics, Scuola Superiore Sant'Anna, piazza Martiri della Libertà 33, 56127, Pisa (IT). Gianluca Pallante acknowledges support from the Italian Ministry of Research, PRIN PNRR 2022 project "NEWS" (Grant No J53D23016680001)

1 Introduction

The role of fiscal policy in macroeconomic stabilization has gained renewed prominence following the 2007-2009 financial crisis, particularly as monetary policy faced constraints at the zero lower bound (Ramey, 2011; Blanchard et al., 2012). A growing body of research has examined the effects of fiscal policy interventions (Ramey, 2019), highlighting their asymmetry in expansions and consolidations (Blanchard & Leigh, 2013; Barnichon et al., 2022), state-dependency (Auerbach & Gorodnichenko, 2012; Ferraresi et al., 2015), and variation across policy regimes (Ilzetzki et al., 2013).[1] Beyond its role in short-term stabilization, fiscal policy has also been instrumental in driving long-term economic transformations. Large-scale government spending — especially in defense and research and development (R&D)— has historically spurred industrial growth and technological advancements (Gross & Sampat, 2023; Kantor & Whalley, 2023; Pallante et al., 2023; Moretti et al., 2025). Empirical evidence further suggests that government spending can have persistent effects on innovation, productivity, and private-sector investment (Surico & Antolin-Diaz, 2022; Fieldhouse & Mertens, 2023; Ilzetzki, 2024), reinforcing its relevance for long-run economic performance.

Given the importance of fiscal interventions, measuring their impact remains a key empirical challenge. The fiscal multiplier — defined as the change in output per dollar of government spending — have been traditionally estimated using time-series approaches. While useful, these methods face limitations: large exogenous spending shocks are rare, and disentangling their effects from concurrent policy changes can be difficult (Nakamura & Steinsson, 2018). An alternative approach, increasingly used in empirical macroeconomics, leverages cross-sectional variation in government spending across smaller geographies. By comparing regions that experience different levels of spending shocks, cross-sectional studies can overcome some of the limitations of time-series models, providing also more detailed insights on transmission mechanisms and distributional issues (Chodorow-Reich, 2019).

A seminal contribution in this literature is Nakamura and Steinsson (2014, NS14, henceforth), which estimates local fiscal multipliers using variation in U.S. military procurement across states. Their identification strategy assumes that changes in aggregate military spending are exogenous to local economic conditions, while its allocation across states follows historical procurement patterns. This approach provides a quasi-experimental setting to estimate how regional economies respond to government demand shocks. Their findings suggest a multiplier of approximately 1.5, meaning that a $1 increase in defense spending raises state GDP by $1.5. Given its methodological rigor, NS14 has become a benchmark study, stimulating debates on

[1] The list is far from being exhaustive. For a more recent review of the major empirical contributions see Castelnuovo and Lim (2019).

whether or not there is "an applied micro free lunch for macroeconomists" (Ramey, 2019).

Indeed, studies suggest that cross-sectional multipliers may be sensitive to heterogeneous regional responses (Andrews, 2005; Pesaran, 2006; Almuzara & Sancibrián, 2024). If geographical units responds differently following policy shocks, standard cross-sectional methods may produce unstable or biased estimates, even when adopting a valid research design.

This paper extends and assesses the robustness of NS14's empirical analysis by reconstructing their dataset and expanding it to the period 1966–2019, including 13 more years, and thus capturing additional waves of military spending expansions and contractions and covering a period where the US monetary policy regime is constrained by the zero-lower bound. One contribution is the harmonization of military procurement data from two distinct sources: the Defense Contract Action Data System (DCADS), used in NS14, and USAspending.gov, which reports post-2006 US government expenditures, including those appropriated by the Department of Defense. To provide further reliability of cross-sectional estimates, a methodological contribution of this paper is to assess the stability of cross-sectional multiplier estimates using the diagnostics proposed by Canova (2024) and implemented in Canova and Pappa (2025). This procedure tests whether dispersion in the location-specific estimated multipliers is sufficiently low to validate the cross-sectional approach.

The findings confirm that fiscal multipliers remain positive, statistically different from zero and greater than one. However, they are lower when narrowly compared with NS14. Yet, they increase when the post-2006 period is included in the analysis. The state-level aggregation that matches the NS14 estimates the closest and most plausibly isolates deficit-financed government stimulus excludes contracts designed for foreign military sales and uniforms the minimum threshold reporting requirement across the whole sample. Finally, using Canova (2024)'s proposed test, results indicate that when accounting for dynamic persistence, the state-level cross-sectional multipliers are relatively stable and heterogeneity in the dynamic response is not statistically detected, validating the cross-sectional approach.

The paper is structured as follows. Section 2 details how the defense-related spending series is constructed and aggregated at state level. Section 3 recaps the empirical framework and the research design in NS14. Sections 4 and 5 replicate and test the stability of the cross-sectional multipliers, respectively. Section 6 concludes.

2 Data and aggregation filters

2.1 Data Description

The empirical analysis relies on two primary data sources to construct a measure of defense-related spending across U.S. states. The first is the Defense Contract Action Data System (DCADS). This is Department of Defense's data collection system for reporting contract actions to the Federal Procurement Data System (FPDS), which is the authoritative source of contract information for all US government agencies. DCADS reports prime contracts for goods and services between the private sector and agencies of the Department of Defense. In particular, each contract information relies on the DD 350 form, i.e., the Individual Contracting Action Report, including detailed information on the contracting office, recipient firm, place of performance, whether it is targeted for a foreign sale or not, and the total contract value.[2]

The second source of data is the USAspending.gov platform, the official open-data repository that provides more detailed records about contracts awarded by U.S. government expenditures, including contracts, grants, and financial assistance programs. I select information on prime contracts from the Department of Defense. This source provides richer information than DCADS but it can straightforwardly be matched along the many dimensions such as location, dates, contractor and contract characteristics.

Following NS14, the military spending dataset is complemented with additional macroeconomic variables, including measures of state output, sectoral output, population, inflation, employment, and military compensation. Table 1 provides a summary of these complementary data sources, specifying the relevant economic indicators and their sources. The final dataset covers 51 U.S. states, including Washington, D.C., and spans the period 1966–2019. The extended dataset builds upon NS14 by incorporating 13 additional years of military contract data.

Table 1: Sources of complementary data

Variable	Measure	Source
State Output	Annual GDP by State (SAGDP)	BEA Regional Economic Accounts
Sectoral State Output	Annual GDP by State by SIC(NAICS) Industry (SAGDPS-SAGDPN)	BEA Regional Economic Accounts
Inflation	Consumer Price Index: Total for United States (CPALTT01USA661S)	FRED
Population	Personal Income, Population (SAINC51)	BEA Regional Data
DoD wages	Compensation of Employees by industry: Military compensation (SAGDP4N)	BEA Regional Data
Employment	Current Employment Statistic (SM)	BLS

[2] The DCADS archival records can be found in two different series: "Records of Prime Contracts Awarded by the Military Services and Agencies", (ARC Identifier 606901) for the period 1965-1975 and available at https://catalog.archives.gov/id/606901; for fiscal years 1976 - 2006, instead I use the "Records of Prime Contracts Awarded by the Military Services and Agencies" (ARC Identifier 578589), available at https://catalog.archives.gov/id/578589

2.2 State-level aggregation

After collecting and merging the contract-level data from DCADS and USAspending.gov, the dataset comprises approximately 41.1 million individual contract observations. Since NS14 estimate state-level fiscal multipliers, the contract-level data must be aggregated at the state-year level to construct a comparable measure of defense spending. I first perform the easiest form of aggregation by summing all observations grouped by the 51 state names and year. I label this dataset as "RAW", as it provides the unadjusted measure of military spending. Given the characteristics of the two data sources assembled, to ascertain the role played by measurement errors, I apply three different data filters.

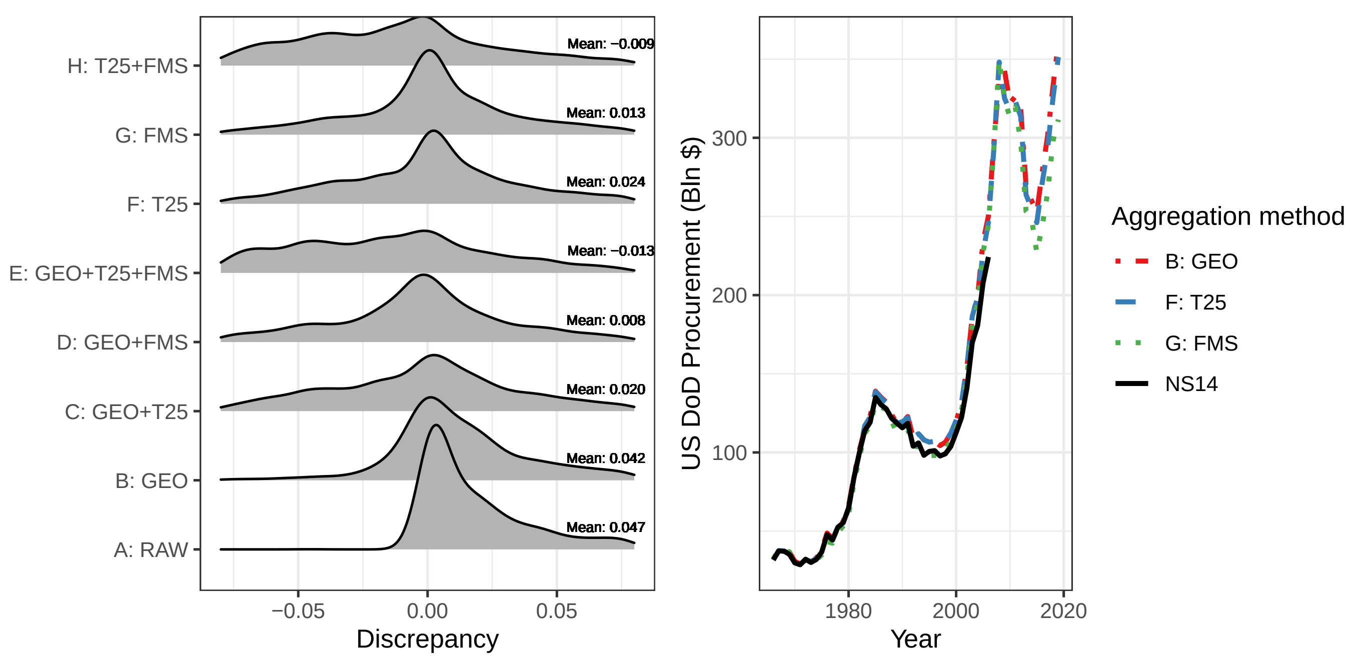
[Threshold reporting requirements (T25)]. From 1966 to 1983, the records include contracts with a value of $10,000 or larger. From fiscal year FY 1983 onwards, the records include contracts with a value of $25,000 or more. However, contracts from USAspending.gov are characterized by transactions lower than the corresponding reporting requirements. Many of these transactions include but are not limited to revisions (see also Auerbach et al., 2020, for a discussion). For this reason, all contracts that reports amounts less than $25,0000 in absolute value are removed. Despite the authors acknowledge changes in the threshold for reporting, I document that excluding from the analysis contracts exceeding the 25k threshold may avoid spurious correlation due to possible concentration of small or big contracts in particular States.

[Geographical adjustment (GEO)]. The dataset shows inconsistencies in state name coding, arising from misspellings, incorrect abbreviations, or missing entries. To assign a geographical reference to these observations, I manually retrieve the correct state name based on county and place of performances names or identifiers. Moreover, the USAspending.gov dataset contains many contracts awarded to the same contractor, performed in the same place, with deobligated amounts very close (or identical) to the original (firstly-awarded) one. Similar to Demyanyk et al. (2019), both contracts are removed from the dataset whenever the difference between the original contract and the corresponding deobligation is lower than the 0.5% of deobligated amount. This filter eliminates about 4.4% of the observations (very close to Demyanyk et al., 2019), which are concentrated in the post 2008 period. To ensure consistency, I apply the same filter to the DCADS records, where none of such instances occur. On this regard, whenever I observe deobligations, they correspond exactly to the original amount.

[Foreign funding for military acquisitions (FMS)]. There are cases in which a foreign government, international organization, or foreign military organizations bear some of the cost of the acquisition. The Foreign Military Sale (FMS) is a U.S. program for transferring military equipment or services to international partners. Since FMS contracts are not financed through domestic U.S. government spending, they do not represent a true fiscal stimulus within the U.S. economy. Including FMS transactions in the dataset would therefore introduce measurement error in the estimation of fiscal multipliers. Consequently, all FMS-related transactions are excluded, ensuring that the state-level aggregate of military contracts reflects only domestically financed military spending. In some phases of the US foreign policy, contracts under FMS were not negligible and therefore I suggest that removing this type of spending may help avoid error in the measurement of a deficit-financed public spending variable.

Figure 1 shows the discrepancies when accounting for different methods of data filtering and adjustments. I acknowledge the fact that none of the filters exactly mirror the data published along NS14. Filter A (No adjustments in the geographical imputations, i.e, RAW) deviates quite a lot from the dataset I aim to replicate, with a rightly skewed discrepancy that averages 4.7%. On the contrary, filter D (GEO and FMS adjustments) matches NS14 data the closest, with a discrepancy that is equally distributed around zero. Thus, in order to achieve the closest replicability for data transformation, the researcher should exclude contracts stipulated for foreign sales and should not neglect inconsistencies in the geographical accounting of the entries. To provide more context, the battery of estimations that I perform clarify whether these factors affect or not substantially the size of the fiscal multiplier.

Figure 1: Differences with original data from NS14 and with different state-level aggregation filters.



Notes. Left panel: densities plots of percentage difference between state-year observations in NS14 and using the data filtering procedures described in Section 2.2. The panel also reports the sample average of the discrepancy. Right panel: US DoD procurement series in NS14 and for selected filtering methods.

3 Model & Research design

To estimate the fiscal multiplier, I adopt the NS14 regression specification, which is also widely used in the estimation of local fiscal multipliers (Auerbach et al., 2020; Muratori et al., 2023; Auerbach et al., 2024):

$$\frac{Y_{i,t} - Y_{i,t-2}}{Y_{i,t-2}} = \beta \frac{G_{i,t} - G_{i,t-2}}{Y_{i,t-2}} + \alpha_i + \gamma_t + \varepsilon_{i,t} \tag{1}$$

Closely following NS14, $Y_{i,t}$ measures the real per capita output (or employment) in state i in year t , while $G_{i,t}$ captures the real per capita value of public spending, measured by Department of Defense (DoD) procurement contracts. The specification includes state fixed effects α_i to account for state-specific trends in output and military spending, as well as year fixed effects γ_t to control for aggregate economic fluctuations and policies carried at the national level. The coefficient of interest, β, represents the fiscal multiplier, which quantifies the dollar value of output generated by a one-dollar increase in government expenditures. The model follows the two-year difference approach implemented by NS14, which reduces short-term volatility while preserving sufficient variation in the spending variable. Standard errors are clustered at the state level to account for serial correlation induced by the overlapping observation windows.

A central challenge in estimating the fiscal multiplier is the endogeneity of government spending, as defense procurement may be influenced by state-specific economic conditions. Political and strategic factors often play a role in allocating defense contracts, either by directing spending to struggling regions as an implicit form of economic support or by prioritizing politically influential states. If unobserved economic factors influence both government spending and economic performance, OLS estimates of β will be biased.

To address this concern, NS14 adopt an instrumental variable (IV) strategy, leveraging the fact that state-level military spending responds systematically in a different way to fluctuations in the total U.S. defense budget. The identification assumption is that while total military spending is driven by national security considerations and geopolitical events, its state-level allocation also follows historical procurement patterns. This allows the isolation of the systematic response of state-level expenditures to total U.S. military spending.

The two-stage least square interpretation of the estimated coefficient is that only systematic state-level variation in military spending is predicted. The first-stage equation is $\Delta G_{i,t} = \theta_i \Delta G_t + \delta_i + \lambda_t + \nu_{i,t}$, where G_t represents total U.S. military spending, and θ_i captures state-specific exposure to national defense appropriations. This exposure coefficient reflects the systematic tendency of state i to receive military contracts when the federal defense budget changes. Throughout the text, only 2SLS regression results are reported.

4 Results

Do the filters that more closely match the empirical distribution of the original NS14 data deliver the same estimates of the fiscal multiplier? Figure 2 presents a summary of the multiplier β, estimated using Equation (1), under different data aggregation approaches.[3] The figure contrasts the narrowest replication (dashed lines), retrieving the original data from NS14 and covering the 1966–2006 period, with estimates obtained incorporating data up to 2019.[4]

To assess the influence of measurement error in complementary variables, which may be due to revisions, I report estimates (dotted lines) based on a dataset assembled using the original NS14 spending series and retrieving all additional variables from their respective primary sources, as documented in Table 1. These estimates serve as a benchmark against those obtained using the replicated, harmonized, and extended dataset, which varies depending on the filtering applied to the newly constructed spending series. As in NS14, all output multipliers are reported both including and excluding military compensation from the spending series (in red circles and blue triangles, respectively).

Across all specifications, the estimates are very close to those reported in NS14, systematically lower, yet greater than one. Once the sample is extended to 2019 and incorporates the harmonized procurement series, the estimated multipliers increase and get closer to the original NS14 benchmark (right-hand panels of Figure 2). This finding is somewhat consistent with the interpretation that the effects of government spending are more prominent during periods of economic slack or when monetary policy hits the zero lower bound, an economic regime effectively covered by the post-2006 extended sample.

Interestingly, the aggregation method that minimizes the average discrepancy with NS14's data—filter D (GEO+FMS)—does not necessarily yield the closest point estimates of the fiscal multiplier. Nonetheless, excluding contracts flagged as foreign military sales (FMS) appears to have the most pronounced effect in raising the estimated multiplier. This effect is followed, in magnitude, by the implementation of a $25,000 threshold on contract reporting (T25). Improving the geographical allocation of contracts (GEO) seem to exert a downward pressure to the estimates, revealing the presence of a mild attenuation bias. Among these, the FMS filter is of particular interest, as it isolates procurement more directly attributable to domestic demand stimulus and likely financed through deficit spending. Given its theoretical and empirical implications, results

[3]For clarity, I limit the focus of the paper's findings to a subset of all NS14's results, which include a series of robustness checks such as the use of a classic Bartik instrument, the estimation of the fiscal multipliers under alternative specifications and in periods of high versus low unemployment. However, I report the full set of replication results—including all alternative specifications and filter combinations—in the Supplementary Material (link) available online.

[4]To broad the scope of the replication study, all estimations are performed using the R software, with the `fixest` package available at https://cran.r-project.org/web/packages/fixest/index.html

suggest that this specification is the preferred one.

The estimated multipliers also replicate one of the key empirical regularities reported in NS14: the size of the fiscal multiplier increases with the geographic scope of aggregation. That is, estimates computed at the Census region level are consistently larger than those at the state level. Additionally, excluding military compensation from the spending series reduces the estimated multiplier at the state level but increases it at the regional level, a pattern also observed in NS14. Figure 3 displays the estimated multipliers for employment. These estimates are generally lower than those for output and also lower than the employment multipliers reported in NS14. As for output, dropping contracts that are designated for foreign sales (FMS) is important for matching the NS14 estimates. Somewhat differently, employment responses to fiscal shocks are less sensitive to variation in the data aggregation methods proposed.

Finally, Table 4 presents the same exercise for state-level sectoral output multipliers.[5] Sector-level responses reveal notable heterogeneity. Despite the composition of military procurement has evolved over time, with services and transportation accounting for a growing share of contracts (Muratori et al., 2023), similarly to NS14, multipliers remain statistically different from zero in the construction, manufacturing, retail and services sectors. Multipliers do not seem to be very much sensitive to the data filtering proposed and, among those that are statistically significant, they tend to be lower when the database is extended up to 2019.

5 Testing dynamic heterogeneity

Recent methodological contributions have raised concerns about the reliability of cross-sectional macroeconomic estimates, particularly in settings where policy interventions induce heterogeneous dynamic responses across units. If regions or states respond differently to the same policy shock, cross-sectional estimation methods may fail to produce stable and reliable multiplier estimates.

To evaluate the robustness of the U.S. defense-related fiscal multiplier, I apply the diagnostic test proposed by Canova (2024), that relies on the coefficient of variation (CV) of the distribution of region-specific estimated multipliers. The underlying rationale is that if responses to policy shocks are homogeneous across states, the CV should be small, indicating that estimates are stable. Conversely, if responses are highly heterogeneous, the CV will be large, signalling that cross-sectional methods may not provide an accurate representation of the underlying economic effects. The diagnostics relies on computing the distribution of *state-by-state* estimated

[5]Data are available following the SIC classification before 1997 and NAICS after 1997. Since data for 1997 are available under the two classifications, growth rate series can be smoothly pasted together. For better readability, Figure 4 shows state-level estimates only. The macro-regional counterpart are available in the online Supplementary Material (link).

Figure 2: Revisited estimates of cross-sectional fiscal spending output multipliers

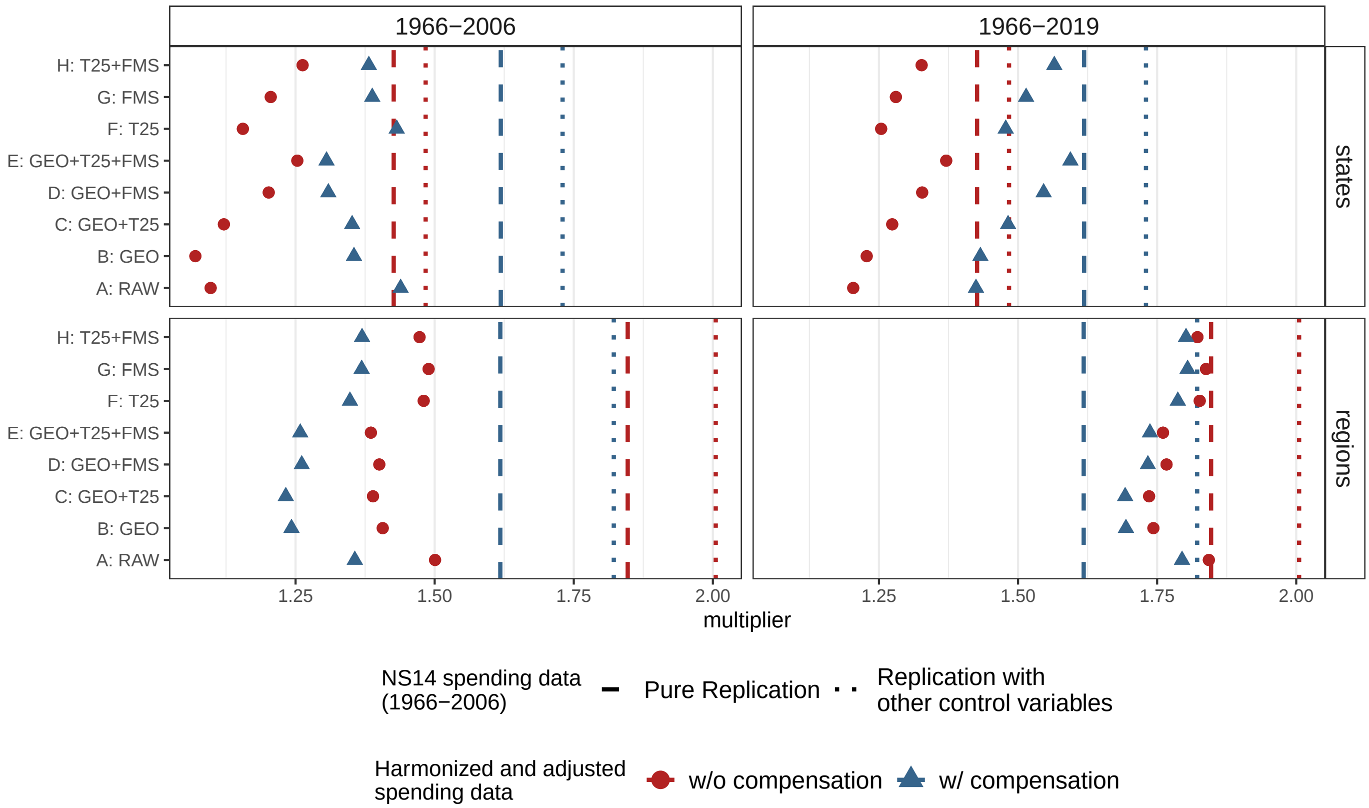


Figure 3: Revisited estimates of cross-sectional fiscal spending employment multipliers

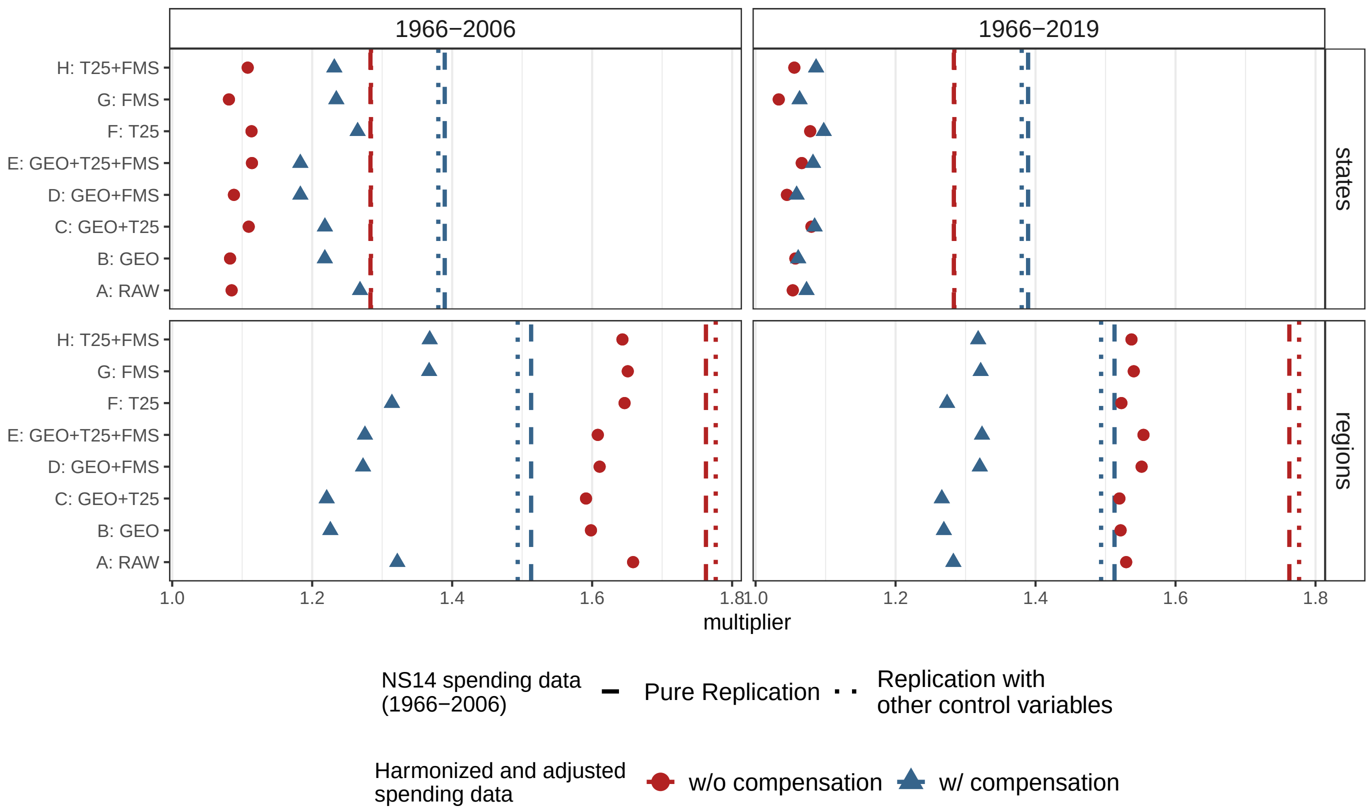
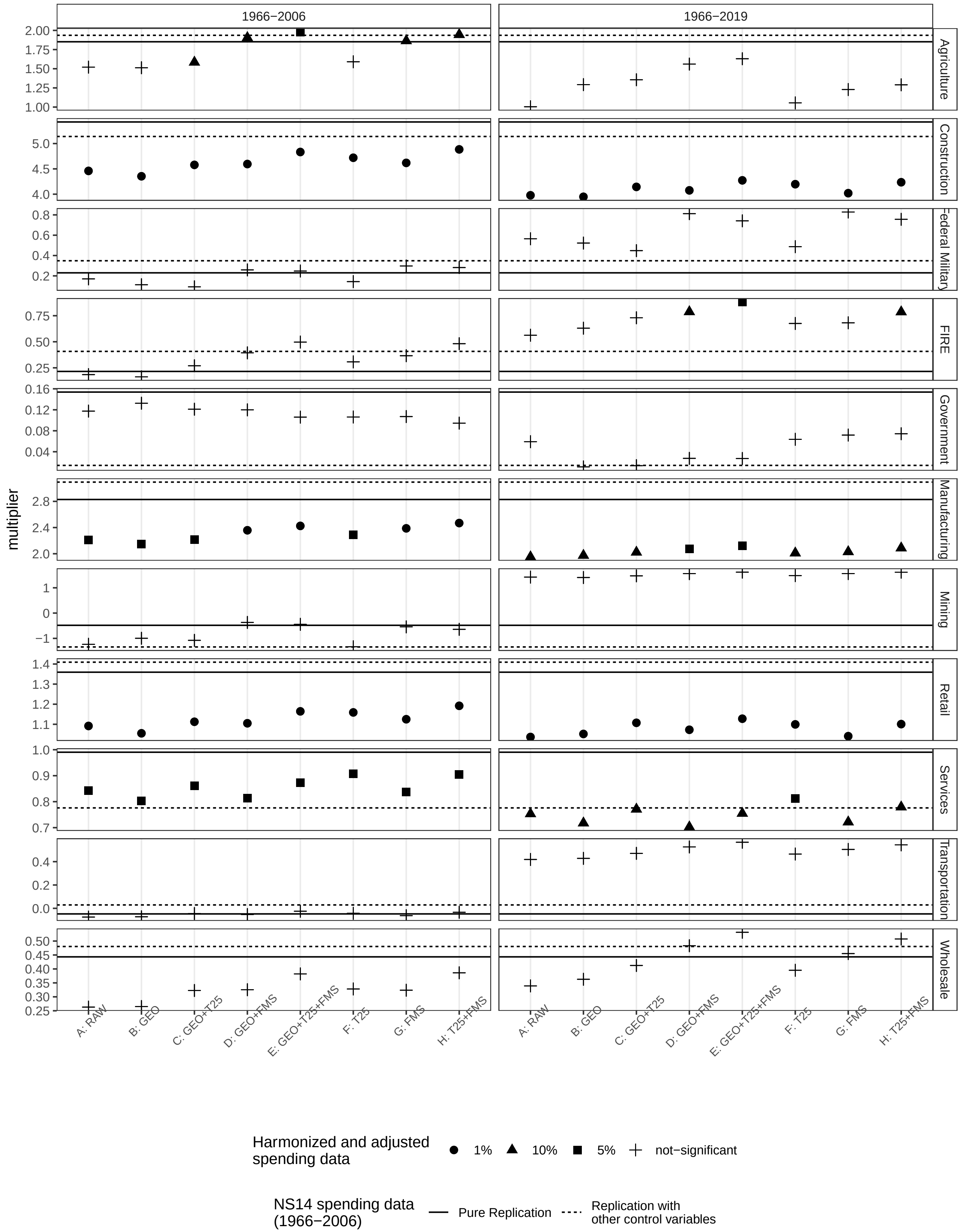


Figure 4: Revisited estimates of cross-sectional (state-level) fiscal spending sectoral output multipliers



multipliers, and compare the CV with bootstrap-generated critical values used to assess the null hypothesis of homogeneous regional response. The degree to which estimates vary across states serves as an indicator of the extent of heterogeneity in fiscal policy responses.

Since the number of regional aggregates (10) is too small to conduct a meaningful CV test, I restrict the analysis to the state-level dataset, which consists of 51 states over 53 years. I adopt two M specifications commonly used in the estimation of fiscal multipliers. The first specification (M = NS) estimates the policy effect using the standard NS14 approach, while the second specification (M = LagDep) introduces lagged dependent variables to account for dynamic persistence. In both specifications, the policy variable enters in first differences and the *state-by-state* estimate of the two-year multiplier is $\beta_{i,M}$.[6] More formally:

$$\text{(NS)} \quad \frac{Y_{i,t} - Y_{i,t-2}}{Y_{i,t-2}} = \beta_{i,NS} \frac{G_{i,t} - G_{i,t-2}}{Y_{i,t-2}} + \alpha_i + \varepsilon_{i,t} \tag{2}$$

$$\text{(LagDep)} \quad \frac{Y_{i,t} - Y_{i,t-2}}{Y_{i,t2}} = \beta_{i,LagDep} \frac{G_{i,t} - G_{i,t-2}}{Y_{i,t-2}} + \delta_1 \frac{Y_{i,t-1} - Y_{i,t-3}}{Y_{i,t-3}} + \delta_2 \frac{G_{i,t-1} - G_{i,t-3}}{Y_{i,t-3}} + \alpha_i + \varepsilon_{i,t} \tag{3}$$

Table 2 reports the results of the CV analysis. The findings indicate that the NS specification exhibits greater dispersion, with higher CV values compared to the LagDep specification. This suggests that the assumption of homogeneous responses across states is less defensible when lagged economic conditions are not controlled for explicitly. The corresponding critical values for the null of homogeneity are 1.425 (90%), 1.576 (95%), and 1.903 (99%) for the two-year-ahead multiplier. The estimated CV values for the LagDep specification remain lower than the reported critical values, suggesting that controlling for the persistence in economic condition is important for taming the heterogeneity in state-level responses. While some degree of variation across states is observed, the results indicate that state-level heterogeneity does not fundamentally undermine the cross-sectional approach. However, the differences between the NS and LagDep specifications suggest that modeling dynamic persistence explicitly can reduce dispersion in multiplier estimates, making the latter specification preferable.

6 Conclusion

This paper revisits the cross-sectional estimates of the US defense-related fiscal multiplier, retrieving and extending the empirical exercise of Nakamura and Steinsson (2014, NS14). Data from the Defense Contract Action Data System (DCADS) and USAspending.gov have been harmonized, so to build a consistent series of

[6] To make the two approaches more easily comparable, only the contemporaneous two-year difference in $G_{i,t}$ is instrumented as detailed in Section 3.

Table 2: Coefficient of variation diagnostics for heterogenous dynamic responses

	NS			LagDep				NS			LagDep		
	β	$\overline{\beta}$	$\lvert IQR \rvert$	β	$\overline{\beta}$	$\lvert IQR \rvert$		β	$\overline{\beta}$	$\lvert IQR \rvert$	β	$\overline{\beta}$	$\lvert IQR \rvert$
A: RAW	1.10	-1.04	2.63	1.37	2.95	0.90	E: GEO+T25+FMS	1.25	-0.95	4.83	1.49	3.30	0.94
B: GEO	1.07	-0.83	2.77	1.40	3.20	0.77	F: T25	1.16	-1.08	2.75	1.45	3.09	0.91
C: GEO+T25	1.12	-0.89	3.00	1.48	3.34	0.84	G: FMS	1.21	-1.05	3.41	1.37	3.00	1.02
D: GEO+FMS	1.20	-0.89	3.78	1.42	3.16	0.94	H: T25+FMS	1.26	-1.12	3.91	1.45	3.12	1.00

|IQR| Critical Values		(2-year ahead multiplier)	
	90% = 1.425	0.95% = 1.576	0.99% = 1.903

The table report, for each data filter and model specification as in Equations 2 and 3, the cross-sectional aggregate estimate β, the average multiplier computed across state-by-state regressions $\overline{\beta}$, and the coefficient of variation defined as the absolute value of the interquartile range of the estimated distribution of multipliers divided by the midpoint of the interquartile range, as in Canova (2024). The critical values for testing the null hypothesis of homogeneity in the dynamic response are taken from Canova (2024), Table 3.

defense-related expenditures across U.S. states from 1966 to 2019.

The analysis shows that, when using this detailed source of data, the estimation of the cross-sectional multiplier is influenced by changes in minimum threshold for procurement reporting, geographical assignment of the performance and whether the signed contracts are designated to foreign military sales. In particular, researchers are advised to exclude contracts with a foreign sale destination as this adjustment brings the empirical estimates closer to the NS14 benchmark, but it also better captures spending more directly attributable to a domestic, deficit-financed, demand stimulus.

Indeed, the results indicate that the fiscal multipliers estimated using the harmonized and extended dataset remain broadly consistent with those originally reported by NS14. While the narrow replication yields slightly lower estimates, extending the dataset to include post-2006 observations leads to larger multipliers—consistent with the literature emphasizing stronger fiscal effects during periods of economic slack and where monetary policy is also stuck at the zero lower bound.

Furthermore, challenged by recent methodological contributions suggesting the conditions under which cross-sectional macro elasticities are unstable in presence of heterogeneous dynamic response, I test the stability of the NS14 empirical framework, using a coefficient of variation (CV) diagnostic, recently proposed by Canova (2024). The analysis shows that, although some heterogeneity exists across states, the dispersion of estimates generally remains within the critical bounds. The inclusion of lagged dependent variables reduces variability and enhances the stability of the estimates, reinforcing the importance of accounting for dynamic persistence. These findings confirm that cross-sectional methods, when properly implemented and tested, continue to provide a valid empirical framework for evaluating the macroeconomic effects of fiscal interventions.

References

Andrews, D. W. K. (2005). Cross-Section Regression with Common Shocks. *Econometrica*, *73*(5), 1551–1585. https://doi.org/10.1111/j.1468-0262.2005.00629.x

Pesaran, M. H. (2006). Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure. *Econometrica*, *74*(4), 967–1012. https://doi.org/10.1111/j.1468-0262.2006.00692.x

Ramey, V. A. (2011). Can Government Purchases Stimulate the Economy? *Journal of Economic Literature*, *49*(3), 673–685. https://doi.org/10.1257/jel.49.3.673

Auerbach, A. J., & Gorodnichenko, Y. (2012). Measuring the Output Responses to Fiscal Policy. *American Economic Journal: Economic Policy*, *4*(2), 1–27. https://doi.org/10.1257/pol.4.2.1

Blanchard, O. J., Romer, D., Spence, A. M., & Stiglitz, J. E. (Eds.). (2012). *In the Wake of the Crisis: Leading Economists Reassess Economic Policy*. The MIT Press. Retrieved March 4, 2025, from https://www.jstor.org/stable/j.ctt5vjq9h

Blanchard, O. J., & Leigh, D. (2013). Growth Forecast Errors and Fiscal Multipliers. *American Economic Review*, *103*(3), 117–120. https://doi.org/10.1257/aer.103.3.117

Ilzetzki, E., Mendoza, E. G., & Végh, C. A. (2013). How big (small?) are fiscal multipliers? *Journal of Monetary Economics*, *60*(2), 239–254. https://doi.org/10.1016/j.jmoneco.2012.10.011

Nakamura, E., & Steinsson, J. (2014). Fiscal Stimulus in a Monetary Union: Evidence from US Regions. *American Economic Review*, *104*(3), 753–792. https://doi.org/10.1257/aer.104.3.753

Ferraresi, T., Roventini, A., & Fagiolo, G. (2015). Fiscal Policies and Credit Regimes: A TVAR Approach. *Journal of Applied Econometrics*, *30*(7), 1047–1072. https://doi.org/10.1002/jae.2420

Nakamura, E., & Steinsson, J. (2018). Identification in Macroeconomics. *Journal of Economic Perspectives*, *32*(3), 59–86. https://doi.org/10.1257/jep.32.3.59

Castelnuovo, E., & Lim, G. (2019). What Do We Know About the Macroeconomic Effects of Fiscal Policy? A Brief Survey of the Literature on Fiscal Multipliers. *Australian Economic Review*, *52*(1), 78–93. https://doi.org/10.1111/1467-8462.12313

Chodorow-Reich, G. (2019). Geographic Cross-Sectional Fiscal Spending Multipliers: What Have We Learned? *American Economic Journal: Economic Policy*, *11*(2), 1–34. https://doi.org/10.1257/pol.20160465

Demyanyk, Y., Loutskina, E., & Murphy, D. (2019). Fiscal Stimulus and Consumer Debt. *The Review of Economics and Statistics*, *101*(4), 728–741. https://doi.org/10.1162/rest_a_00796

Ramey, V. A. (2019). Ten Years after the Financial Crisis: What Have We Learned from the Renaissance in Fiscal Research? *Journal of Economic Perspectives*, *33*(2), 89–114. https://doi.org/10.1257/jep.33.2.89

Auerbach, A., Gorodnichenko, Y., & Murphy, D. (2020). Local Fiscal Multipliers and Fiscal Spillovers in the USA. *IMF Economic Review*, *68*(1), 195–229. https://doi.org/10.1057/s41308-019-00102-3

Barnichon, R., Debortoli, D., & Matthes, C. (2022). Understanding the Size of the Government Spending Multiplier: It's in the Sign. *The Review of Economic Studies*, *89*(1), 87–117. https://doi.org/10.1093/restud/rdab029

Surico, P., & Antolin-Diaz, J. (2022, July). The Long-Run Effects of Government Spending. Retrieved June 5, 2024, from https://cepr.org/publications/dp17433

Fieldhouse, A., & Mertens, K. (2023, May). The Returns to Government R&D: Evidence from U.S. Appropriations Shocks. https://doi.org/10.24149/wp2305r1

Gross, D. P., & Sampat, B. N. (2023). America, Jump-Started: World War II R&D and the Takeoff of the US Innovation System. *American Economic Review*, *113*(12), 3323–3356. https://doi.org/10.1257/aer.20221365

Kantor, S., & Whalley, A. (2023, July). *Moonshot: Public R&D and Growth* (tech. rep. No. w31471). National Bureau of Economic Research. Cambridge, MA. https://doi.org/10.3386/w31471

Muratori, U., Juarros, P., & Valderrama, D. (2023). Heterogeneous Spending, Heterogeneous Multipliers. *IMF Working Papers*, *2023*(052). https://doi.org/10.5089/9798400233777.001.A001

Pallante, G., Russo, E., & Roventini, A. (2023). Does public R&D funding crowd-in private R&D investment? Evidence from military R&D expenditures for US states. *Research Policy*, *52*(8), 104807. https://doi.org/10.1016/j.respol.2023.104807

Almuzara, M., & Sancibrián, V. (2024). *Micro Responses to Macro Shocks* (Staff Reports - Federal Reserve Bank of New York). Federal Reserve Bank of New York. Retrieved March 4, 2025, from https://www.newyorkfed.org/medialibrary/media/research/staff_reports/sr1090.pdf

Auerbach, A. J., Gorodnichenko, Y., & Murphy, D. (2024). Macroeconomic Frameworks: Reconciling Evidence and Model Predictions from Demand Shocks. *American Economic Journal: Macroeconomics*, *16*(3), 190–229. https://doi.org/10.1257/mac.20210033

Canova, F. (2024). Should we trust cross-sectional multiplier estimates? *Journal of Applied Econometrics*, *39*(4), 589–606. https://doi.org/10.1002/jae.3041

Ilzetzki, E. (2024). Learning by Necessity: Government Demand, Capacity Constraints, and Productivity Growth. *American Economic Review*, *114*(8), 2436–2471. https://doi.org/10.1257/aer.20230033

Canova, F., & Pappa, E. (2025). The Macroeconomic Effects of EU Regional Structural Funds. *Journal of the European Economic Association*, *23*(1), 327–360. https://doi.org/10.1093/jeea/jvae031

Moretti, E., Steinwender, C., & Van Reenen, J. (2025). The Intellectual Spoils of War? Defense R&D, Productivity, and International Spillovers. *The Review of Economics and Statistics*, *107*(1), 14–27. https://doi.org/10.1162/rest_a_01293