When Does Government Spending Matter? It's All in the Measurement

Online Appendix

Edoardo Briganti   Gillian Brunet   Victor Sellemi

July 3, 2025

Replication Instructions. The results of the paper are derived following the replication package guidelines suggested by Professor Ingar Haaland in his post on X.

- All results are replicated by running with one click the `master_dofile.do`, which runs multiple do-files stored in a `code` directory.
- All raw data are stored in a `raw_data` directory which is used to create the datasets used in the analysis, stored in a `data` directory.
- All output is stored in an `output` directory organized in a figures and tables directory.
- Tables' contents is automatically uploaded from `Stata` to `Overleaf`, also following his code.

Contents

A Constructing Defense Contract Data

Here we describe the data sources of defense contracts and the construction of the defense contracts series, available going back to 1940Q1.

A.1 Business Conditions Digest

The first data source comes from contract data used in Ramey (1989), which is originally from the periodical Business Conditions Digest, or "*BCD*".[1] BCD contains monthly data of prime military contracts from January 1951 until November 1988. The series was then discontinued and migrated to the Survey of Current Business (SCB). However, data from SCB is only available from January 1990 until September 1995, with systematic omission of values from Q4 of each year. Given these missing data and reliability concerns, we choose not to use SCB data.

Figure A1 plots real BCD contract awards data, both in levels (left panel) and as a share of GDP (right panel), at the quarterly frequency (aggregated from monthly and annualized by calendar year).The vertical line indicates 1988Q4, when BCD data was migrated to SCB.

A.2 Federal Procurement Summary Report

The second data source originates from the annual Official Federal Procurement Summary Report ("*FPSR*"), produced by the Directorate for Information Operations and Reports (DIOR). FPSR contains data on both annual and quarterly federal procurement contracts, and spans the inception of the Federal Procurement Data System (FPDS) in Q1 1981 to Q3 2003.

The annual reports provide both (i) the value of military prime contract awards by fiscal year, and (ii) bar charts illustrating the quarterly values of *total* federal procurement contracts, covering contracts in both federal defense and non-defense sectors.

[1] Military contracts became part of the set of instruments known as the Hall-Ramey Instruments (from Hall (1990) and Ramey (1989)).

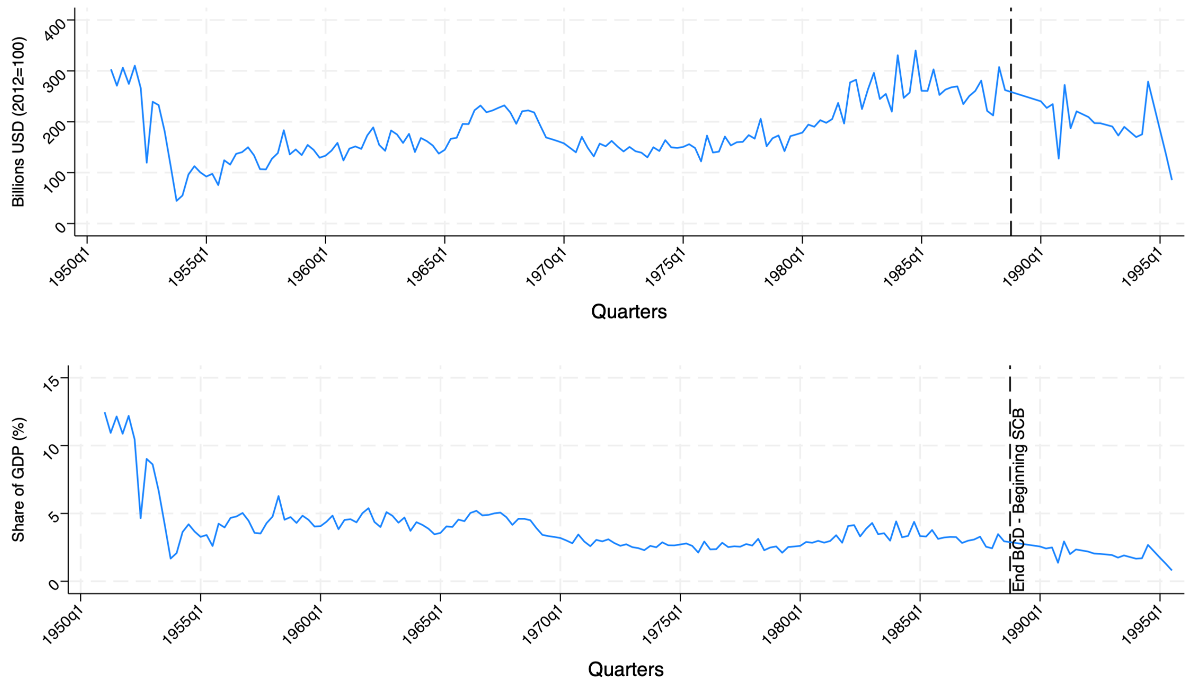


Figure A1: Military Prime Contracts Data from the Business Conditions Digest

Notes: Monthly data is aggregated at quarterly frequency and then annualized by calendar year. Price deflator is the GDP price deflator available from NIPA.

The left panel of Figure A2 compares the annual (fiscal year) values of military contracts with the quarterly values of federal contracts, aggregated to the fiscal year frequency. Federal contracts include both military and non-military procurement. The quarterly federal values, when aggregated by fiscal year, show a strong correlation with the annual military values. Notably, approximately 80% of federal procurement during these years is attributed to military procurement. This is illustrated in the right panel of the same figure, which depicts the share of military contracts as a percentage of total federal contracts by fiscal year.

To interpolate the annual military contract data, we use information from the quarterly values of total federal contracts in a two-step process.

First, we seasonally adjust the series of quarterly (annualized) real federal contracts using the X-13 ARIMA-SEATS program from the Census Bureau.[2] The top-left panel of Figure A3 displays

[2]The program is available for download here.

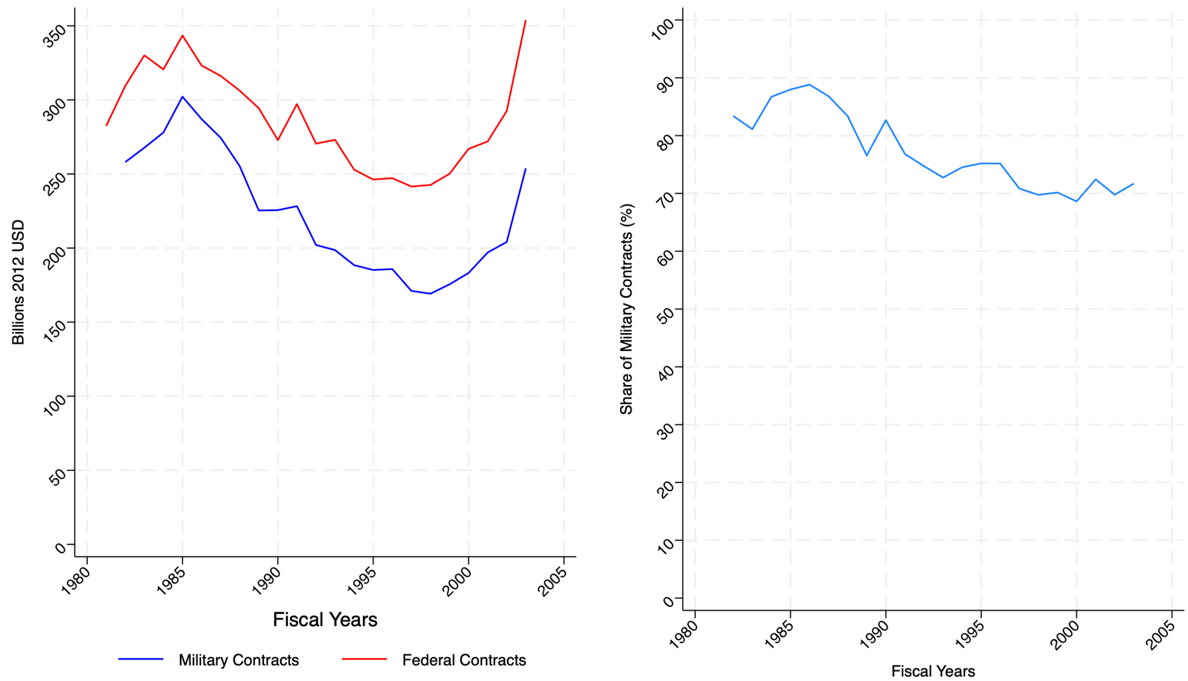


Figure A2: Federal Procurement Summary Report Annual Data - Federal vs. Military Contracts

Notes: Quarterly data on federal (i.e., military plus non-military) procurement is aggregated to the fiscal year frequency.

the original quarterly non-seasonally-adjusted series of all federal procurement contracts alongside its seasonally adjusted version.

Second, we interpolate the fiscal-year series of military procurement contracts by ensuring that the values of the nominal quarterly annualized and seasonally adjusted series of all federal procurement contracts within a fiscal year average out to the original values of military procurement contracts for that fiscal year, as reported in the FPSR. This involves adjusting the values of quarterly contracts to account for the gap between their fiscal year average and the reported military contracts for that fiscal year. The top-right panel of Figure A3 compares the quarterly series before (red line) and after (blue line) the fiscal year adjustment. Since military contracts account for approximately 80% of federal contracts, the fiscal year adjustment is generally small, representing only a slight downward shift in the quarterly series.

Finally, the bottom-left panel of Figure A3 shows the new quarterly series of military contracts

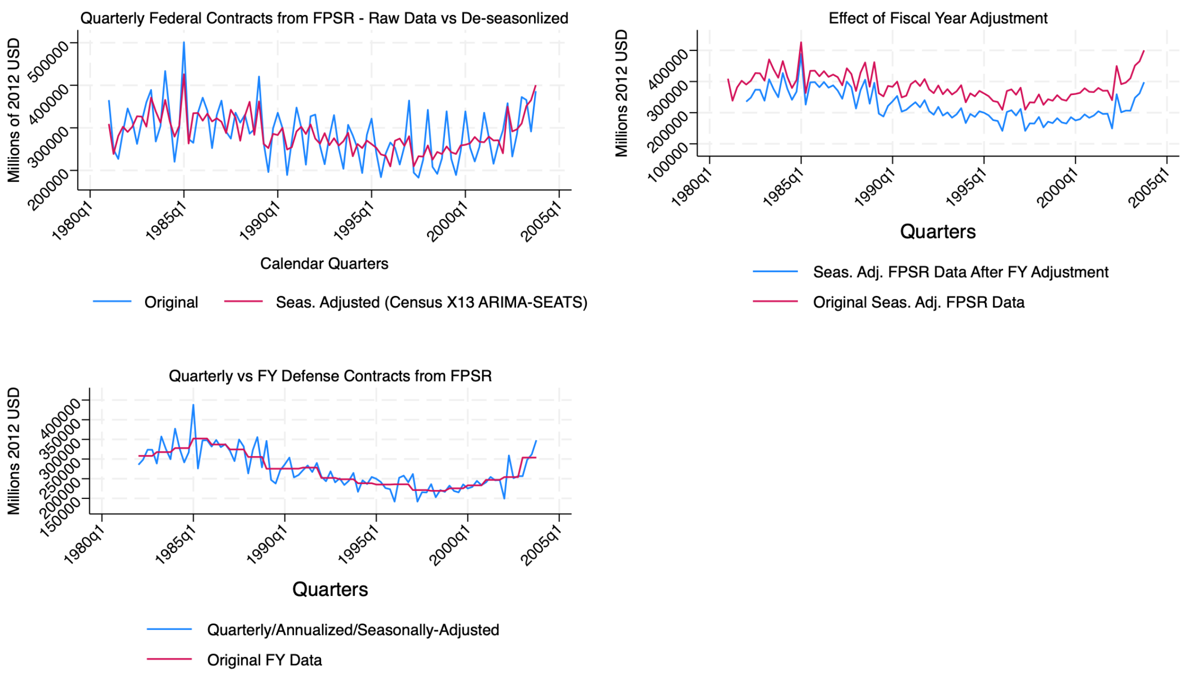


Figure A3: Federal Procurement Summary Report: Seasonal and Fiscal Year Adjustments

Notes: Price deflator is the GDP price deflator. Series are annualized.

in blue, alongside their fiscal year averages, which coincide by construction with the original military contract values from the FPSR.

In conclusion, the top panel of Figure A4 shows the real value of the quarterly series of military procurement contracts, while the bottom panel presents these quarterly values as a share of nominal GDP.

A.3 Federal Procurement Data System (FPDS-NG)

After the fourth quarter of 2000, all daily federal procurement transactions are observed from the Next Generation of the Federal Procurement Data System (or "*FPDS-NG*").

We identify military contracts by focusing on those awarded by the Department of Defense. These contracts are identified in FPDS-NG by observing a value of 97 in the field `awarding_agency_code`, which represents approximately 56% of all federal procurement transactions.

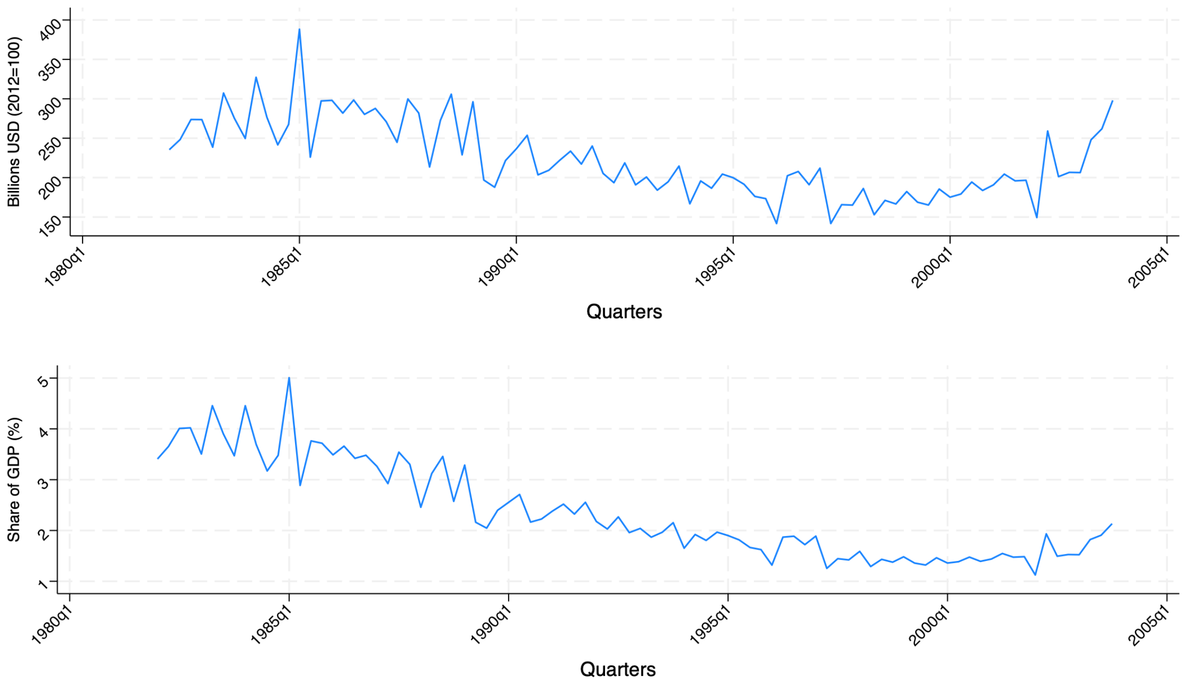


Figure A4: Federal Procurement Summary Report Time Series of Contracts

Notes: Price deflator is the GDP price deflator. Quarterly data is annualized.

We aggregate military contracts by fiscal year and present them in Figure A5. Notice the initial increase in contracts, reflecting the military buildup following the terrorist attacks on September 11, 2001. A significant spike in contracts occurs in 2010, followed by a decline after the enactment of President Obama's expenditure-based fiscal adjustment, the Budget Control Act of 2011. This downward trend in contracts reverses following the election of President Trump, who increased military spending.

We interpolate the series using quarterly variation in newly awarded military contracts. This approach is preferred over aggregating all contracts at a quarterly frequency because large de-obligations can occur in the quarters following the initial positive obligation, which they aim to cancel. The existence of large de-obligations in subsequent quarters increases the noise in the series. Aggregation by fiscal year helps balance out de-obligated contracts which have no real economic effect.

The interpolation procedure mirrors that used for the FPSR data. First, we deseasonalize the

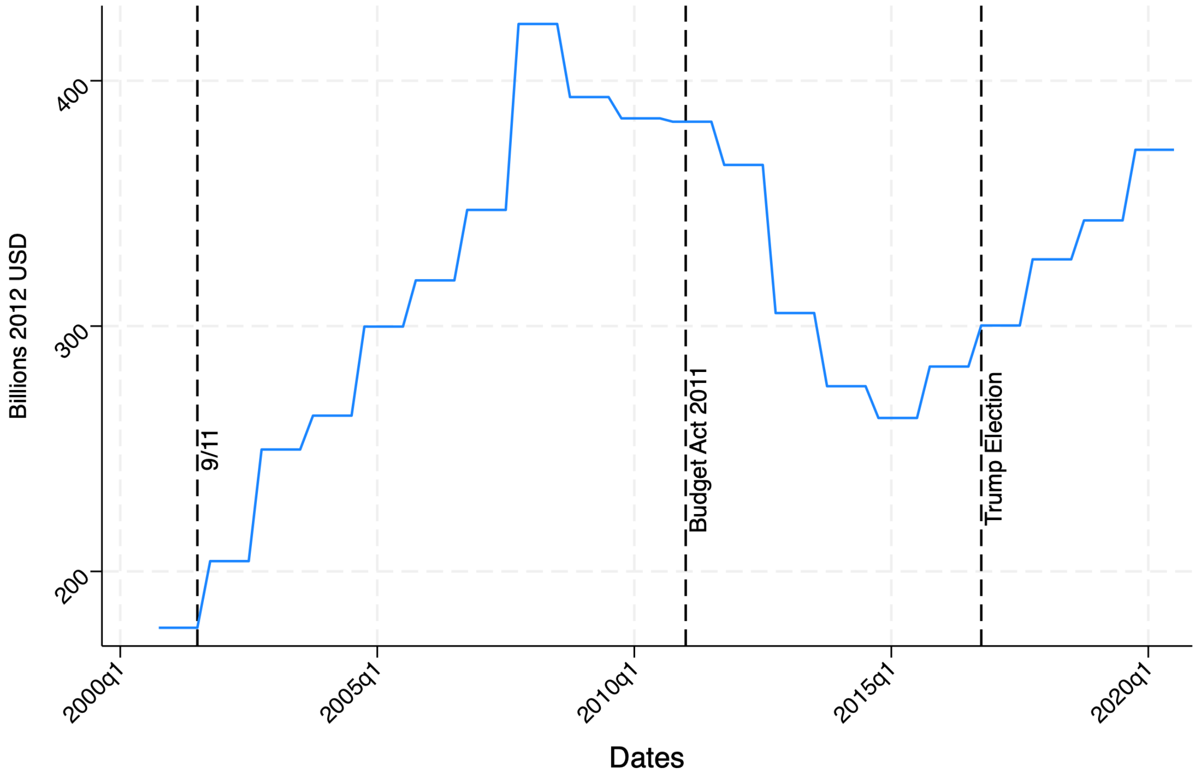


Figure A5: Military Contracts by Fiscal Year in the Federal Procurement Data System

series of quarterly new military contracts using the X-13 ARIMA-SEATS program from the Census Bureau. The top-left panel of Figure A6 compares the seasonally adjusted series (red line) with the original series (blue line). Second, we adjust the series to ensure that the average values within each fiscal year align with the original fiscal year values of military contracts. This fiscal year adjustment is shown in the top-right panel of Figure A6. Note that the series of contracts after the fiscal year adjustment (blue line) lies above the seasonally adjusted series of new contracts, as total fiscal year procurement contracts include both new contracts and modifications. Finally, the bottom-left panel of Figure A6 compares the fiscal year series of military contracts (Figure A5) with the newly created quarterly, seasonally adjusted series. By construction, the average of the quarterly values within a fiscal year coincide with the original/raw fiscal year data (i.e., Figure A5), while quarterly variation reflects changes in newly awarded, seasonally adjusted military contracts.

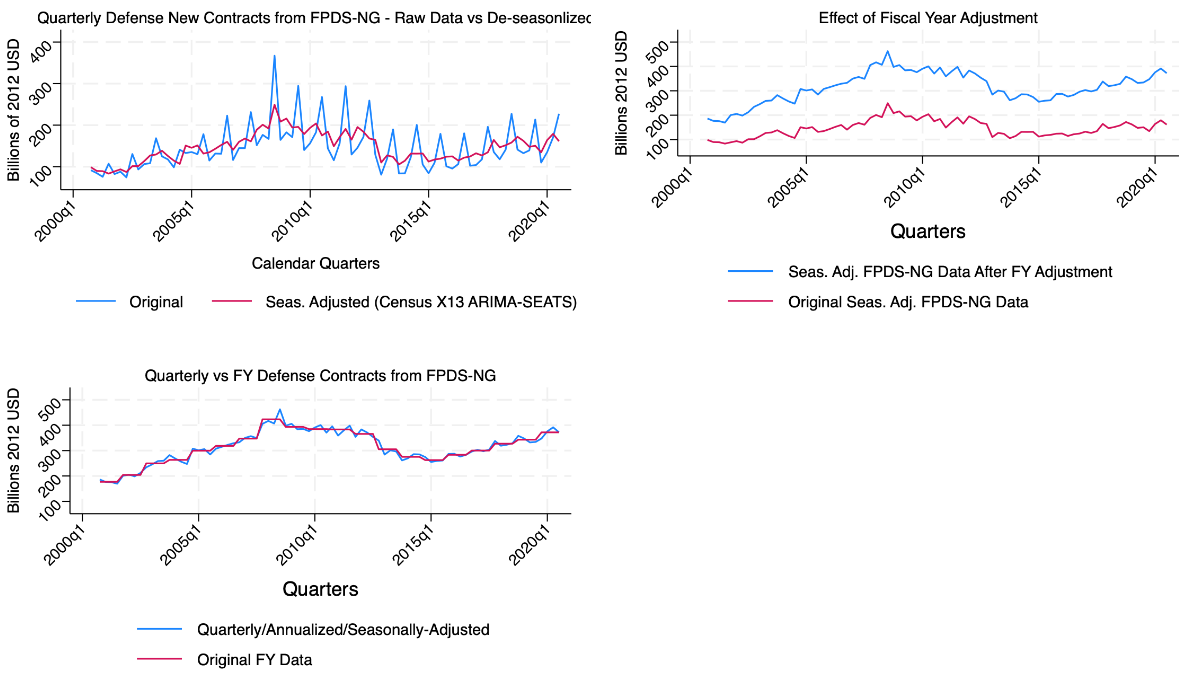


Figure A6: Construction of Quarterly Military Contracts from the Federal Procurement Data System

Notes: Price deflator is the GDP price deflator.

A.4 BCD Extrapolation

Data from Business Condition Digest (and Business Cycle Developments for years before 1961) are available starting from January 1951, immediately following the outbreak of the Korean War, the largest military shock in the post-WWII sample. We reconstruct the series of military procurement contracts from 1947:1 to 1950:4 using information from two variables highly correlated with military prime contracts: (i) average weekly hours of production workers in aircraft manufacturing and (ii) defense procurement spending data from the NIPA.

Average Hours in Aircraft Military production in the 20th century was heavily driven by the aircraft industry, as most military items were aircraft and their engine and navigation components. This is confirmed by the Top 100 companies reports from the Department of Defense, where the top defense contractors were predominantly aircraft and parts manufacturers (see Fisher and Peters

(2010), McGrattan and Ohanian (2010), Nekarda and Ramey (2011), and Ilzetzki (2023)). Average weekly hours of production workers is an excellent proxy for capital utilization and tracks ongoing production (see Bils and Cho (1994) and Fernald (2012)). The bottom panel of Figure A7 displays the time series of real military prime contracts per capita in blue (left axis) and average weekly hours of production workers in aircraft manufacturing in red (right axis).

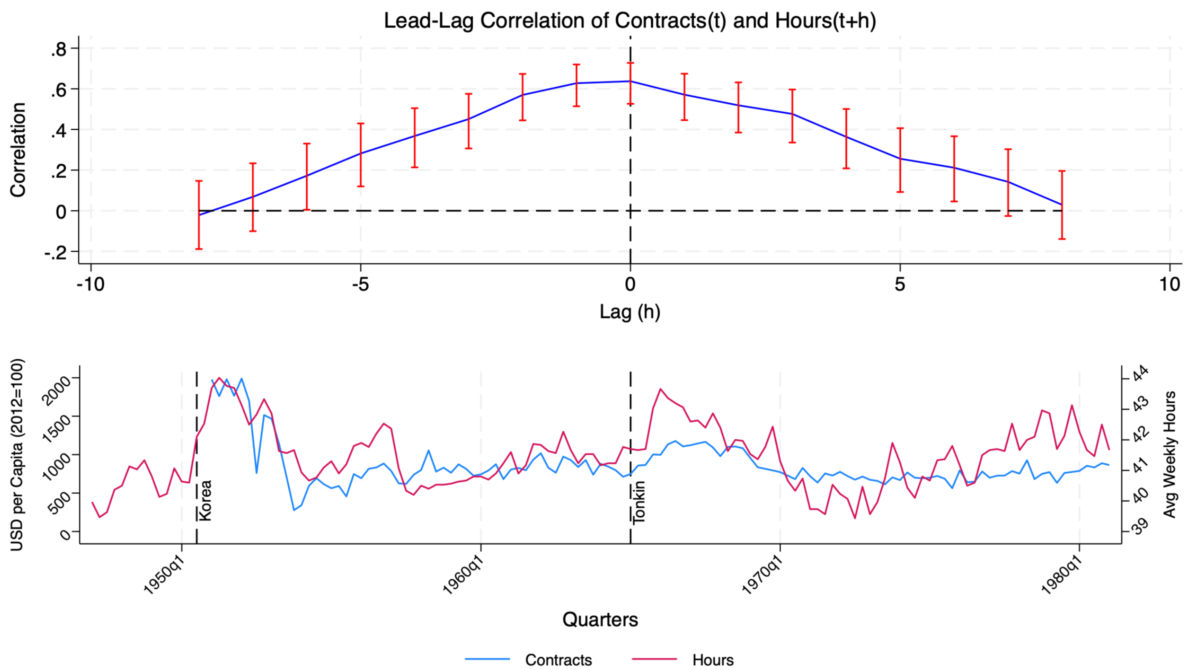


Figure A7: Comovement in Weekly Hours of Aircraft Production Workers and Military Contracts

Notes: Top panel shows the lead-lag correlation map between real military prime contracts (MPC) per capita and average weekly hours of production workers in aircraft manufacturing (aircraft). The price deflator is the GDP price deflator. Bottom panel plots of the two series.

The co-movement between the two series is evident. We quantify this co-movement using a lead-lag correlation map, shown in the top panel of the figure. Notice that the correlation between the two series reaches a global maximum when the timing of hours and contracts coincide. The correlation diminishes when hours are either delayed (h is positive) or anticipated (h is negative), indicating that the two series co-move.

NIPA Defense Procurement Spending Secondly, we construct a proxy for defense procurement spending following Cox, Muller, et al. (2023), who sum NIPA's (i) defense intermediate goods and services purchased and (ii) defense gross investments in structures, equipment, and software. Defense procurement spending from the NIPA closely tracks military procurement contracts. The former is the accounting field in the NIPA which accrues to military contracts. In fact, all (obligated) military contracts will eventually be accounted for in G as defense procurement spending by NIPA, following the initial award. The two variables differ only in timing: contracts are recorded at the award date, while most items are recorded into NIPA based on delivery times and/or payments to contractors, which follow the award date and production by military contractors.

The bottom panel of Figure A8 shows the time series of real military prime contracts per capita in blue and real NIPA defense procurement spending per capita in red.

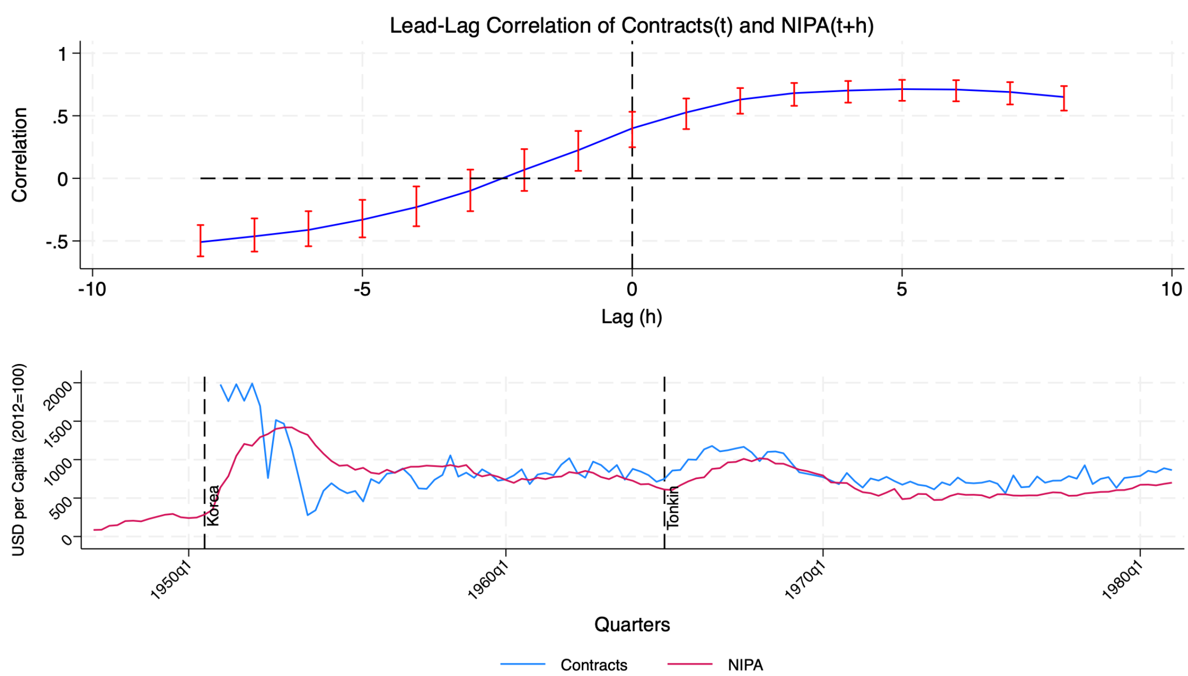


Figure A8: Military Contracts Lead NIPA Defense Procurement Spending

Notes: Top-panel shows lead-lag correlation map between real military prime contracts (MPC) per capita and real defense procurement spending per capita (NIPA). Formally: $\text{Corr}(\text{MPC}_t, \text{NIPA}_{t+h})$. Price deflator is the GDP price deflator. Bottom-panel plots of the two series.

Military prime contracts appear to lead NIPA defense procurement spending. In fact, the lead-lag correlation map, illustrated in the top panel of Figure A8, shows a positive correlation between the series only when NIPA is delayed.

Extrapolation Given these data characteristics, we predict real military prime contracts per capita using (i) contemporaneous values of average hours of production workers in aircraft manufacturing and (ii) both contemporaneous and future values of real defense procurement spending per capita.

In particular, we estimate the following equation via OLS, spanning from 1951:1 to 1980:4:

$$\text{MPC}_t = \kappa + \beta \cdot (\text{Avg.Hours.Aircraft})_t + \sum_{h=0}^{4} \psi_h \cdot \text{NIPA}_{t+h} + \varepsilon_t \tag{1}$$

Table A1 reports the OLS estimates of equation (1), and Figure A9 displays the original series of real military prime contracts per capita in blue, along with the resulting predicted series of contracts in red.

Table A1: Predicting Military Contracts with Aircraft Production Hours (1947Q1-1980Q4)

Dependent: Military Contracts	*Coefficient*	*Std. err.*	*t*	*p-value*	*[95% conf.*	*interval]*
Avg. Weekly Hours (Aircraft)	51.646	23.745	2.175	.032	4.585	98.707
NIPA						
–.	-1.277	.38	-3.361	.001	-2.029	-.525
F1.	.172	.578	.298	.766	-.973	1.317
F2.	.923	.59	1.564	.121	-.247	2.093
F3.	.877	.593	1.479	.142	-.299	2.053
F4.	-.065	.431	-.151	.88	-.919	.79
R^2 64.84%						
T 116						

Notes: Constant is not reported in the output-table. NIPA refers to real defense procurement spending per capita. Price deflator is the GDP price deflator. Average weekly hours of production workers in Aircraft manufacturing are available monthly from 1947 from the discontinued database of the Bureau of Labor Statistics.

Notice that the predicted series closely tracks the original one. We will use the predicted values of military prime contracts from 1947:1 to 1950:4 to fill the gap. We will refer to this data as "*BCD Extrapolated*".

Among these extrapolated data points, it is important to analyze the year 1950, which includes

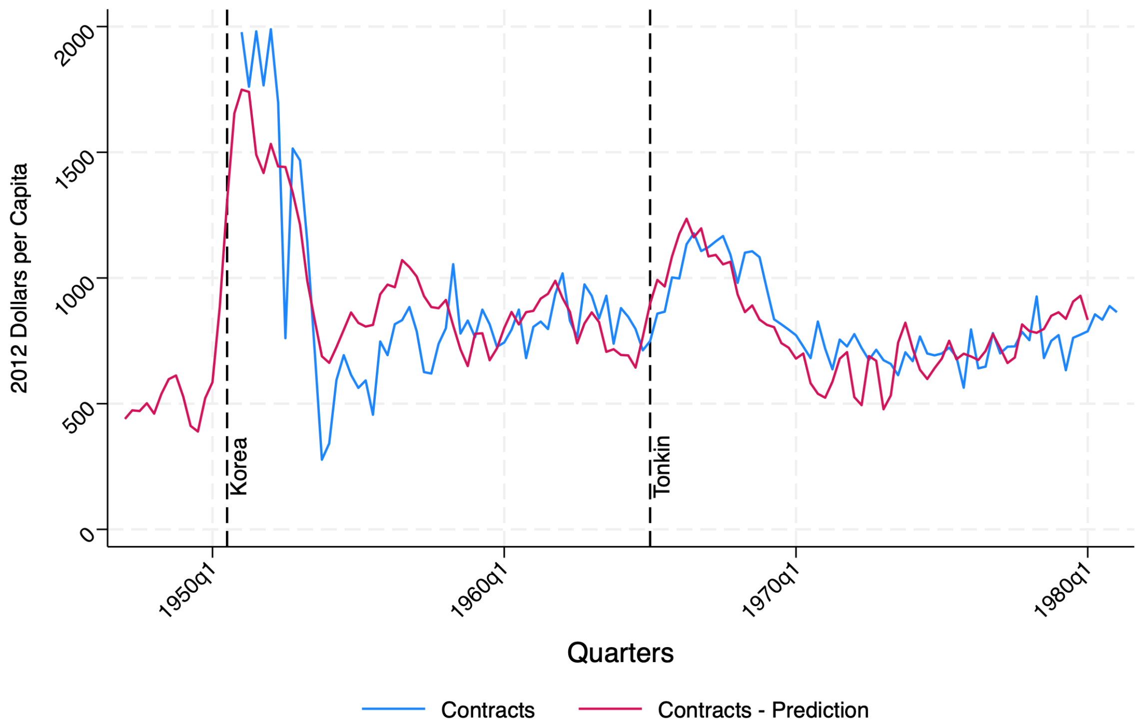


Figure A9: Predicted Military Contracts using Aircraft Production Hours

Notes: Variables are in real per capita values. Price deflator is the GDP price deflator.

the outbreak of the Korean War. According to the extrapolated data, military prime contracts increased in the second quarter of 1950. This is consistent with the narrative from the Survey of Current Business Edition of August 1950, which states that military production of aircraft had already started increasing. From page 4:[3]

“ *Chart 2 shows the trends in production of 16 finished products over the past 2-1/2 years, as well as comparable data for 1940. The increases from the first to the second quarter of 1950 were particularly striking for the durable finished goods shown in the chart. In most cases the rates of output represented new peaks, which generally ranged from 60 to well over 100 percent above the prewar volume.* ***The largest second quarter gains were in aircraft, a reflection of the substantial orders placed for military account,*** *and in passenger cars [...].* ”

[3] See Chart 2 on page 3 of SCB August 1950 as well as page 4, section “Expansion centered in durables”.

Following the invasion of South Korea on June 25, 1950, extrapolated military contracts exhibit a sharp increase in the third quarter of 1950. This is again in line with the narrative from the Survey of Current Business; from page 9:[4]

“ *The outlook for these programs [military expenditures of the Defense Department] subsequent to the second quarter, however, has of course been altered radically upward by the United States response to the Korean hostilities. Both* ***procurement*** *and military payrolls, as well as a wide variety of supporting outlays,* ***will mount rapidly*** *as the announced expansion of our military strength gets under way.*”

[4]The same section of the SCB at page 9 (Government Purchases maintained) also discusses a slight drop in military expenditures before the outbreak of the Korean war. However, for military expenditure they refer to a combination of (i) outlays, (ii) payrolls and (iii) procurement and not necessarily to prime contract awards. On the contrary, at page 4 they clearly mention an increase in 1950:2 of military orders of aircraft, witnessed by the documented increase in aircraft production.

A.5 World-War II Data

Our last source of military prime contract data comes from Brunet (2024), who utilized digitized documents from the Civilian Production Administration. The dataset encompasses all contracts valued at $50,000 or more awarded between June 1940 and September 1945, excluding only food products. The contracts include airplanes and victory ships, as well as intermediate goods (e.g., propellers and gun fittings), raw materials (e.g., aluminum and leather), and smaller items (e.g., mattresses, gloves, insect repellent, and toilet paper).

We aggregate military prime contracts by quarter and annualize their value. Figure A10 displays the value of real military contracts and real government spending, i.e., G, available from Ramey and Zubairy (2018)'s online database.

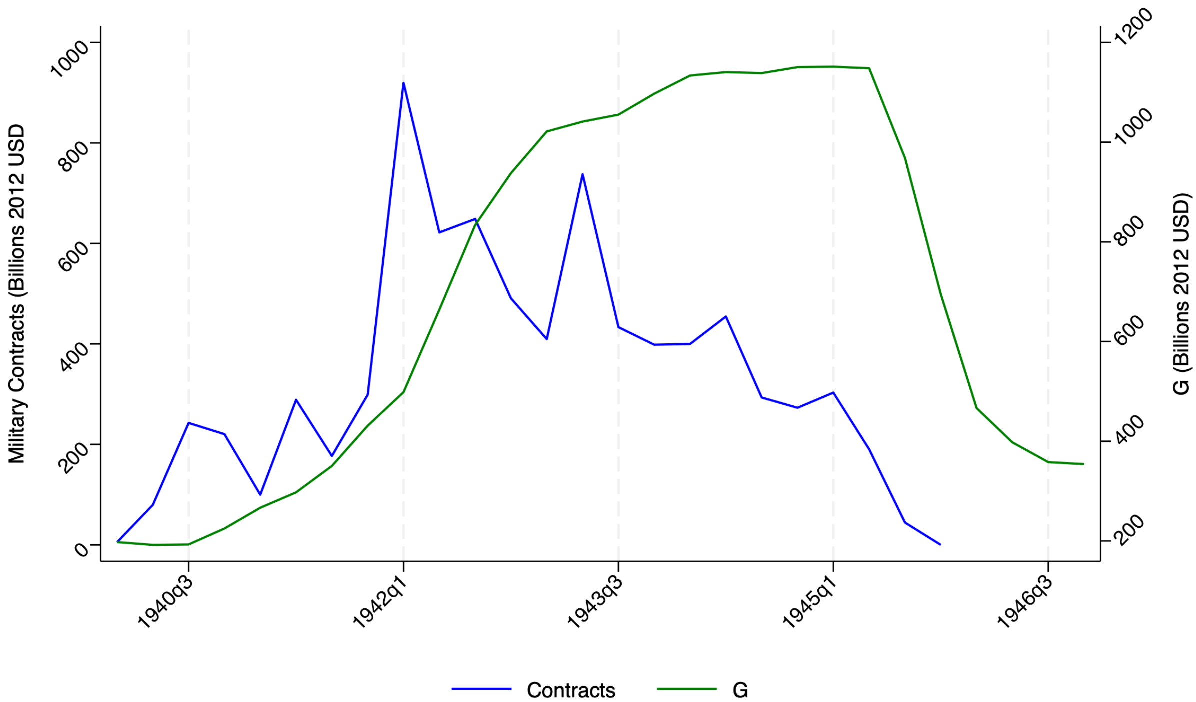


Figure A10: Military Contracts Lead NIPA G during WWII

Notes: The price deflator is the GDP price deflator. The price deflator from Ramey and Zubairy (2018) uses 2009 as the reference year, while the currently available price deflator uses 2012 as the reference year. We divide the price deflator in RZ18 by its value in 2012 to convert the price index relative to the year 2012.

Note that contracts lead the increase in G. Unfortunately, the series does not report contracts after 1945:4, and the series of extrapolated military contracts from BCD only starts in 1947:1. Information from the United States Budget for the fiscal year 1948 suggests that military prime contracts were very small, if not null, following the end of WWII. In particular, Table 7 from page A110 reports that the value of military contract authorizations was even negative, on aggregate, indicating the de-obligation of previously awarded military contracts.[5]

Following the significant restructuring of the federal government mandated by Truman's National Security Act of July 1947, which disbanded the War and Navy departments and created a unified Department of Defense, precise data on military contracts for those years is missing.

We will fill the gap in military prime contract data from 1945:4 to 1947:1, i.e., the calendar year 1946, using a simple linear interpolation.

[5]Specifically, Table 7 reports a value of negative $1,715,103,983 for the fiscal year 1946 and negative $1,352,058,297 for the fiscal year 1947.

A.6 Constructing a Quarterly Series of Military Prime Contracts

Lastly, we assemble a quarterly series of military prime contracts using the previously discussed series. Specifically, the series composition is as follows: (i) data from 1940:1 to 1945:4 are derived from contracts documented by Brunet (2024); (ii) data for calendar year 1946 are constructed using linear interpolation; (iii) data from 1947:1 to 1950:4 are constructed using extrapolated data from BCD, average weekly hours of production workers in aircraft manufacturing, and NIPA defense procurement spending; (iv) data from 1951:1 to 1980:4 are sourced from BCD; (v) data from 1981:4 to 2003:3 are sourced from FPSR; and (vi) data after 2003:3 are sourced from FPDS-NG.

Figure A11 displays real prime contract measures alongside Ramey and Shapiro (1998)'s war dates, augmented with the events of the fall of Paris in the second quarter of 1940, the 9/11, the Budget Control Act of 2011, and the election of President Trump.

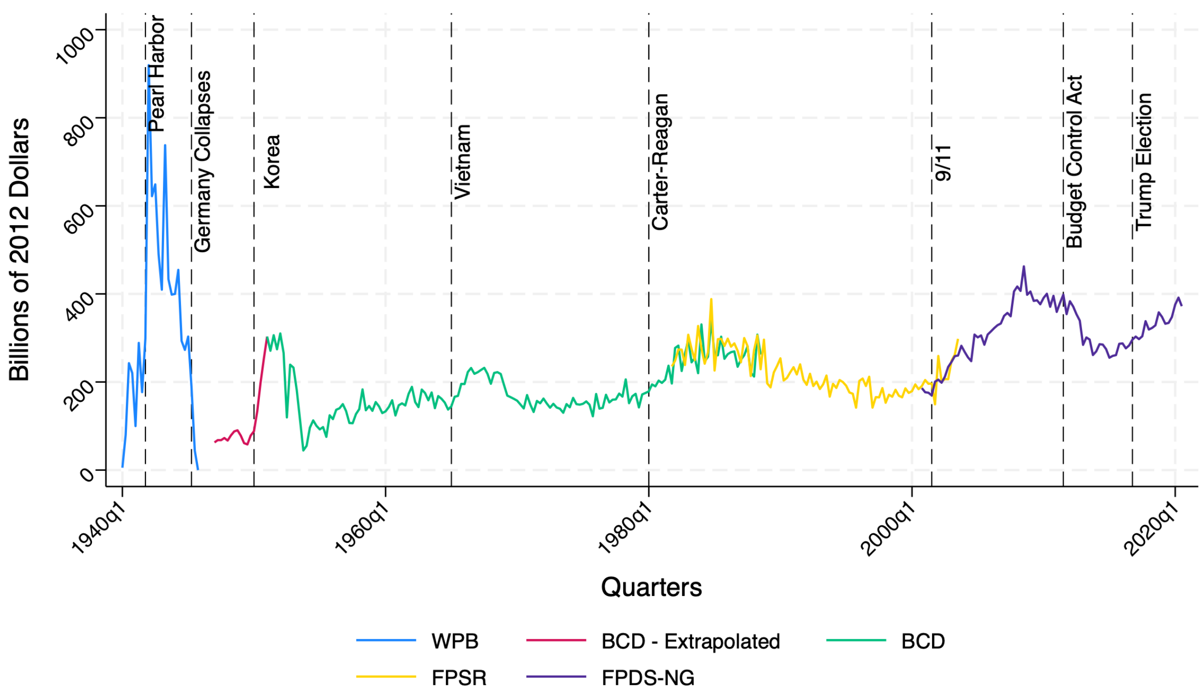


Figure A11: Measures of Military Prime Contracts

Notes: The price deflator is the GDP price deflator. Price deflator before 1947:1 comes from Ramey and Zubairy (2018), adjusted to be consistent with 2012 as base year.

Notice how well the measures overlap in the 1980s and at the beginning of the 2000s, indicating remarkable consistency across the different data sources.

Lastly, Figure A12 displays the appended series of real defense contracts (red line) alongside the value of NIPA's defense procurement spending, available only from 1947:1 (blue line). From this point forward, this new variable will be referred to simply as "*(defense) contracts*".

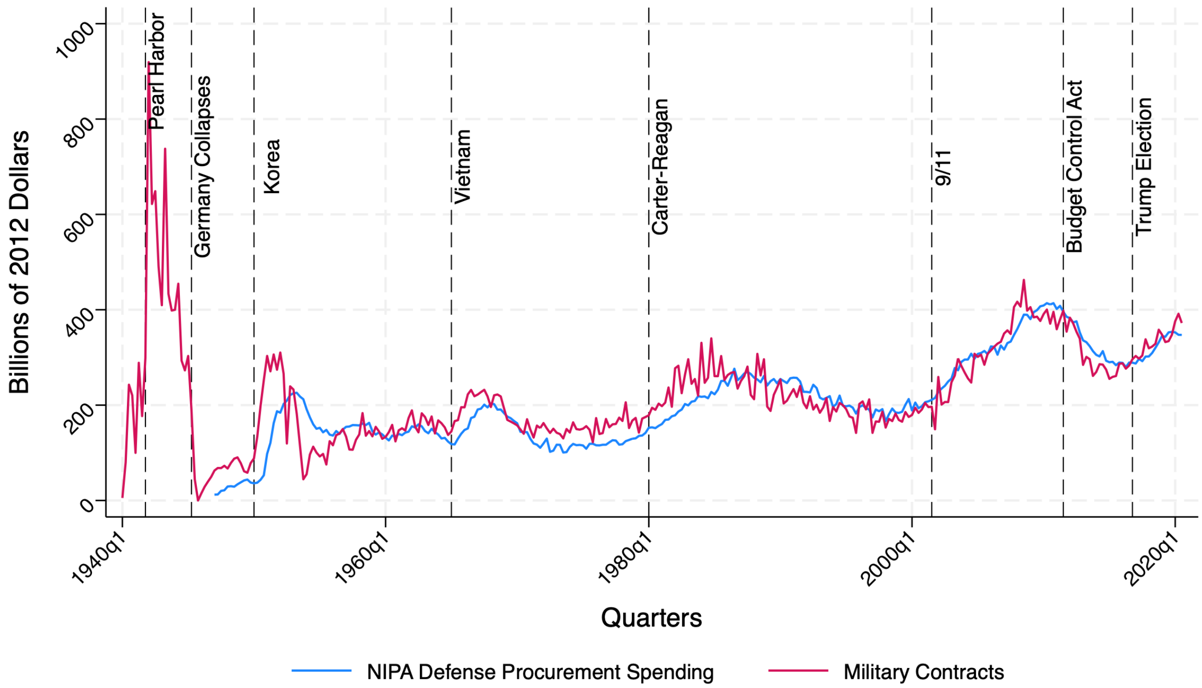


Figure A12: Measures of Military Prime Contracts

Notes: The price deflator is the GDP price deflator. Price deflator before 1947:1 comes from Ramey and Zubairy (2018), adjusted to be consistent with 2012 as base year. Data for calendar year 1946 is constructed with linear interpolation.

A.7 Comparison of Contract Series with Dupor&Guerrero 2017

In this section we compare our newly constructed series of defense contracts, aggregated by fiscal year, with the series of defense contracts from Dupor and Guerrero (2017), referred to it as DG17. This is a sanity check which compares our data with a benchmark in the literature. Figure A13 shows the two series.

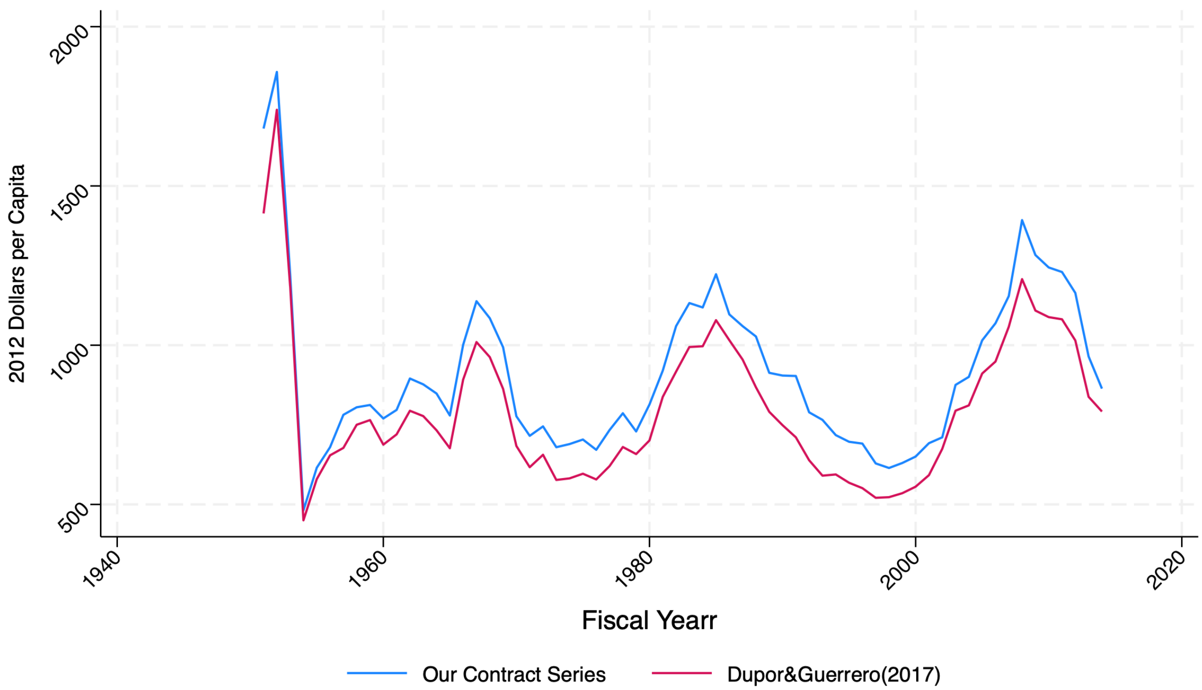


Figure A13: Contract Series is consistent with Dupor and Guerrero (2017), but is slightly higher due to inclusion of small (<$10k) contracts

Notes: Data from Dupor and Guerrero (2017) comes from their gun4 series.

DG17 contains all contracts with nominal value larger than $ 10,000 before FY1984 and all contracts with nominal value larger than $25,000 from FY1984, and, in turn, closely tracks Nakamura and Steinsson (2014) series. Notice that our series of contracts follows the same dynamics as DG17. Since our measure of contracts does not exclude small contracts, it always lies above DG17.

A.8 Further Details on Timing of Defense Spending in NIPA "G"

When possible, the BEA uses the direct pricing method to estimate government spending on many types of military goods, including purchases of aircraft, missiles, vehicles, and petroleum products. Under the direct pricing method, the timing of spending recorded in NIPA reflects the delivery of goods to the military. As explained by the Bureau of Economic Analysis:

> An example of the direct-pricing method is the derivation of the F-15 aircraft estimates. The F-15 aircraft is usually purchased by DOD [the Department of Defense] as component parts. The components—such as engines, airframes, guns, and various electronic subsystems—are supplied as government-furnished equipment to the airframe contractor who performs the final integration and assembly. Quarterly estimates of the value of component parts are obtained by multiplying the quantity delivered to DOD by the price paid by DOD. The estimates for each component, including integration and assembly, are summed.[6]

Under the direct pricing method, government spending is recorded as finished components are delivered, which occurs after—sometimes significantly after—the components are produced. In the interim, these government purchases show up in NIPA as inventory investment by the firms producing the products. When the products are delivered, NIPA records a negative inventory investment (as firms' inventories of military goods are reduced) and positive government purchases.

The other two methods for measuring military spending, the ratio method and direct estimation, both rely on the disbursement of funds from the Treasury—that is, when contractors are paid. While the details depend on the contract, often final payments are not made until after goods are delivered to the government. Again, NIPA often measures government spending long after production occurs. This is not a flaw in NIPA, but rather by design: NIPA measures the flow of physical goods and dollars through economic transactions, not the production of goods.

[6] BEA (2005), II: 33-34.

Since budget authority captures authorizations for spending, it necessarily leads disbursements (payments), since contracts cannot be made without prior authorization and payments cannot start until contracts are granted. Deliveries often come later, sometimes much later. In many cases final payments are made after delivery

B Measurement Delays and Timing Relationship Across Measures of Spending

In this section of the Appendix, we analyze in detail the timing relationships across different measures of government spending. Specifically, we utilize (i) lead-lag correlation maps and (ii) Granger causality tests across variables to quantify the time delays and formalize the concept of anticipation.

B.1 Budget Authority and Contracts

The top panel of Figure A14 displays the time series of contracts (blue line - left axis) and Budget Authority (red line - right axis) in real per capita values, beginning from 1947, with the price deflator being the GDP price deflator.

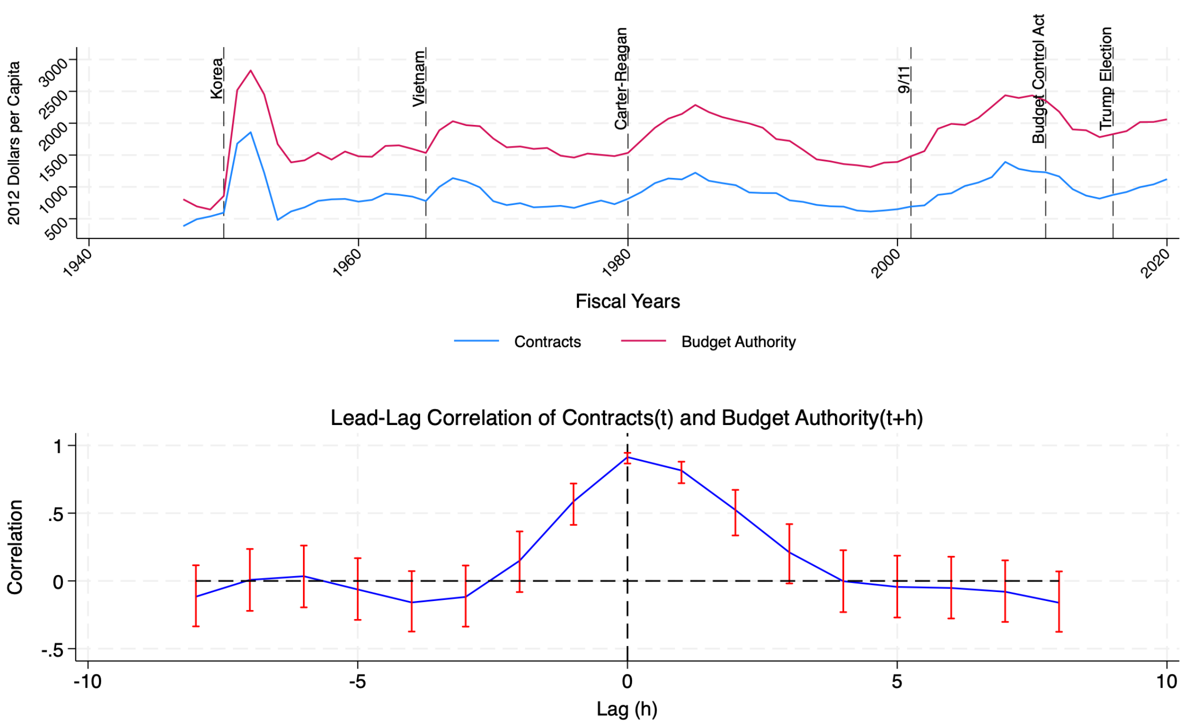


Figure A14: Comovement in Budget Authority and Military Prime Contracts (without WWII)

The two series demonstrate synchronized dynamics, with low-frequency fluctuations driven by military build-ups. The annual time-variation appears to be identical. This synchronization is corroborated by the bottom panel of Figure A14, which presents the lead-lag correlation map between the two time-series: the correlation is maximized when the timing of the two series aligns.

Figure A15 extends the same analysis to include WWII in the series of contracts and Budget Authority.

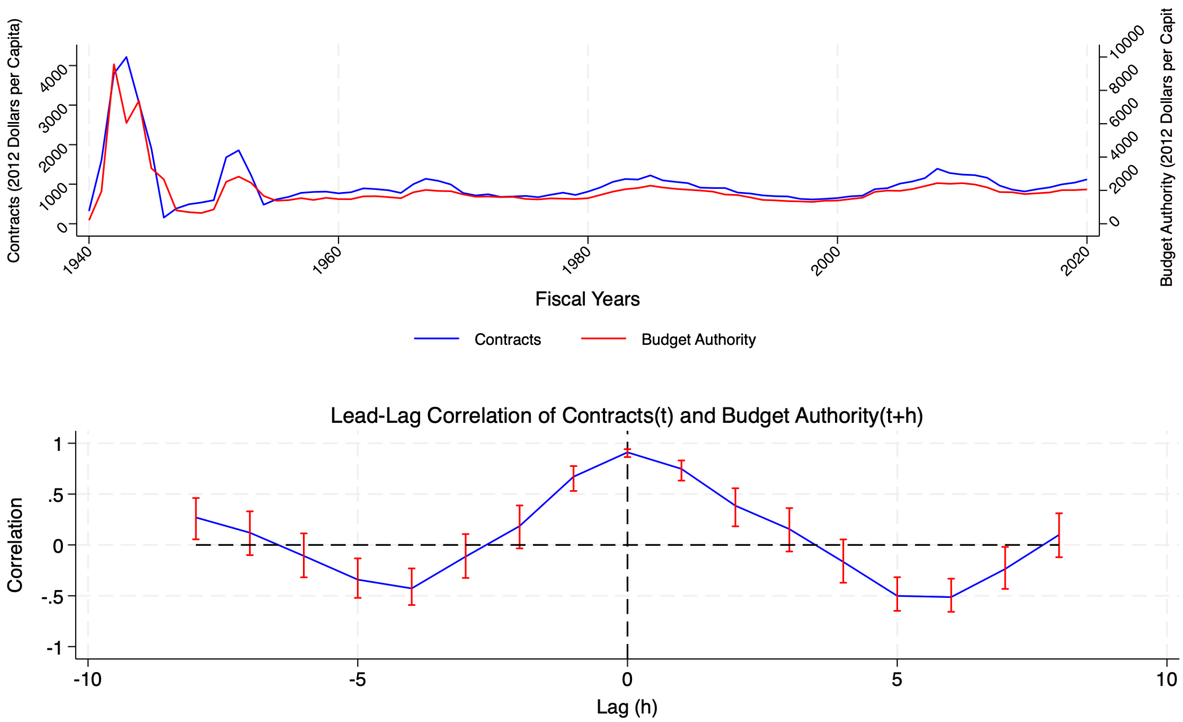


Figure A15: Comovement in Budget Authority and Military Prime Contracts (with WWII)

The results are consistent. The lead-lag correlation map continues to exhibit a global maximum at zero, indicating aligned timing of the variables.

To ensure robustness, we also examine the lead-lag correlation map between the two variables in first differences—i.e., $Corr\,(x_t - x_{t-1}, y_{t+h} - y_{t+h-1})$—which also shows a global maximum at $h = 0$. Figure A16 illustrates this with the top panel depicting results for the sample starting in 1947, excluding WWII, while the bottom panel includes WWII.

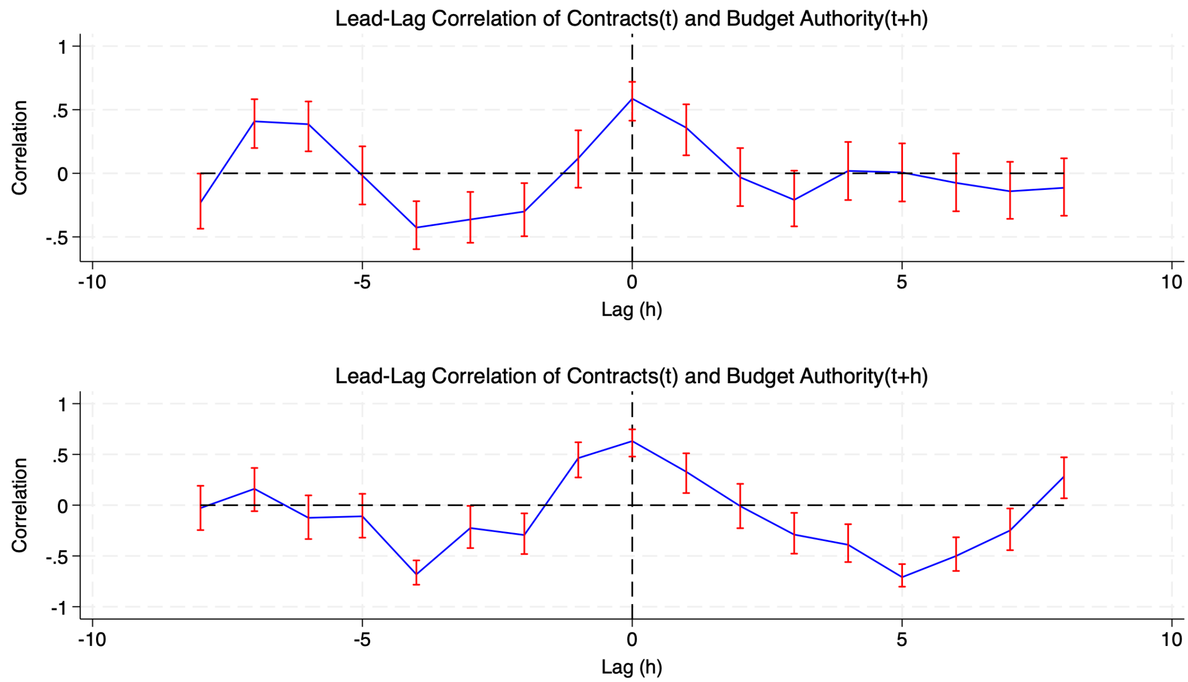


Figure A16: Comovement in First Differences of Budget Authority and Military Prime Contracts, both with (bottom panel) and without (top panel) WWII

B.2 Contracts and NIPA Defense Procurement Spending

We now examine the timing relationship between NIPA's measure of defense procurement spending and military prime contracts. The top panel of Figure A17 presents the time series for contracts (blue line) and NIPA defense procurement spending (red line) in real per capita values. The bottom panel of the figure illustrates the lead-lag correlation map between the quarterly year-to-year changes — i.e., $Corr\left(x_t - x_{t-4}, y_{t+h} - y_{t+h-4}\right)$ — of the two variables.

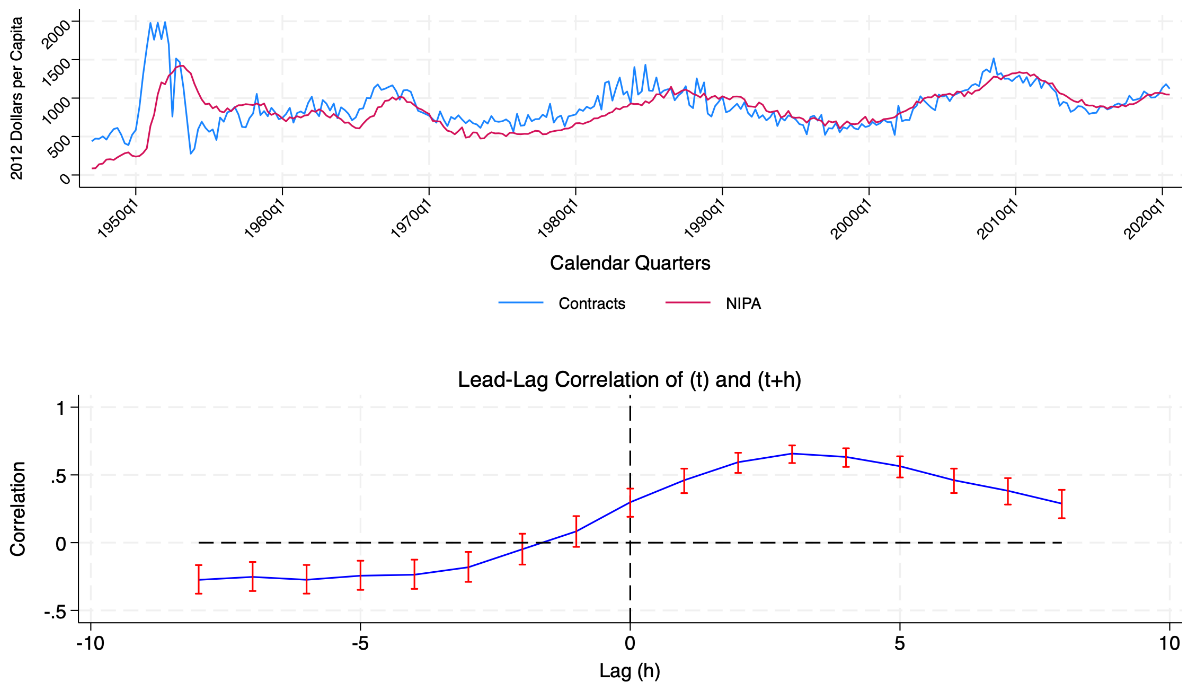


Figure A17: Contracts Lead NIPA Defense Procurement Spending by 9 Months on Average (including the Korean War)

The lead-lag correlation map indicates positive values when NIPA's series is delayed. Particularly, the correlation maximizes at $h = 3$, suggesting an average delay of about 9 months between the awarding of new contracts and their reflection in NIPA measures.

For robustness, Figure A18 excludes the Korean War from the sample, starting from the first quarter of 1954.

The plot of the two time series in the top panel shows that contracts lead NIPA. This delay is

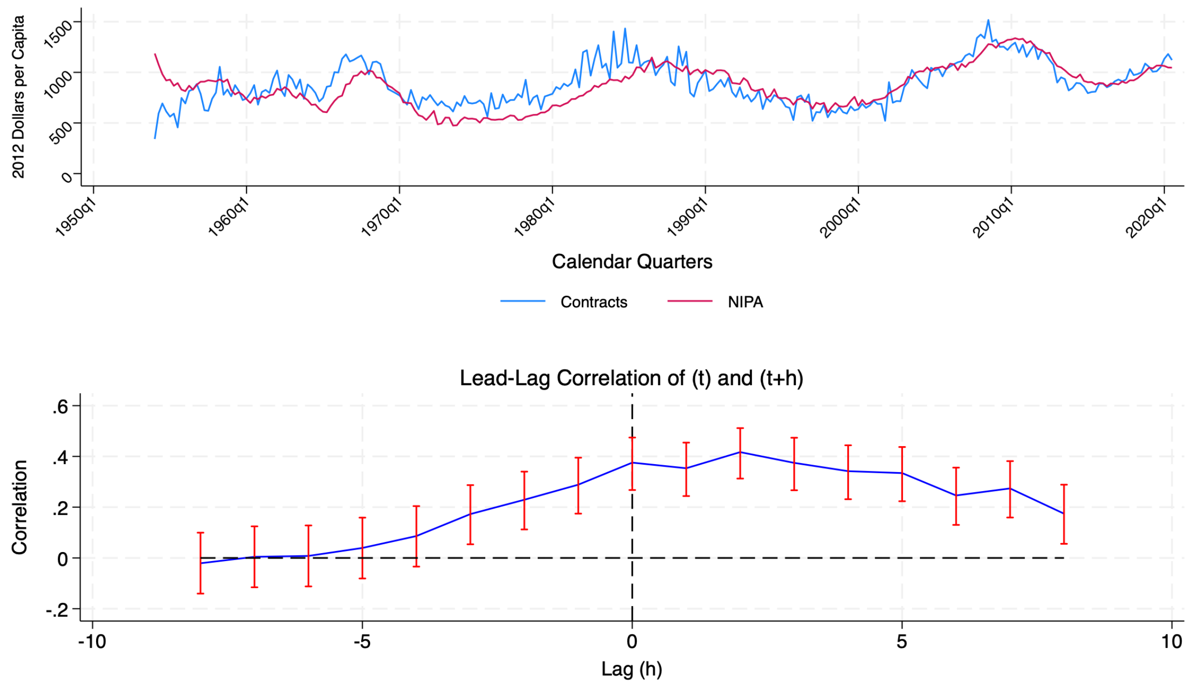


Figure A18: Contracts Lead NIPA Defense Procurement Spending by 6 Months on Average (excluding the Korean War)

evident around significant events like the Vietnam War (1965), the Carter-Reagan military buildup (1980), Obama's Budget Control Act (2011), and the subsequent reversal in 2015. The bottom panel's lead-lag correlation map generally exhibits larger positive values when NIPA is delayed, with a global maximum at $h = 2$.

We chose to present the lead-lag correlation maps in quarterly year-to-year differences instead of levels, observing that correlations measured at quarterly frequencies are very persistent. For completeness, Figure A19 includes the lead-lag correlation maps in levels, with the top panel including the Korean War and the bottom panel excluding it.

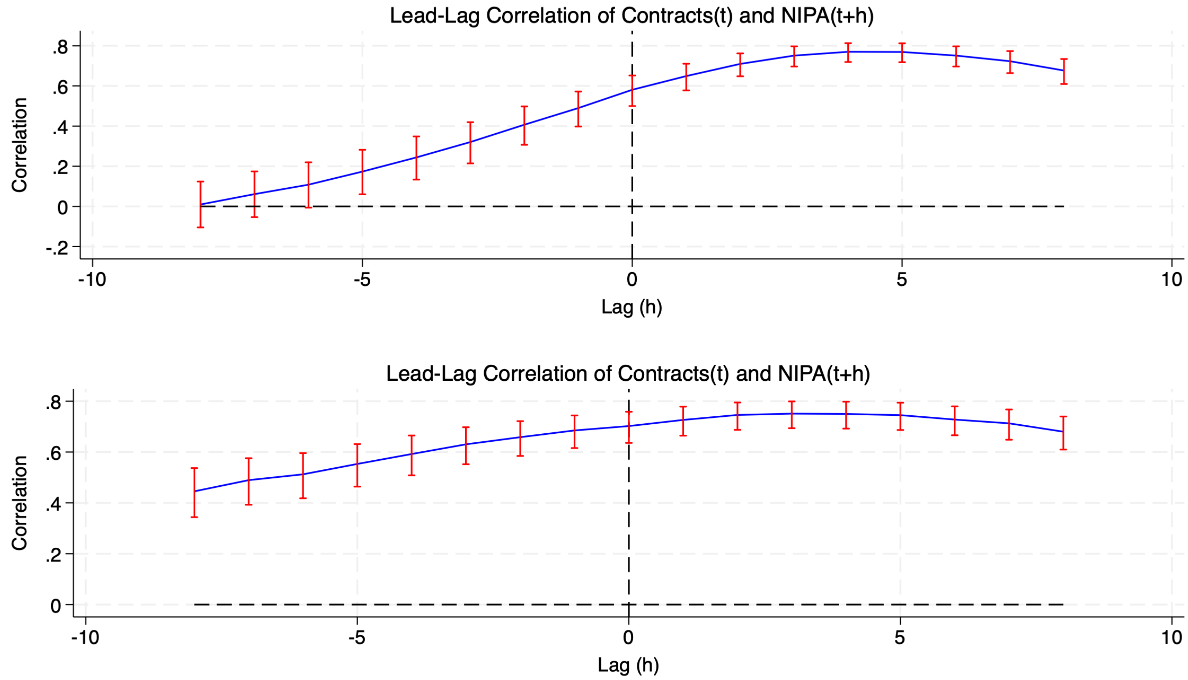


Figure A19: Contracts Lead NIPA Defense Procurement Spending by 3-4 quarters, both with (bottom panel) and without (top panel) the Korean War

Results remain qualitatively consistent: contracts appear to lead NIPA defense procurement spending since the correlation values are more positive and reach a global maximum when $h > 0$, indicating a delay in NIPA. The maximum correlation is achieved at $h = 4$ with the Korean War included, and $h = 3$ when excluded.

B.3 Budget Authority and NIPA Defense Spending

In this section, we compare the timing of Budget Authority with NIPA's measure of defense spending. The top panel of Figure A20 presents NIPA defense spending (blue line - left axis) alongside Budget Authority (red line - right axis) by fiscal year, starting from 1947, as we lack earlier data on NIPA's defense spending. Both variables are adjusted to real per capita values using the GDP price deflator.

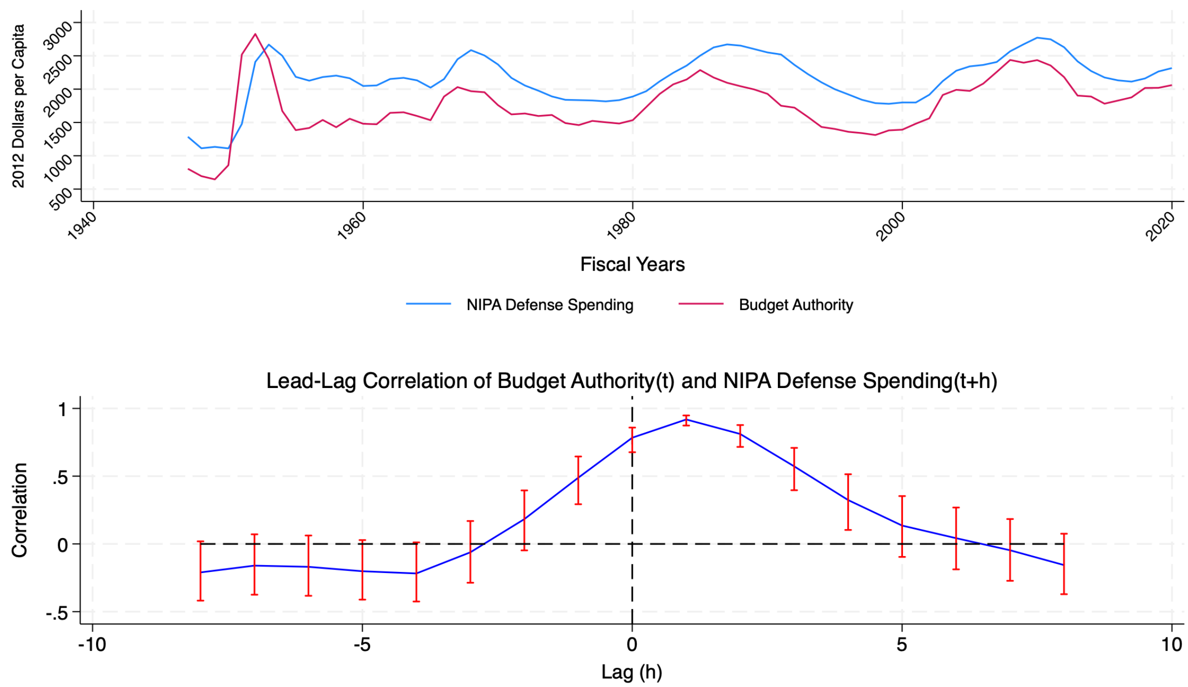


Figure A20: Budget Authority Leads NIPA Defense Spending by one year (with Korean War)

It is evident that variations in Budget Authority lead those in NIPA defense spending, as seen at the onset, peak, and end of each military buildup. The bottom panel of the figure quantifies the average delay by displaying their lead-lag correlation map. The correlation is generally larger when NIPA reporting is delayed—i.e., $h > 0$—and it reaches a global maximum after one year, indicating an average delay of twelve months for the government to spend authorized funds.

For robustness, the analysis is repeated, excluding the Korean War, with the sample starting in 1954. Figure A21 illustrates these results.

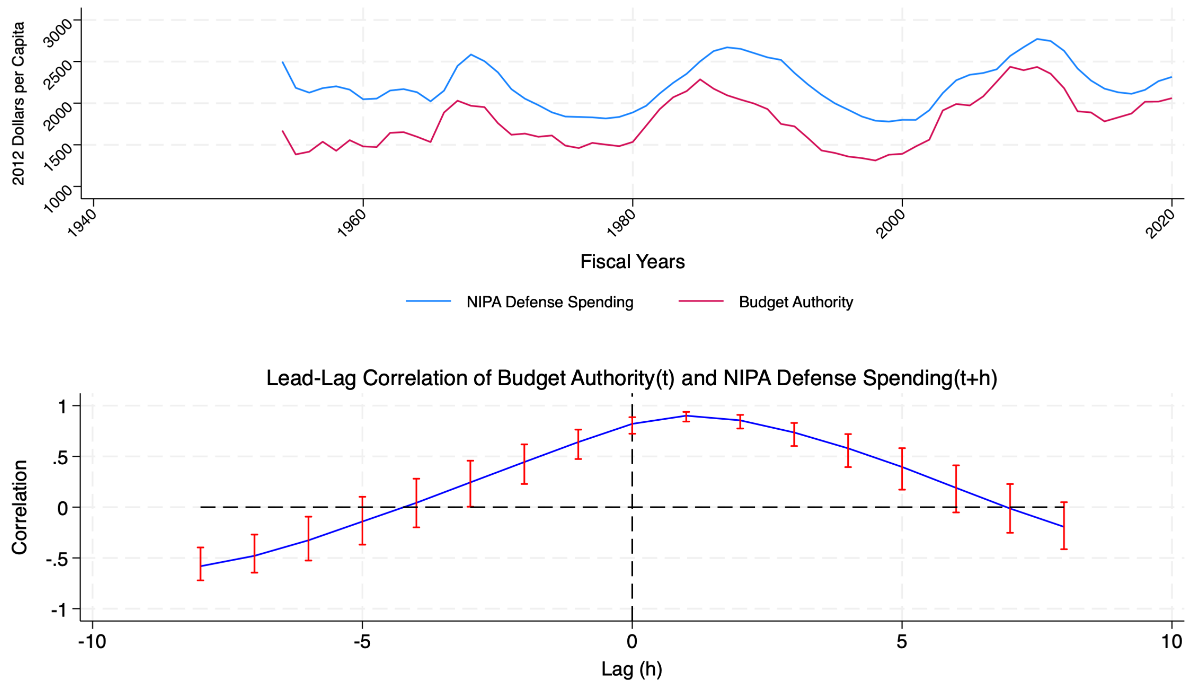


Figure A21: Budget Authority Leads NIPA Defense Spending by one year (without Korean War)

The top panel again makes the delay between Budget Authority and NIPA defense spending apparent. The lead-lag correlation map resembles the previous one, with the maximum still reached when NIPA reporting is delayed by one year.

Lastly, we report the lead-lag correlation maps calculated using first differences for both samples. The top panel of Figure A22 depicts the lead-lag correlation for the sample starting in 1947, which includes the Korean War. In this instance, the correlation between the two variables is significant and positive only when NIPA reporting is delayed, reaching a maximum at either 2 or 3 years. The bottom panel shows the scenario excluding the Korean War, with a similar pattern, although the correlation in their first differences is maximized when $h = 6$.

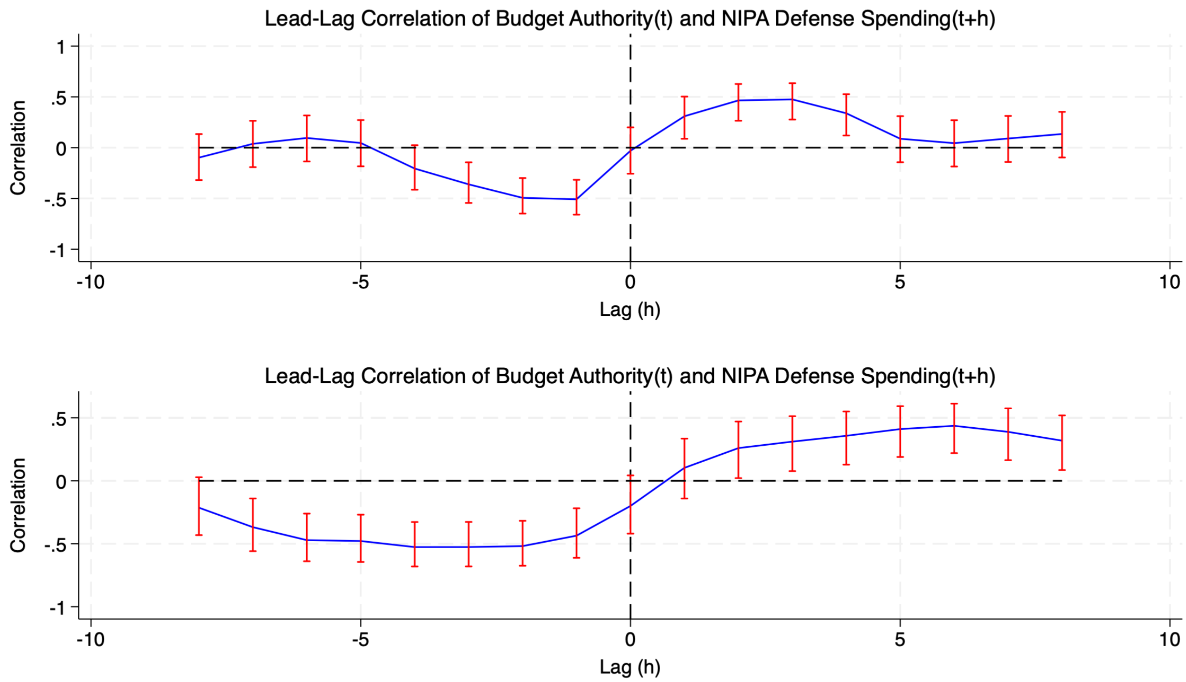


Figure A22: Budget Authority Leads NIPA Defense Spending in First Differences, with (top panel) and without (bottom panel) the Korean War

This analysis underscores the predictive nature of Budget Authority over NIPA defense spending across different contexts and periods, even when accounting for methodological adjustments like first differences.

B.4 Budget Authority and NIPA Defense Outlays

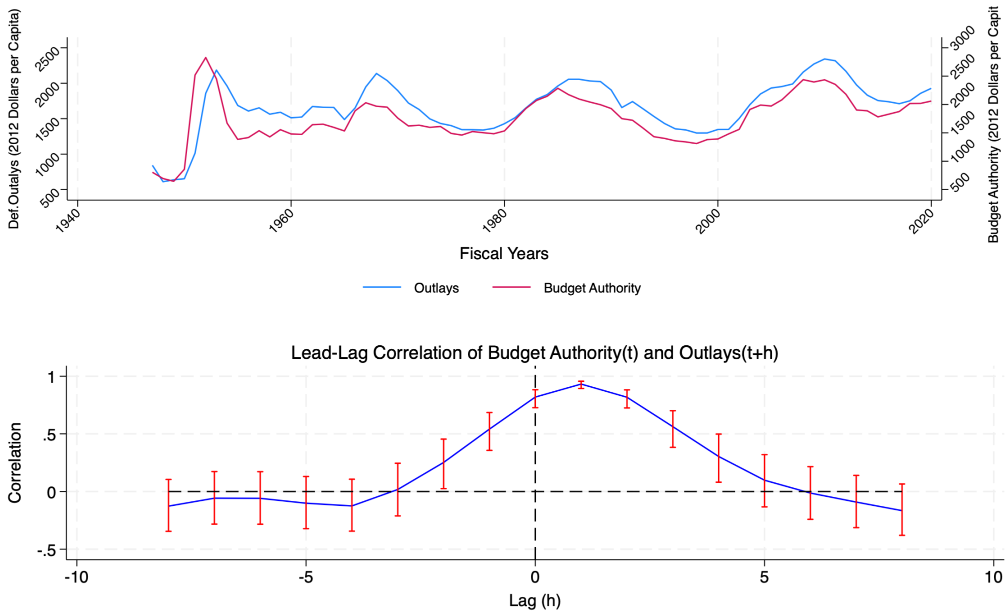


Figure A23: Budget Authority Leads Defense Outlays (with Korean War)

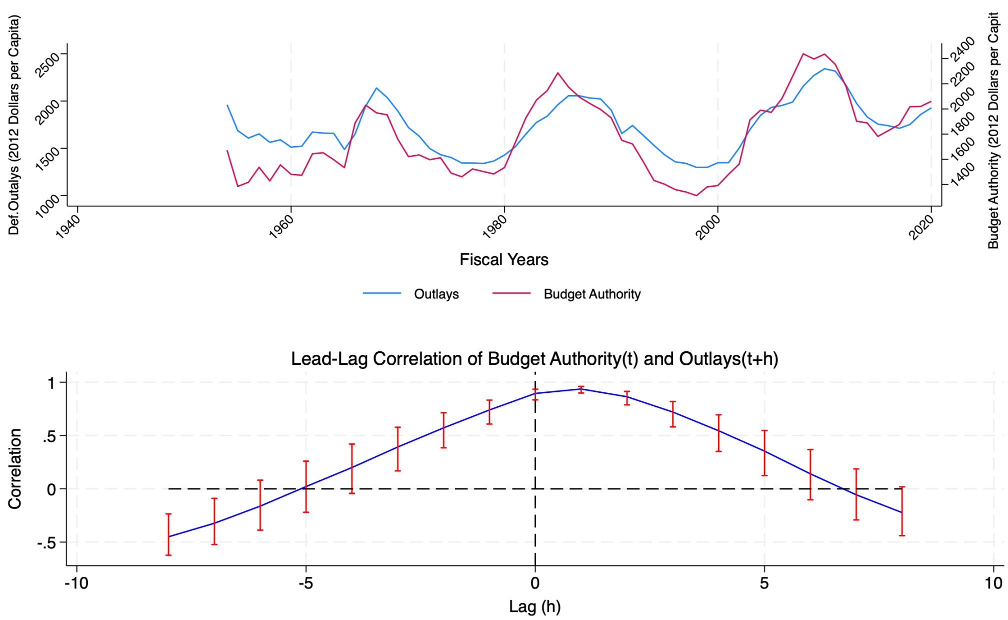


Figure A24: Budget Authority Leads Defense Outlays (without Korean War)

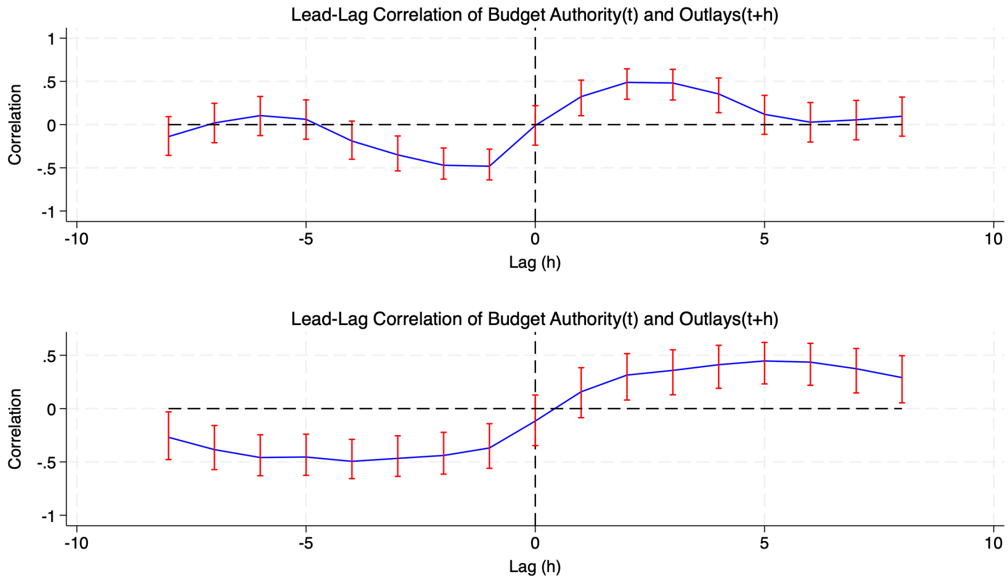


Figure A25: Budget Authority Leads Defense Outlays in First Differences, both with (top panel) and without (bottom panel) the Korean War

B.5 Spending Authorizations and NIPA Defense Spending

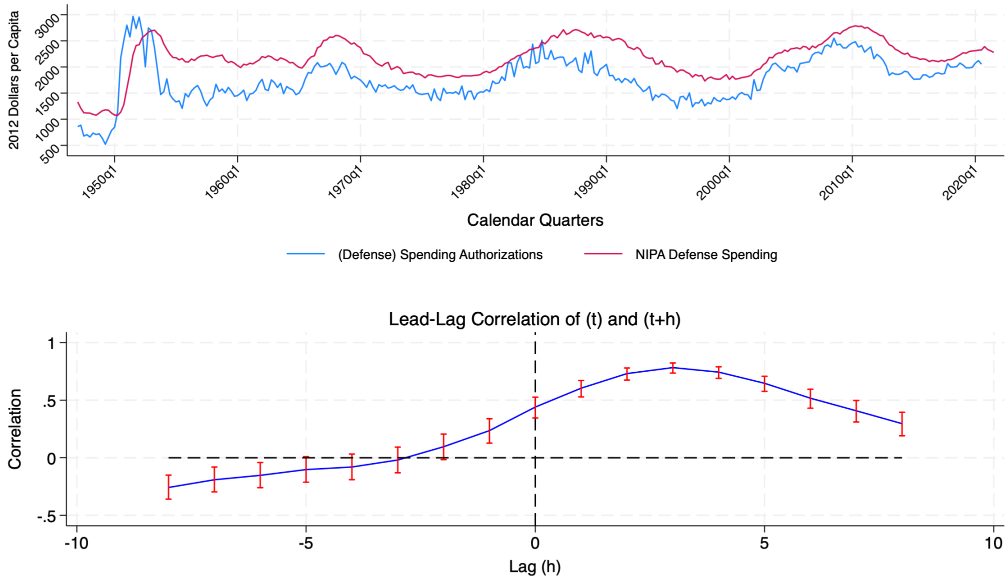


Figure A26: Spending Authorizations Lead NIPA Defense Spending (with Korean War)

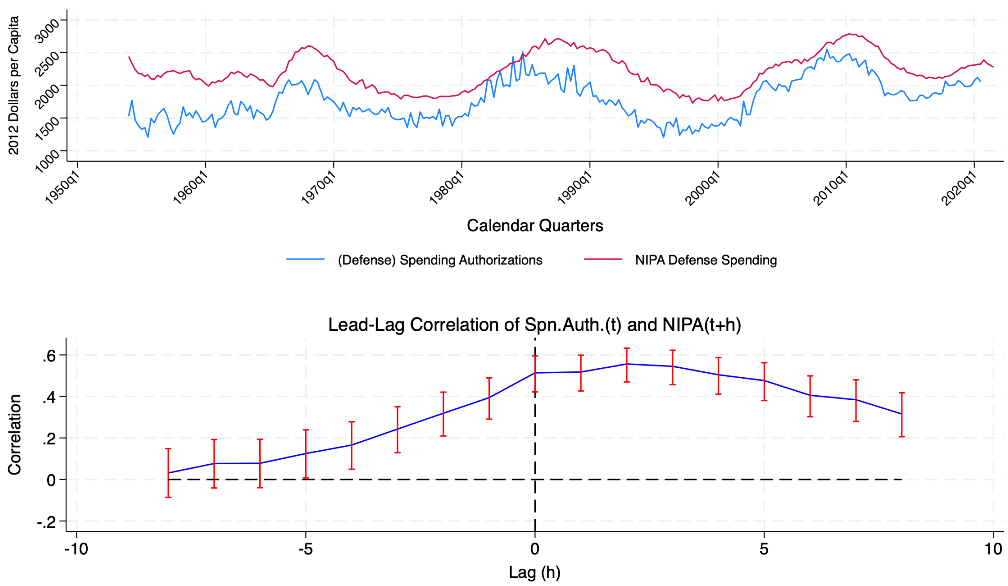


Figure A27: Spending Authorizations Lead NIPA Defense Spending (without Korean War)

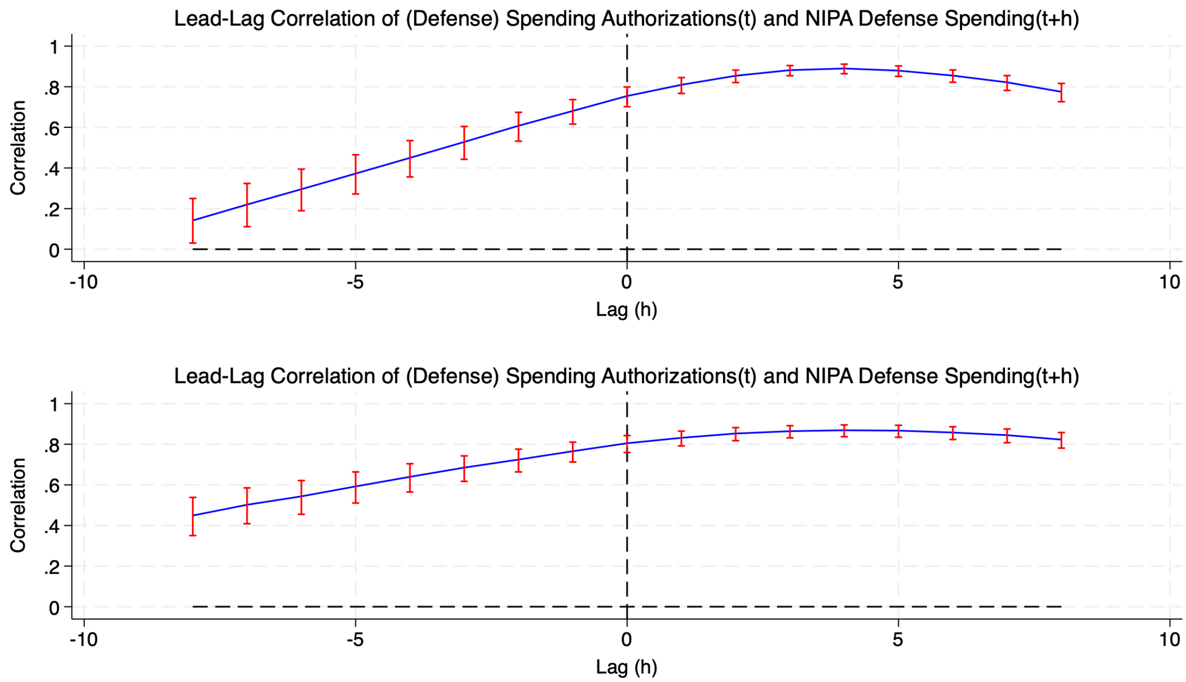


Figure A28: Spending Authorizations Lead NIPA Defense Spending in First Differences, both with (top panel) and without (bottom panel) the Korean War

B.6 Granger Causality Tests

In this section we carry out further Granger Causality tests to explore the timing relationship across different measures of government spending other than Spending Authorizations. Table A2 reports the results.

In the first panel of Table A2, we assess the predictability of NIPA's defense procurement spending using military prime contracts and vice versa. The large value of the F-statistics and its nearly zero p-value indicate that contracts Granger cause NIPA spending. This result is verified using data starting from the first quarter of 1947, the point at which NIPA data on defense procurement becomes available. Since data on military contracts (i) is extrapolated before 1951 and (ii) is significantly influenced by the outbreak of the Korean War, specifically in the third quarter of 1950 (Perotti (2014) and Ramey (2016)), we also present results starting from the first quarter of 1951. In this scenario, contracts retain predictive power over NIPA defense procurement spending. Conversely, NIPA defense procurement spending shows limited predictive power for contracts, with values of the F-statistics an order of magnitude smaller than those of contracts, and p-values above the 10% significance level. Overall, the results indicate that contracts consistently Granger cause NIPA's defense procurement spending measure, but the opposite is not true.

In the second panel of Table A2, we examine the timing relationship between Budget Authority and contracts. To facilitate this analysis, we (i) extend the contracts series back to fiscal year 1940 using data from Brunet (2024), and (ii) aggregate contracts by fiscal year.[7] Contracts demonstrate predictive power over Budget Authority at a fiscal year frequency only when either WWII or the Korean War is included in the sample. This predictability vanishes when the sample starts after the outbreak of the Korean War. A similar pattern is observed for the predictive power of Budget Authority on contracts, although in this case, the F-statistics are smaller, indicating weaker predictive ability. Overall, the results suggest that neither Budget Authority nor contracts consistently predict the other.

[7] The fiscal year started in July prior to 1977 and from October thereafter.

Table A2: Granger Causality Tests across Government Spending Measures

Predicted	*Predictor*	*Sample*	*Frequency*	F	*p value*
NIPA Defense Procurement Spending - Contracts					
NIPA Def. Proc. Spending	Military Prime Contracts	1947:1 - 2019:4	Cal. Quarter	10.571	0
NIPA Def. Proc. Spending	Military Prime Contracts	1955:1 - 2019:4	Cal. Quarter	9.758	0
Military Prime Contracts	NIPA Def. Proc. Spending	1947:1 - 2019:4	Cal. Quarter	1.308	.24
Military Prime Contracts	NIPA Def. Proc. Spending	1955:1 - 2019:4	Cal. Quarter	1.634	.115
Budget Authority - Contracts					
Budget Authority	Military Prime Contracts	1940 - 2019	Fiscal Year	33.123	0
Budget Authority	Military Prime Contracts	1947 - 2019	Fiscal Year	13.865	0
Budget Authority	Military Prime Contracts	1955 - 2019	Fiscal Year	1.452	.242
Military Prime Contracts	Budget Authority	1940 - 2019	Fiscal Year	5.508	.006
Military Prime Contracts	Budget Authority	1947 - 2019	Fiscal Year	4.632	.013
Military Prime Contracts	Budget Authority	1955 - 2019	Fiscal Year	1.482	.235
Budget Authority - NIPA Defense Spending					
NIPA Def. Spending	Budget Authority	1947 - 2019	Fiscal Year	61.364	0
NIPA Def. Spending	Budget Authority	1955 - 2019	Fiscal Year	52.407	0
Budget Authority	NIPA Def. Spending	1947 - 2019	Fiscal Year	.247	.782
Budget Authority	NIPA Def. Spending	1955 - 2019	Fiscal Year	.738	.482
Budget Authority - Defense Outlays					
Def. Outlays	Budget Authority	1940 - 2019	Fiscal Year	168.37	0
Def. Outlays	Budget Authority	1947 - 2019	Fiscal Year	9.815	0
Def. Outlays	Budget Authority	1955 - 2019	Fiscal Year	52.407	0
Budget Authority	Def. Outlays	1940 - 2019	Fiscal Year	9.917	0
Budget Authority	Def. Outlays	1947 - 2019	Fiscal Year	9.817	0
Budget Authority	Def. Outlays	1955 - 2019	Fiscal Year	.167	.847
Spending Authorizations - Defense News Shocks					
Spending Authorizations	Defense News Shocks	1940:1 - 2019:4	Cal. Quarter	14.319	0
Spending Authorizations	Defense News Shocks	1947:1 - 2020:4	Cal. Quarter	4.683	0
Spending Authorizations	Defense News Shocks	1955:1 - 2020	Fiscal Year	4.843	0
Defense News Shocks	Spending Authorizations	1940:1 - 2020:4	Cal. Quarter	7.705	0
Defense News Shocks	Spending Authorizations	1947:1 - 2020:4	Cal. Quarter	2.988	.003
Defense News Shocks	Spending Authorizations	1955:1 - 2020:4	Cal. Quarter	2.745	.006

Notes: Granger causality tests are conducted using Stata's `vargranger` post-estimation command from the `var` command, which estimates a bivariate VAR model for the predicted variable and the predictor with `p` lags. The VAR configuration employs the `dfk` and `small` options of the `var` command to adjust the F statistics for small-sample bias. The testing procedure involves (i) running an OLS regression of the predicted variable on `p` lags of itself along with `p` lags of the predictor, and (ii) conducting a Wald test to assess the null hypothesis that the `p` lags of the predictor are jointly non-significant. The lag `p` is set to *four* for quarterly data and to *two* for annual data. Variables are in real per capita values. Price deflator is the GDP price deflator (2012=100).

In the third panel of Table A2, we conduct a similar analysis exploring the predictability of Budget Authority versus NIPA's measure of defense spending, both aggregated at the fiscal year frequency. This analysis considers the sample with and without the outbreak of the Korean War, due to the limited availability of NIPA data on defense spending. The results are clear: Budget Authority Granger-causes NIPA defense spending, but not vice versa. This finding supports our narrative that Budget Authority predates NIPA's measure of defense spending.

In the fourth panel, we assess the predictability between Budget Authority and Defense Outlays. Across all samples, Budget Authority Granger causes Defense Outlays. Conversely, while Outlays seem to predict Budget Authority, this result (i) exhibits much smaller F statistics and (ii) lacks robustness when the outbreak of the Korean War is excluded from the analysis. Overall, the results clearly indicate that Budget Authority predicts Outlays, whereas the opposite relationship remains unclear. This is consistent with the understanding that outlays follow the authorization to spend in the government spending process.

Lastly, we check the predictability between defense news shocks (Ramey and Zubairy (2018)) and spending authorizations. Even if F statistics appear to be larger when defense news shock is the predictor, both news and spending authorizations appear to have statistical significant predicting power on each other. This result suggests that the timing of the two variables is similar.

C Stylized Model - Proof of Theorem

Stylized Model. Recall that the stylized model in the paper was represented by three equations:

$$B_t = \rho \cdot B_{t-1} + \eta_t, \quad \eta_t \sim_{iid} WN(0, \sigma_B^2) \tag{2}$$

$$G_t = (1-\lambda) \cdot B_t + \lambda \cdot B_{t-1}, \quad \lambda \in [0,1] \tag{3}$$

$$Y_t = \phi \cdot Y_{t-1} + \gamma \cdot B_t + \varepsilon_t, \quad \phi \in (0,1),\ \gamma > 0,\ \varepsilon_t \sim_{\text{iid}} WN(0, \sigma_y^2), \tag{4}$$

Moreover, by combining equations (2) and (3), we can rewrite NIPA G as an ARMA(1,1) process:

$$G_t = \rho \cdot G_{t-1} + \underbrace{(1-\lambda) \cdot \eta_t + \lambda \cdot \eta_{t-1}}_{:=\xi_t}, \tag{5}$$

Mispecified Model. The misspecified model for output can also be written as follows:

$$Y_t = \tilde{\phi} \cdot Y_{t-1} + \tilde{\gamma} \cdot G_t + u_t, \tag{6}$$

where u_t is an error term orthogonal to lagged output and government spending, and the misspecified parameters (denoted by tilde) are functions of the true underlying parameters of the model (i.e., $\lambda, \phi, \gamma, \rho, \sigma_b$, and σ_y).

Under the mis-specified model we estimate Equation (6). The OLS estimators converge to the tilde-parameters, $\tilde{\phi}$ and $\tilde{\gamma}$, which in turn are a function of the underlying parameters' set of the model (i.e. equations (2), (5) and (4)): $\phi, \gamma, \rho, \lambda, \sigma_\eta, \sigma_\varepsilon$.

In order to compare the multiplier estimated using budget authority, $\mathcal{M}(H)$ with the one estimated using NIPA, $\tilde{\mathcal{M}}(H)$, we need to find the exact analytical expressions for these values.

Characterize the Bias. In order to find the asymptotic values of the tilde-parameters, let's start from the DGP equation linking NIPA to authorizations and re-write it in terms of current authorizations as a function of past authorizations and NIPA:

$$B_t = \frac{1}{1-\lambda} \cdot G_t - \frac{\lambda}{1-\lambda} \cdot B_{t-1}$$

Now substitute this expression into the equation for output (Equation (4)):

$$Y_t = \phi \cdot \underbrace{Y_{t-1}}_{f(\eta_{t-1},\eta_{t-2},\ldots)} + \gamma \cdot \frac{1}{1-\lambda} \cdot \underbrace{G_t}_{g(\eta_{t-1},\eta_{t-2},\ldots)} - \gamma \cdot \frac{\lambda}{1-\lambda} \cdot \underbrace{B_{t-1}}_{h(\eta_{t-1},\eta_{t-2},\ldots)} + \varepsilon_t$$

As shown, lagged budget authority, B_{t-1}, behaves as an omitted variable when we only use NIPA to estimate the effects of government spending with Equation (6). This occurs because past shocks η_{t-1} affect both lagged output, Y_{t-1} and current values of NIPA, G_t which record them with delay.

We can characterize the bias using a control function approach. In other words, we need to orthogonalize the omitted variable with respect to the variables included in the mis-specified model:

$$B_{t-1} = \theta_1 \cdot Y_{t-1} + \theta_2 \cdot G_t + e_t \tag{7}$$

where θ_1 and θ_2 are the asymptotic values of an OLS regression of B_{t-1} on Y_{t-1} and G_t respectively while e_t is an orthogonal error term.

At this point, replace equation (7) into the previous expression to get:

$$Y_t = \underbrace{\left(\phi - \underbrace{\gamma \cdot \frac{\lambda}{1-\lambda} \cdot \theta_1}_{\text{Bias}}\right)}_{:=\tilde{\phi}} \cdot Y_{t-1} + \underbrace{\left(\gamma \cdot \underbrace{\frac{1-\lambda \cdot \theta_2}{1-\lambda}}_{\text{Bias}}\right)}_{:=\tilde{\gamma}} \cdot G_t + \underbrace{\gamma \cdot e_t + \varepsilon_t}_{\text{New orthogonal error term}}$$

First of all, notice the following important remark.

Remark. *The bias in the mis-specified parameters vanishes as $\lambda \to 0$ and $G_t \approx B_t$.*

In other words, in absence of measurement delays, authorizations and NIPA are identical and the approaches estimate the same structural parameters.

At this point, we need to characterize the analytical expressions of θ_1 and θ_2. We can do this using the Frisch-Waugh-Lovel theorem (FWL), or partitioned regression theorem.

Characterization of θ_1 and θ_2. Suppose we estimate the mis-specified model, Equation (6), via OLS. By the FWL Theorem, we have that:

$$\hat{\theta}_1^{\text{OLS}} \xrightarrow{p} \theta_1 = \frac{\mathbb{C}ov\left(B_{t-1} - \dfrac{\mathbb{C}ov\left(B_{t-1}, G_t\right)}{\mathbb{V}ar(G_t)} \cdot G_t \ , \ Y_{t-1} - \dfrac{\mathbb{C}ov\left(Y_{t-1}, G_t\right)}{\mathbb{V}ar(G_t)} \cdot G_t\right)}{\mathbb{V}ar\left(Y_{t-1} - \dfrac{\mathbb{C}ov\left(Y_{t-1}, G_t\right)}{\mathbb{V}ar(G_t)} \cdot G_t\right)}.$$

Similarly, for θ_2 we have:

$$\hat{\theta}_2^{\text{OLS}} \xrightarrow{p} \theta_2 = \frac{\mathbb{C}ov\left(B_{t-1} - \dfrac{\mathbb{C}ov\left(B_{t-1}, Y_{t-1}\right)}{\mathbb{V}ar(Y_{t-1})} \cdot Y_{t-1} \ , \ G_t - \dfrac{\mathbb{C}ov\left(G_t, Y_{t-1}\right)}{\mathbb{V}ar(Y_{t-1})} \cdot Y_{t-1}\right)}{\mathbb{V}ar\left(G_t - \dfrac{\mathbb{C}ov\left(G_t, Y_{t-1}\right)}{\mathbb{V}ar(G_t)} \cdot Y_{t-1}\right)}.$$

Proposition 1. *The population value of θ_1 can be simplified as follows:*

$$\theta_1 = \frac{\mathbb{V}ar(G_t) \cdot \mathbb{C}ov(B_{t-1}, Y_{t-1}) - \mathbb{C}ov(B_{t-1}, G_t) \cdot \mathbb{C}ov(G_t, Y_{t-1})}{\mathbb{V}ar(Y_t) \cdot \mathbb{V}ar(G_t) - \mathbb{C}ov(Y_{t-1}, G_t)^2}$$

Proof. Let's start by simplifying the denominator:

$$\begin{aligned}
\text{Denominator} &= \mathbb{V}ar\left(Y_{t-1} - \frac{\mathbb{C}ov\left(Y_{t-1}, G_t\right)}{\mathbb{V}ar(G_t)} \cdot G_t\right) \\
&= \mathbb{V}ar(Y_{t-1}) + \left(\frac{\mathbb{C}ov\left(Y_{t-1}, G_t\right)}{\mathbb{V}ar(G_t)}\right)^2 \cdot \mathbb{V}ar(G_t) - 2\frac{\mathbb{C}ov\left(Y_{t-1}, G_t\right)}{\mathbb{V}ar(G_t)} \cdot \mathbb{C}ov(Y_{t-1}, G_t) \\
&= \frac{1}{\mathbb{V}ar(G_t)} \cdot \left(\mathbb{V}ar(Y_t) \cdot \mathbb{V}ar(G_t) - \mathbb{C}ov(Y_{t-1}, G_t)^2\right)
\end{aligned}$$

Concerning the numerator:

$$\begin{aligned}
\text{Numerator} &= \mathbb{C}ov\left(B_{t-1} - \frac{\mathbb{C}ov\left(B_{t-1}, G_t\right)}{\mathbb{V}ar(G_t)} \cdot G_t\ ,\ Y_{t-1} - \frac{\mathbb{C}ov\left(Y_{t-1}, G_t\right)}{\mathbb{V}ar(G_t)} \cdot G_t\right) \\
&= \mathbb{C}ov(B_{t-1}, Y_{t-1}) - \frac{\mathbb{C}ov\left(Y_{t-1}, G_t\right)}{\mathbb{V}ar(G_t)} \cdot \mathbb{C}ov(B_{t-1}, G_t) + ... \\
&\quad ... + \frac{\mathbb{C}ov\left(B_{t-1}, G_t\right)}{\mathbb{V}ar(G_t)} \cdot \frac{\mathbb{C}ov\left(Y_{t-1}, G_t\right)}{\mathbb{V}ar(G_t)} \cdot \mathbb{C}ov(G_t, G_t) - \frac{\mathbb{C}ov\left(B_{t-1}, G_t\right)}{\mathbb{V}ar(G_t)} \cdot \mathbb{C}ov(G_t, Y_{t-1}) \\
&= \frac{1}{\mathbb{V}ar(G_t)} \cdot \left(\mathbb{C}ov(B_{t-1}, Y_{t-1}) \cdot \mathbb{V}(G_t) - \mathbb{C}ov(B_{t-1}, G_t) \cdot \mathbb{C}ov(G_t, Y_{t-1})\right)
\end{aligned}$$

By dividing the two terms we obtain the result.

■

By symmetry, we can derive the population value of θ_2.

Proposition 2. *The population value of θ_2 can be simplified as follows:*

$$\theta_2 = \frac{\mathbb{V}ar(Y_{t-1}) \cdot \mathbb{C}ov(B_{t-1}, G_t) - \mathbb{C}ov(B_{t-1}, Y_{t-1}) \cdot \mathbb{C}ov(Y_{t-1}, G_t)}{\mathbb{V}ar(Y_t) \cdot \mathbb{V}ar(G_t) - \mathbb{C}ov(Y_{t-1}, G_t)^2}$$

Relevant Population Moments *At this point we need to characterize all the population moments reported above in terms of structural parameters.*

Proposition 3. *The variance of output σ_y^2 is equal to:*

$$\sigma_y^2 := \mathbb{V}ar(Y_t) = \frac{1}{1-\phi^2} \cdot \left(\sigma_\varepsilon^2 + \gamma^2 \cdot \frac{\sigma_\eta^2}{1-\rho^2}\right)$$

Proof. From Equation (4) (output) and Equation (2) (authorizations) we have:

$$Y_t = (1 - \phi\, L)^{-1} \cdot \gamma \cdot B_t + (1 - \phi\, L)^{-1} \cdot \varepsilon_t$$

$$Y_t = \gamma \cdot (1 - \phi\, L)^{-1} \cdot (1 - \rho\, L)^{-1} \cdot \eta_t + (1 - \phi\, L)^{-1} \cdot \varepsilon_t$$

A standard result in the lag-operator theory is:[8]

$$(1-\phi\, L)^{-1}\cdot(1-\rho\, L)^{-1}=\sum_{j=0}^{\infty}\frac{\phi^{j+1}-\rho^{j+1}}{\phi-\rho}\cdot L^j.$$

This can be seen by recursively plugging lagged values of output and authorizations into the equation of output:

$$\begin{aligned}
Y_t &= \phi\cdot Y_{t-1}+\gamma\cdot B_t+\varepsilon_t\\
&= \phi^2\cdot Y_{t-2}+\gamma\cdot(\phi+\rho)\cdot B_{t-1}+\varepsilon_t+\phi\cdot\varepsilon_{t-1}+\gamma\cdot\eta_t\\
&= \phi^3\cdot Y_{t-3}+\gamma\cdot\left(\phi^2+\phi\cdot\rho+\rho^2\right)\cdot B_{t-1}+\varepsilon_t+\phi\cdot\varepsilon_{t-1}+\phi^2\cdot\varepsilon_{t-2}+\gamma\cdot(\eta_t+(\phi+\rho)\cdot\eta_{t-1})\\
&[...]\\
&= \phi^h\cdot Y_{t-h}+\gamma\cdot\left(\sum_{j=0}^{h-1}\phi^{h-1-j}\cdot\rho^j\right)\cdot B_{t-h+1}+\sum_{j=0}^{h-1}\phi^j\cdot\varepsilon_{t-j}+\gamma\sum_{n=0}^{h-2}\left(\sum_{j=0}^{n}\phi^{n-j}\cdot\rho^j\right)\cdot\eta_{t-n}\\
&= \phi^h\cdot Y_{t-h}+\gamma\cdot\frac{\phi^h-\rho^h}{\phi-\rho}\cdot B_{t-(h-1)}+\sum_{j=0}^{h-1}\phi^j\cdot\varepsilon_{t-j}+\gamma\sum_{n=0}^{h-2}\frac{\phi^{n+1}-\rho^{n+1}}{\phi-\rho}\cdot\eta_{t-n}
\end{aligned}$$

Therefore, the moving average representation of output is equal to:

$$Y_t=\sum_{j=0}^{\infty}\phi^j\cdot\varepsilon_{t-j}+\gamma\cdot\sum_{j=0}^{\infty}\frac{\phi^{j+1}-\rho^{j+1}}{\phi-\rho}\cdot\eta_{t-j}$$

Since $\eta_t\perp\varepsilon_t$ for all t and they are both serially uncorrelated, we have that:

$$\mathbb{V}ar(Y_t)=\frac{\sigma_\varepsilon}{1-\phi^2}+\gamma^2\cdot\left(\frac{\phi^2}{(\phi-\rho)^2}\cdot\frac{\sigma_\eta^2}{1-\phi^2}+\frac{\rho^2}{(\phi-\rho)^2}\cdot\frac{\sigma_\eta^2}{1-\rho^2}-\frac{2\cdot\phi\cdot\rho}{(\phi-\rho)^2}\cdot\frac{\sigma_\eta^2}{1-\phi\rho}\right)$$

which can be simplified as:

$$\sigma_y^2:=\mathbb{V}ar(Y_t)=\frac{1}{1-\phi^2}\cdot\left(\sigma_\varepsilon^2+\gamma^2\cdot\frac{(1+\phi\rho)\cdot\sigma_\eta^2}{(1-\rho^2)(1-\phi\rho)}\right)$$

[8] Provided that $|\phi|,|\rho|<1$. See Hamilton (1994).

∎

Proposition 4. *The variance of NIPA is given by:*

$$\mathbb{V}ar(G_t) = (1 - 2\lambda \cdot (1-\lambda) \cdot (1-\rho)) \cdot \mathbb{V}ar(B_t), \quad \mathbb{V}ar(B_t) = \frac{\sigma_\eta^2}{1-\rho^2}$$

Proof. The variance of authorizations follows from the variance of an AR(1) process.

Now, using the equation for NIPA, Equation (5), we have that its moving average representation is equal to:

$$\begin{aligned}
G_t &= \rho \cdot G_{t-1} + (1-\lambda) \cdot \eta_t + \lambda \cdot \eta_{t-1} \\
&= \rho^2 \cdot G_{t-2} + (1-\lambda) \cdot \eta_t + (\lambda + \rho \cdot (1-\lambda)) \cdot \eta_{t-1} + \rho \cdot \lambda \cdot \eta_{t-2} \\
&[...] \\
&= \rho^h \cdot G_{t-h} + (1-\lambda) \cdot \eta_t + (\lambda + \rho \cdot (1-\lambda)) \cdot \sum_{j=0}^{h-2} \rho^j \cdot \eta_{t-j-1} + \lambda \cdot \rho^{h-1} \cdot \eta_{t-h} \\
&= (1-\lambda) \cdot \eta_t + (\lambda + \rho \cdot (1-\lambda)) \cdot \sum_{j=0}^{\infty} \rho^j \cdot \eta_{t-j-1}
\end{aligned}$$

Therefore the variance of NIPA is:

$$\begin{aligned}
\mathbb{V}ar(G_t) &= (1-\lambda)^2 \cdot \sigma_\eta^2 + (\lambda + \rho \cdot (1-\lambda))^2 \cdot \sum_{j=0}^{\infty} \rho^{2j} \cdot \sigma_\eta^2 \\
&= \left((1-\lambda)^2 \cdot (1-\rho^2) + (\lambda + \rho \cdot (1-\lambda))^2\right) \cdot \underbrace{\frac{\sigma_\eta^2}{1-\rho^2}}_{=\mathbb{V}ar(B_t)} \\
&= (1 - 2\lambda \cdot (1-\lambda) \cdot (1-\rho)) \cdot \mathbb{V}ar(B_t)
\end{aligned}$$

∎

Corollary 5. *The variance of authorizations is larger than the one of NIPA:*

$$\forall \lambda \in [0,1], \quad \mathbb{V}ar(G_t) \leq \mathbb{V}ar(B_t)$$

Proof. From the expression for the variance of NIPA:

$$\mathbb{V}ar(G_t) = \left(1 - \underbrace{2\lambda \cdot (1-\lambda) \cdot (1-\rho)}_{\in[0,\frac{1}{2}]}\right) \cdot \mathbb{V}ar(B_t) \leq \mathbb{V}ar(B_t)$$

■

Proposition 6. *The covariance between contemporaneous values of output and authorizations is given by:*

$$\mathbb{C}ov(Y_t, B_t) := \frac{\gamma \cdot \sigma_\eta^2}{(1-\rho^2)\cdot(1-\phi\cdot\rho)}$$

Proof. From Equation (4) (output) and Equation (2) (authorizations) we have:

$$\begin{aligned}
\mathbb{C}ov(Y_t, B_t) &= \mathbb{C}ov\left(\phi \cdot Y_{t-1} + \gamma \cdot B_t + \varepsilon_t\ ,\ B_t\right) \\
&= \gamma \cdot \mathbb{V}ar(B_t) + \phi \cdot \mathbb{C}ov(Y_{t-1}, B_t) \\
&= \gamma \cdot \frac{\sigma_\eta^2}{1-\rho^2} + \phi \cdot \mathbb{C}ov(Y_{t-1}, \rho \cdot B_{t-1} + \eta_t) \\
&= \gamma \cdot \frac{\sigma_\eta^2}{1-\rho^2} + \phi \cdot \rho \cdot \mathbb{C}ov(Y_{t-1}, B_{t-1}) \\
&= \frac{\gamma \cdot \sigma_\eta^2}{(1-\rho^2)\cdot(1-\phi\cdot\rho)} \\
&= \frac{\gamma}{1-\phi\cdot\rho} \cdot \mathbb{V}ar(B_t)
\end{aligned}$$

where we have used the fact that as long as $|\phi|, |\rho| < 1$, the process of Y_t is covariance stationary and $\mathbb{C}ov(Y_t, B_t) = \mathbb{C}ov(Y_{t-1}, B_{t-1})$. ■

Proposition 7. *The covariance between lagged authorizations and NIPA is given by:*

$$\mathbb{C}ov(B_{t-1}, G_t) = (\lambda + \rho \cdot (1 - \lambda)) \cdot \mathbb{V}ar(B_t).$$

Proof. Using the definition of NIPA, Equation (3), we have:

$$\begin{aligned}
\mathbb{C}ov(B_{t-1}, G_t) &= \mathbb{C}ov(B_{t-1}, (1 - \lambda) \cdot B_t + \lambda \cdot B_{t-1}) \\
&= (1 - \lambda) \cdot \mathbb{C}ov(B_{t-1}, \underbrace{\rho \cdot B_{t-1} + \eta_t}_{=B_t}) + \lambda \cdot \mathbb{V}ar(B_t) \\
&= (\lambda + \rho \cdot (1 - \lambda)) \cdot \mathbb{V}ar(B_t)
\end{aligned}$$

■

Proposition 8. *The covariance between lagged output and NIPA is given by:*

$$\mathbb{C}ov(Y_{t-1}, G_t) = \frac{\gamma}{1 - \phi \cdot \rho} \cdot \mathbb{C}ov(B_{t-1}, G_t)$$

where:

$$\mathbb{C}ov(B_{t-1}, G_t) = (\lambda + \rho \cdot (1 - \lambda)) \cdot \mathbb{V}ar(B_t)$$

has been derived in the previous proposition.

Proof. Using the expression for output, Equation (4) we have:

$$\begin{aligned}
\mathbb{C}ov(Y_{t-1}, G_t) &= \mathbb{C}ov(\phi \cdot Y_{t-2} + \gamma \cdot B_{t-1} + \varepsilon_{t-1}, G_t) \quad \text{Using Equation (5)} \\
&= \phi \cdot \mathbb{C}ov(Y_{t-2}, \underbrace{\rho \cdot G_{t-1} + (1 - \lambda) \cdot \eta_t + \lambda \cdot \eta_{t-1}}_{=G_t}) + \gamma \cdot \mathbb{C}ov(B_{t-1}, G_t) \\
&= \phi \cdot \rho \cdot \mathbb{C}ov(Y_{t-2}, G_{t-1}) + \gamma \cdot \mathbb{C}ov(B_{t-1}, G_t) \\
&= \frac{\gamma}{1 - \phi \cdot \rho} \cdot \mathbb{C}ov(B_{t-1}, G_t).
\end{aligned}$$

■

At this point we have all the required information to prove the actual main theorem.

Theorem 1. *Let the data be generated by Equations* (2), (3) *and* (4). *If* $\lambda > \frac{1}{2}$, *i.e. there are severe measurement delays in NIPA, the (large sample) fiscal multiplier obtained with authorizations is larger than the one obtained with NIPA:*

$$\mathcal{M}(H) > \tilde{\mathcal{M}}(H) \quad \forall H > 0.$$

Proof. First, recall the formulas of the multipliers estimated with the two methods. First, if we identify fiscal shocks using authorizations, η_t, we have:

$$\frac{\partial G_{t+h}}{\partial \eta_t} = \begin{cases} 1-\lambda & \text{if } h = 0, \\ \rho^{h-1} \cdot (\lambda + \rho \cdot (1-\lambda)) & \text{if } h > 0. \end{cases}$$

As long as $\lambda > \phi$, the multiplier is decreasing from impact, $H = 0$, to the long-run: $H \to \infty$.

On the contrary, if we identify the fiscal shocks using NIPA, ξ_t, the multiplier is equal to:

$$\tilde{\mathcal{M}}(H) := \frac{\sum_{h=0}^{H} \frac{\partial Y_{t+h}}{\partial \xi_t}}{\sum_{h=0}^{H} \frac{\partial G_{t+h}}{\partial \xi_t}} = \begin{cases} \tilde{\gamma} & \text{if } H = 0, \\ \frac{\tilde{\gamma} \cdot \frac{\rho \cdot (1-\tilde{\phi}) \cdot (1-\rho^{H+1}) - \tilde{\phi} \cdot (1-\rho) \cdot (1-\tilde{\phi}^{H+1})}{(1-\tilde{\phi}) \cdot (1-\rho) \cdot (\rho - \tilde{\phi})}}{\frac{1-\rho^{H+1}}{1-\rho}} & \text{if } H > 0, \\ \frac{\tilde{\gamma}}{1-\tilde{\phi}} & \text{if } H \to \infty. \end{cases} \tag{8}$$

which is an increasing function in H, moving from an impact multiplier of $\tilde{\gamma}$ to a long-run multiplier of $\frac{\tilde{\gamma}}{1-\tilde{\phi}}$. When there is no delay, $\lambda = 0$, the multiplier is increasing because the ratio between two consecutive values is larger than one:

$$\frac{\tilde{\mathcal{M}}(H+1)}{\tilde{\mathcal{M}}(H)} = 1 + \underbrace{\frac{\tilde{\phi} \cdot \left((1-\rho+\tilde{\phi}) \cdot \rho^{H+2} - \phi^{H+2} \right)}{\rho \cdot \frac{1-\rho^{H+1}}{1-\rho} - \tilde{\phi} \cdot \frac{1-\tilde{\phi}^{H+1}}{1-\tilde{\phi}}} \cdot \frac{1-\rho^{H+1}}{1-\rho^{H+2}}}_{>0 \quad \forall \rho, \tilde{\phi} \in (0,1)} > 1$$

Finally, recall that the tilde-parameters, $\tilde{\phi}$ and $\tilde{\gamma}$ are the coefficients of a regression of output on NIPA and lagged output (mis-specified model), Equation (6).

The proof is made of two parts. First, we show that the impact multiplier is larger when using authorizations: $\mathcal{M}(0) > \tilde{\mathcal{M}}(0)$. Second, we show that the long-run multiplier is larger when using authorizations: $\mathcal{M}(\infty) > \tilde{\mathcal{M}}(\infty)$. Third, since both multipliers are monotonic functions, multipliers obtained with authorizations will be larger than multipliers obtained using NIPA at any horizon H.

Impact Multipliers. The (large sample) impact multiplier obtained with authorizations is $\gamma/(1-\lambda)$, while the one obtained with NIPA is $\tilde{\gamma}$. We have showed that the latter term is equal to:

$$\tilde{\gamma} = \frac{\gamma \cdot (1 - \lambda \cdot \theta_2)}{1 - \lambda}$$

Therefore, we have:

$$\begin{aligned}
&\mathcal{M}(0) > \tilde{\mathcal{M}}(0) \\
&\frac{\gamma}{1-\lambda} > \frac{\gamma \cdot (1 - \lambda \cdot \theta_2)}{1 - \lambda} \\
&\rightarrow \theta_2 > 0.
\end{aligned}$$

Therefore, we need to show that $\theta_2 > 0$. From Proposition 2, we have that:

$$\theta_2 = \frac{\mathbb{V}ar(Y_{t-1}) \cdot \mathbb{C}ov(B_{t-1}, G_t) - \mathbb{C}ov(B_{t-1}, Y_{t-1}) \cdot \mathbb{C}ov(Y_{t-1}, G_t)}{\mathbb{V}ar(Y_t) \cdot \mathbb{V}ar(G_t) - \mathbb{C}ov(Y_{t-1}, G_t)^2}$$

Therefore, $\theta_2 > 0$ if and only if:

$$\mathbb{V}ar(Y_{t-1}) \cdot \mathbb{C}ov(B_{t-1}, G_t) > \mathbb{C}ov(B_{t-1}, Y_{t-1}) \cdot \mathbb{C}ov(Y_{t-1}, G_t).$$

We can use the the formulas derived earlier for these population moments to verify whether the condition is met.

$$\underbrace{\frac{1}{1-\phi^2}\cdot\left(\sigma_\varepsilon^2+\gamma^2\cdot\frac{(1+\phi\rho)\cdot\sigma_\eta^2}{(1-\rho^2)(1-\phi\rho)}\right)}_{\mathbb{V}ar(Y_t)}\cdot\mathbb{C}ov(B_{t-1},G_t) > \underbrace{\frac{\gamma}{1-\phi\rho}\cdot\mathbb{V}ar(B_t)}_{\mathbb{C}ov(B_t,Y_t)}\cdot\underbrace{\frac{\gamma}{1-\phi\rho}\cdot\mathbb{C}ov(B_{t-1},G_t)}_{=\mathbb{C}ov(Y_{t-1},G_t)}$$

$$\frac{\sigma_\varepsilon^2}{1-\phi^2}+\gamma^2\cdot\frac{(1+\phi\rho)}{1-\phi\rho}\cdot\underbrace{\frac{\sigma_\eta^2}{1-\rho^2}}_{=\mathbb{V}ar(B_t)} > \frac{\gamma^2}{(1-\phi\rho)^2}\cdot\mathbb{V}ar(B_t)\quad\text{since } \mathbb{C}ov(B_{t-1},G_t)>0$$

$$\frac{\sigma_\varepsilon^2}{1-\phi^2}+\gamma^2\cdot\frac{(1+\phi\rho)}{1-\phi\rho}\cdot\frac{\phi^2(1-\rho^2)}{1-\phi^2}\cdot\mathbb{V}ar(B_t)>0$$

The above expression always holds true as long as ϕ and ρ are between 0 and 1. Therefore, impact multipliers obtained with NIPA are lower than those ones obtained from authorizations (large sample estimates).

Long-run Multipliers. The last part of the proof consists in showing that the long-run multipliers estimated with NIPA is smaller than the one obtained with authorizations (large sample estimates). For this purpose, recall that the population value of the autoregressive coefficient of output in the mis-specified model is equal to:

$$\tilde{\phi}=\phi-\gamma\cdot\frac{\lambda}{1-\lambda}\cdot\theta_1.$$

Therefore, we have:

$$\begin{aligned}
\mathcal{M}(\infty) &> \tilde{\mathcal{M}}(0)\\
\frac{\gamma}{1-\phi} &> \frac{\tilde{\gamma}}{1-\tilde{\phi}}\\
\frac{\gamma}{1-\phi} &> \frac{\frac{\gamma\cdot(1-\lambda\cdot\theta_2)}{1-\lambda}}{1-\left(\phi-\gamma\cdot\frac{\lambda}{1-\lambda}\cdot\theta_1\right)}\\
\frac{1}{1-\phi} &> \frac{1-\lambda\theta_2}{(1-\lambda)(1-\phi)+\gamma\lambda\theta_1}\\
\rightarrow \gamma\cdot\theta_1+(1-\phi)\cdot(\theta_2-1) &> 0.
\end{aligned}$$

We need to show that the above condition is always true.

At this point, it becomes necessary to derive the values of the numerator of θ_1 in terms of structural parameters:

$$\theta_1 = \frac{\mathbb{V}ar(G_t) \cdot \mathbb{C}ov(B_{t-1}, Y_{t-1}) - \mathbb{C}ov(B_{t-1}, G_t) \cdot \mathbb{C}ov(G_t, Y_{t-1})}{\mathbb{V}ar(Y_t) \cdot \mathbb{V}ar(G_t) - \mathbb{C}ov(Y_{t-1}, G_t)^2}$$

$$\rightarrow \text{Numerator of } \theta_1 = \mathbb{V}ar(G_t) \cdot \mathbb{C}ov(B_t, Y_t) - \mathbb{C}ov(B_{t-1}, G_t) \cdot \underbrace{(\lambda + (1-\lambda) \cdot \rho) \cdot \mathbb{C}ov(B_t, Y_t)}_{\mathbb{C}ov(G_t, Y_{t-1})}$$

$$\text{Numerator of } \theta_1 = \mathbb{C}ov(B_t, Y_t) \cdot \left(\mathbb{V}(G_t) - (\lambda + (1-\lambda) \cdot \rho) \cdot \underbrace{(\lambda + (1-\lambda) \cdot \rho) \cdot \mathbb{V}ar(B_t)}_{\mathbb{C}ov(B_{t-1}, G_t)} \right)$$

$$\text{Numerator of } \theta_1 = \mathbb{C}ov(B_t, Y_t) \cdot \left[\underbrace{1 - 2\lambda(1-\lambda)(1-\rho) - (\lambda + (1-\lambda) \cdot \rho)^2}_{=(1-\lambda)^2 \cdot (1-\rho^2)} \right] \cdot \mathbb{V}(B_t)$$

$$\text{Numerator of } \theta_1 = \mathbb{C}ov(B_t, Y_t) \cdot (1-\lambda)^2 \cdot (1-\rho^2) \cdot \mathbb{V}(B_t)$$

Therefore, the condition we need to verify becomes:

$$\gamma \cdot \theta_1 + (1-\phi) \cdot (\theta_2 - 1) > 0$$

$$\gamma \cdot \mathbb{C}ov(B_t, Y_t) \cdot (1-\lambda)^2 \cdot (1-\rho^2) \cdot \mathbb{V}(B_t) + (1-\phi)(\theta_2 - 1) \cdot (\text{Denominator of } \theta_1) > 0,$$

where the sign of the inequality is not affected by dividing both sides by the denominator of θ_1. In fact, the denominator of the (large sample) OLS coefficients represents the determinant of a variance-covariance matrix of the sum of the two regressors, Y_{t-1} and G_t:

$$\begin{aligned}
(\text{Denominator of } \theta_1) &= \mathbb{V}ar(Y_t) \cdot \mathbb{V}ar(G_t) - \mathbb{C}ov(Y_{t-1}, G_t)^2 \\
&= det\left(\begin{bmatrix} \mathbb{V}ar(Y_t) & \mathbb{C}ov(Y_{t-1}, G_t) \\ \mathbb{C}ov(Y_{t-1}, G_t) & \mathbb{V}ar(G_t) \end{bmatrix} \right) \\
&= |\Sigma(Y_{t-1}, B_t)|
\end{aligned}$$

which is, by definition, positive since variance-covariance matrices are positive definite matrices as long as regressors are not collinear.

Since the denominator of θ_1 and θ_2 are the same, we have that:

$$
\begin{aligned}
&(\theta_2 - 1) \cdot (\text{Denominator of } \theta_1) \\
&= (\text{Numerator of } \theta_2) - (\text{Denominator of } \theta_2) \\
&= \mathbb{V}ar(Y_{t-1}) \cdot \mathbb{C}ov(B_{t-1}, G_t) - \mathbb{C}ov(B_{t-1}, Y_{t-1}) \cdot \mathbb{C}ov(Y_{t-1}, G_t) - \big(\mathbb{V}ar(Y_t) \cdot \mathbb{V}ar(G_t) - \mathbb{C}ov(Y_{t-1}, G_t)^2\big) \\
&= \mathbb{V}ar(Y_t) \cdot \Big[\mathbb{C}ov(B_{t-1}, G_t) - \mathbb{V}ar(G_t)\Big] + \mathbb{C}ov(Y_{t-1}, G_t) \cdot \Big[\underbrace{\mathbb{C}ov(Y_{t-1}, G_t)}_{(\lambda+\rho(1-\lambda))\cdot\mathbb{C}ov(B_t,Y_t)} - \mathbb{C}ov(B_t, Y_t)\Big] \\
&= \mathbb{V}ar(Y_t)\Big[\lambda + \rho(1-\lambda) - (1 - 2\lambda(1-\lambda)(1-\rho))\Big]\mathbb{V}ar(B_t) - \mathbb{C}ov(Y_{t-1}, G_t)(1-\lambda)(1-\rho)\mathbb{C}ov(B_t, Y_t) \\
&= 2(1-\lambda)(1-\rho)(\lambda - \frac{1}{2})\mathbb{V}ar(Y_t)\mathbb{V}ar(B_t) - (1-\lambda)(1-\rho)(\lambda + \rho(1-\lambda))\mathbb{C}ov(B_t, Y_t)^2
\end{aligned}
$$

Therefore, the condition we have to prove becomes:

$$
\begin{aligned}
&\rightarrow \gamma \cdot \underbrace{\mathbb{C}ov(B_t, Y_t)(1-\lambda)^2 \cdot (1-\rho^2) \cdot \mathbb{V}(B_t)}_{\text{Numerator of } \theta_1} + ... \\
&... + (1-\phi)\Big[\underbrace{2(1-\lambda)(1-\rho)(\lambda - \frac{1}{2})\mathbb{V}ar(Y_t)\mathbb{V}ar(B_t) - (1-\lambda)(1-\rho)(\lambda + \rho(1-\lambda))\mathbb{C}ov(B_t, Y_t)^2}_{\text{Numerator of } \theta_2 - \text{Denominator of } \theta_2}\Big] > 0 \\
&\rightarrow (1-\lambda)(1+\rho)(1-\phi\rho)\mathbb{C}ov(B_t, Y_t)^2 + ... \\
&... + 2(1-\phi)(\lambda - \frac{1}{2})\mathbb{V}ar(Y_t)\mathbb{V}ar(B_t) - (1-\phi)(\lambda + \rho(1-\lambda))\mathbb{C}ov(B_t, Y_t)^2 > 0 \\
&\rightarrow \mathbb{C}ov(B_t, Y_t)^2\Big[\underbrace{(1-\lambda)(1+\rho)(1-\phi\rho) - (1-\phi)(\lambda + \rho(1-\lambda))}_{=1-2\lambda+\lambda\phi-\phi\rho^2+\lambda\phi\rho^2}\Big] + 2(1-\phi)(\lambda - \frac{1}{2})\mathbb{V}ar(Y_t)\mathbb{V}ar(B_t) > 0 \\
&\rightarrow \mathbb{C}ov(B_t, Y_t)^2\Big[\underbrace{1 - 2\lambda + \lambda\phi - \phi\rho^2 + \lambda\phi\rho^2}_{=\phi(1-\lambda)(1-\rho^2)-2(1-\phi)(\lambda-\frac{1}{2})}\Big] + 2(1-\phi)(\lambda - \frac{1}{2})\mathbb{V}ar(Y_t)\mathbb{V}ar(B_t) > 0 \\
&\rightarrow 2\underbrace{(1-\phi)}_{>0}(\lambda - \frac{1}{2})\Big[\underbrace{\mathbb{V}ar(Y_t)\mathbb{V}ar(B_t) - \mathbb{C}ov(B_t, Y_t)^2}_{=|\Sigma(B_t,Y_t)|>0}\Big] + \underbrace{\phi(1-\lambda)(1-\rho^2) \cdot \mathbb{C}ov(B_t, Y_t)^2}_{>0} > 0
\end{aligned}
$$

The inequality holds true if $\lambda \geq \frac{1}{2}$, proving the "*if-statement*" of the Theorem. ■

D Attenuation in Shocks to NIPA Government Spending

In the paper, we demonstrate that measures of NIPA spending lags behind different measures of defense authorizations and contracts. Researchers might be tempted to "correct" this timing problem by using leads of NIPA's measures in their regression equations to identify fiscal shocks, instead of using contracts or Budget Authority as instruments. However, it is important to highlight that contracts and Budget Authority differ from their NIPA counterparts not only in their timing but also in how they measure fiscal shocks. NIPA acts as a delayed moving average of past shocks to either contracts or Budget Authority. For instance, consider an aggregate increase in newly awarded military prime contracts: contracts immediately reflect the lumpy value of new awards, whereas NIPA progressively records these values over time as the purchased items are produced and delivered to the government. Therefore, using leads of NIPA to identify government might address the timing problem, but the so identified shocks would still result in attenuation bias, as the shocks are smaller and spread over time.

We compare the standard deviation of the cyclical components of real per capita values of Spending Authorizations, Budget Authority and Contracts with those ones of their NIPA analogues: NIPA defense spending and NIPA defense procurement spending. The cyclical component is derived by following Hamilton (2018)'s method of regressing a measure of government spending on its h-th lag, set at 5 years (20 quarters), reflecting the average duration of military build-ups, which often exceed the average two-year recession recovery period suggested by Hamilton (2018).

The top panel of Table A3 compares the standard deviations of the cyclical components of spending authorizations and NIPA defense spending. For the same reason we discussed in the Granger Causality test, we present results both including and excluding dates before 1951. Note that the standard deviation for spending authorization is 36.62% higher than that of NIPA when excluding the Korean War outbreak and 36.32% higher when including it.

The rest of the table shows results comparing (i) contracts and NIPA defense procurement spending (middle panel) and (ii) Budget Authority with NIPA defense spending aggregated by

Table A3: Spending Authorizations Exhibit More Cyclical Variation than NIPA Defense Spending

Measure of G	*Frequency*	*Sample*	*St. Dev. - Cyc. Component*	Δ%
Spending Authorizations Vs NIPA Defense Spending				
Spending Authorizations	Quarterly	1947:1-2016:4	3718.18	+ 36.32%
NIPA Defense Spending	Quarterly	1947:1-2016:4	2727.48	-
Spending Authorizations	Quarterly	1951:1-2016:4	3726.34	+ 36.62%
NIPA Defense Spending	Quarterly	1951:1-2016:4	2727.48	-
Contracts Vs NIPA Defense Procurement Spending				
Contracts	Quarterly	1947:1-2016:4	2588.25	+ 26.65%
NIPA Defense Procurement Spending	Quarterly	1947:1-2016:4	2043.57	-
Contracts	Quarterly	1951:1-2016:4	2589.84	+ 26.73%
NIPA Defense Procurement Spending	Quarterly	1951:1-2016:4	2043.57	-
Budget Authority Vs NIPA Defense Spending				
Budget Authority	Fiscal year	1947-2016	4127.99	+ 51.34%
NIPA Defense Spending	Fiscal year	1951-2016	2727.48	-
Budget Authority	Fiscal year	1951-2016	3768.7	+ 38.17%
NIPA Defense Spending	Fiscal year	1951-2016	2727.48	-

Notes: Values are in real per capita terms. The price deflator is the GDP price deflator. Cyclical components for quarterly variables are obtained as OLS residuals from a projection of four lags following the 5-year lag: `reg Y L(20/24).y` and `predict cyclical, resid`. Annual values are derived using `reg Y L5.y` and `predict cyclical, resid`.

fiscal year (bottom panel). The conclusion is identical, contracts and authorizations exhibit much higher standard deviations of their cyclical components.

To conclude, NIPA's measures of government spending are not only delayed relative to, and consistently predicted by, other measures, but they also attenuate fiscal shocks, highlighting the existence of a *measurement problem*.

E Ruling Out Anticipation of Military Contracts at Quarterly Frequencies

While it is possible that a single firm might have some foresight about being awarded a future large non-competed contract, aggregate variation in military contracts reflects unpredictable military events. In this section, we show that military contracts do not appear to be anticipated in aggregate data.

Stock Market Response to Contracts First, we construct an equally weighted portfolio of stock returns of four major defense contractors: Boeing, Lockheed, Raytheon, and Northrop.[9] We construct a time series of quarterly year-over-year returns in excess of returns on the S&P 500. Figure A29 shows the time series. We refer to this time series as "*Top4*".

We then use a VAR with the following variables:

$$\boldsymbol{X}_t := \begin{bmatrix} \text{Contracts}_t & \text{Top4}_t & \text{GDP}_t & G_t & TB3_t & \text{Tax}_t^{\text{R\&R2010}} \end{bmatrix}$$

where nominal variables—i.e., contracts, GDP, and NIPA's G—are in logs of real per capita values; $\text{Tax}_t^{\text{R\&R2010}}$ are the Romer & Romer (2010) exogenous tax shocks divided by lagged GDP. We estimate $\boldsymbol{X}_t = \boldsymbol{\Phi} \cdot \boldsymbol{X}_{t-1} + \boldsymbol{\varepsilon}_t$ and order contracts first, treating their innovations as structural shocks (i.e., Cholesky identification). Figure A30 shows the impulse response function of a shock of contracts to Top4. The sample goes from 1947:1 to 2007:4 in the left panel and from 1951:1 in the right panel.

Notice that a positive shock to contracts leads to a positive response of Top4, which peaks at quarter one from the shock. If shocks to contracts were anticipated by the stock market, we would not expect to see a delayed positive response of Top4 in response to the shock. The large dip following the initial positive response is due to a combination of (i) a mechanical reduction in

[9] This variable is similar in spirit to the excess cumulative returns variable constructed in Fisher and Peters (2010).

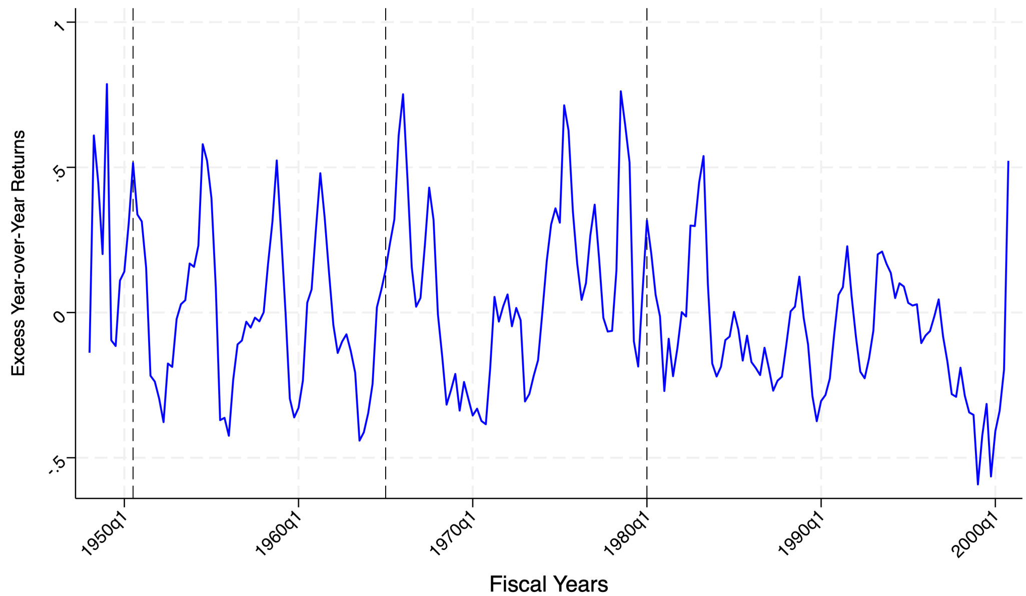


Figure A29: Quarterly Stock Returns of Top 4 Defense Contracts in Excess of the S&P 500 (Top4)

Notes: Top4 is constructed by taking an arithmetic average of the quarterly year-over-year stock market returns of Boeing, Raytheon, Northrop, and Lockheed. From the index, we then subtract the quarterly year-over-year stock market returns of the S&P 500. Dashed lines are Ramey and Shapiro (1998) war dates: 1950:3 (Korean War), 1965:1 (Vietnam War), 1980:1 (Carter-Reagan build-up).

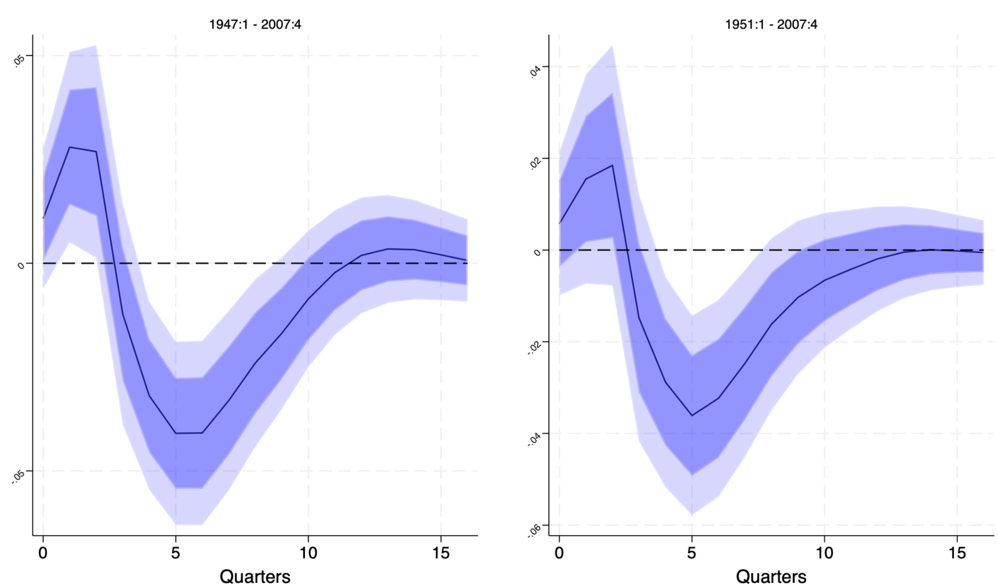


Figure A30: Top4 Stock Returns Respond to Contracts with a Delay

Notes: Confidence bands are 68% and 90%.

stock market returns after an initial period of high returns and (ii) the influence of a capital gain tax introduced during the Korean War (see Fisher and Peters (2010)). The right panel shows results if we exclude the Korean War from the sample: the initial positive response becomes more positive, and the ensuing dip becomes less negative, consistent with Fisher and Peters (2010)'s explanation.

Granger Causality Tests Second, we estimate a bivariate VAR using only real per capita values of military contracts and Top4 and then carry out a Granger Causality Test. Results are reported in Table A4.

Table A4: Contracts Granger Cause Top4 Stock Returns but not Vice-Versa

Predicted	*Predictor*	*Sample*	*Frequency*	F	*p value*
Contracts	Top4	1947:1 - 2000:4	Cal. Quarter	.707	.685
Contracts	Top4	1951:1 - 2000:4	Cal. Quarter	.526	.836
Top4	Contracts	1947:1 - 2000:4	Cal. Quarter	2.026	.045
Top4	Contracts	1951:1 - 2000:4	Cal. Quarter	1.677	.107

Notes: Granger causality tests are conducted using Stata's `vargranger` post-estimation command from the `var` command, which estimates a bivariate VAR model for the predicted variable and the predictor with `p` lags. The VAR configuration employs the `dfk` and `small` options of the `var` command to adjust the F statistics for small-sample bias. The testing procedure involves (i) running an OLS regression of the predicted variable on `p` lags of itself along with `p` lags of the predictor, and (ii) conducting a Wald test to assess the null hypothesis that the `p` lags of the predictor are jointly non-significant. The lag `p` is set to *eight*. Contracts are in real per capita values. The price deflator is the GDP price deflator (2012=100).

Notice that Top4 does not Granger Cause contracts (top-panel), but contracts appear to Granger Cause Top4 (bottom panel). Therefore, we conclude that there is no empirical evidence supporting that aggregate military contracts are anticipated; instead, our evidence suggests the opposite: in aggregate, contracts anticipate stock returns of major contractors.

F Comparisons to Other Approaches

F.1 How Does Our Approach Fit Into Recent Literature?

Spending authorizations primarily affect total G through their impact on defense procurement spending. Cox, Müller, et al. (2024) show that this result also holds for other widely employed fiscal shocks, such as defense news shocks (Ramey and Zubairy, 2018) and shocks identified by ordering NIPA G first in a recursive VAR (Blanchard and Perotti, 2002). This is because defense spending is the most volatile component of G, driven largely by military events. On the one hand, this is beneficial for identification, as defense spending is the most exogenous component of government spending. On the other hand, it raises questions about the external validity of the fiscal multiplier estimates.

For instance, defense procurement is heavily concentrated in manufacturing industries (Ramey and Shapiro, 1998; Perotti, 2007; Nekarda and Ramey, 2011; Cox, Müller, et al., 2024). It is plausible that government purchases from other non-manufacturing sectors could generate different input-output network effects (Bouakez, Rachedi, and Santoro, 2023). One potential solution to this issue is to use a multi-sector model calibrated to match the multiplier estimates obtained with spending authorizations when purchases are operated mainly via manufacturing, and then use the model to analyze the effects of government purchases in non-manufacturing sectors.

F.2 Spending Authorizations vs. Ramey's Defense News Shocks

Spending authorizations are not the first measure to predict both NIPA defense spending and Cholesky shocks derived from NIPA. Ramey (2011) presents a narrative measure of defense news shocks, which systematically anticipate changes to NIPA defense spending. While both spending authorizations and Ramey's defense news shocks anticipate changes in NIPA government spending, there are important conceptual differences between them.

Differences in Measurement of Expectations. The goal of Ramey's news shock measure is to capture broad shifts in expectations about future spending. Her measure includes shocks to expectations about spending several years in the future (discounted using present values), as well as multi-year shocks (e.g. if spending is now expected to increase by $20 billion next year and then stay $20 billion higher for the next three years, Ramey's measured shock will be the PDV of that $20 billion increase in each relevant year, summed across years). In that sense, her narrative measure is fundamentally forward-looking, as any broad-based measure of shocks to expectations must be.

In contrast, to the extent that spending authorizations capture changes in expectations, they do so very narrowly. spending authorizations will not capture changes in expectations about spending several years in the future unless those expectations are reflected in current authorizations. We can think of spending authorizations as a direct measure of spending shocks rather than a measure of expectations. Alternatively, spending authorizations can be understood as implicitly assuming adaptive expectations for future spending: that is, the best (or at least most certain) information about spending after the coming year is information about spending in the coming year.[10] In fact, considering our stylized model of Section **??**, it is possible to show that the expected future value of government spending at time t is equal to its current (discounted) value:

$$\mathbb{E}_t[G_{t+1}] = \theta \cdot G_t, \quad \text{with } \theta \in [\rho, 1]$$

where θ depends on the persistence of authorizations, ρ, and the extent of measurement delays, λ.[11]

[10] Adaptive expectations may be simple, but they are situationally relevant. In the federal budget process, future government spending is usually assumed to follow adaptive expectations. The previous year's spending levels are usually the starting point for negotiations over future bills. In recent years, it has often been politically easier to continue spending at current levels (through a "continuing resolution") for at least some period of time rather than negotiating a new spending bill.

[11] In particular, $\theta = (\lambda + (1-\lambda)\rho)(1-\lambda(1-\rho))/(1-2\lambda\rho(1-\rho)(1-\lambda))$.

Differences in the Treatment of Uncertainty. As a consequence of the differences in how Ramey's news shocks and spending authorizations treat changes in expectations, the two measures also make very different implicit assumptions about uncertainty. Expectations about future discretionary spending levels rarely contain a high degree of certainty because of the annual legislative (and political) process which determines spending. Moreover, changes to expectations about future spending often involve either additional uncertainty or the resolution of uncertainty. Any broad measure of changes in expectations must reflect some view of uncertainty, whether implicitly or explicitly.

Ramey's news shock measure does not take an explicit stance on policy uncertainty. Instead, she measures changes in *Business Week*'s best guesses about the path of future spending. However, some of the shocks in her series reflect much higher degrees of uncertainty than others. Some shocks in Ramey's series reflect changes that have already been incorporated into legislation (and therefore also show up in spending authorizations). Since the government has already committed to these changes in spending, those near-term shocks are very certain. News shocks also incorporate changes to expectations about spending several years in the future, such as beliefs about when Nazi Germany will fall (a key factor in several of Ramey's World War II shocks) and the length of American ground presence in Vietnam. In Ramey's defense news shocks, the only difference between these two types of shocks is discounting to reflect present values (generally small adjustments). The second type of shock involves much higher levels of uncertainty than the first, but they are treated the same. Implicitly, this reflects a view (consistent with many models) that households and firms are forward-looking and optimize based on the best information they have at any time, ignoring uncertainty.

Spending authorizations implicitly handle uncertainty quite differently. Because they measure the government's commitment to spend (via budget authority and contract awards), spending authorizations measure changes to spending more or less at the moment at which uncertainty is resolved. Rather than treating all shocks to expected future spending identically regardless of associated uncertainty, spending authorizations essentially ignore non-adaptive shocks to expectations unless and until uncertainty has been resolved (but incorporate all realized shocks into adaptive expectations of future spending). In that sense, spending authorizations are far more discriminating about the credibility of shocks to expectations.

In addition to the conceptual differences between news shocks and spending authorizations, the two measures also have somewhat different uses. In many circumstances, the choice of which measure to use—Ramey's defense news shocks or spending authorizations—depends on what assumptions researchers want to make about agents' behavior. Ramey's defense news shocks assume that households and firms are forward-looking and make decisions based on the best information they have about the future, ignoring uncertainty. In contrast, spending authorizations reflect an assumption that households and firms care about certainty and put off decisions until uncertainty is resolved.

Timing Relationship. Given the previous definition, one might expect defense news shocks to predate spending authorizations. However, the timing of shocks associated with major wars is quite similar under both measures, and the maximum value of the lead-lag correlation between the two series is only 0.40, at $h = 0$ (see Figure A31).

Second, we show in Online Appendix B.5 that no clear winner emerges from running Granger causality tests between defense news shocks and spending authorizations: both series appear to predict each other, no matter the sample used.[12] Yet while the timing of the two series is strongly coincident, they do not measure the same concept of shocks.

Most importantly, using a recursive VAR approach with defense news shocks ordered first, ahead of shocks derived from NIPA defense spending—as in Auerbach and Gorodnichenko (2012)—does *not* solve the measurement problem stemming from the delayed timing of NIPA. While spending authorizations can be understood as purely correcting for the timing of government purchases in NIPA, defense news shocks cannot because they also incorporate changes to expectations about future spending. Because defense news shocks include pure shocks to expectations (including

[12] Alternate samples include the full sample starting in 1940 (including WWII), the sample beginning in 1947 (post-WWII, using only official quarterly GDP data), and from 1951 onwards (after the initial Korean War shock).

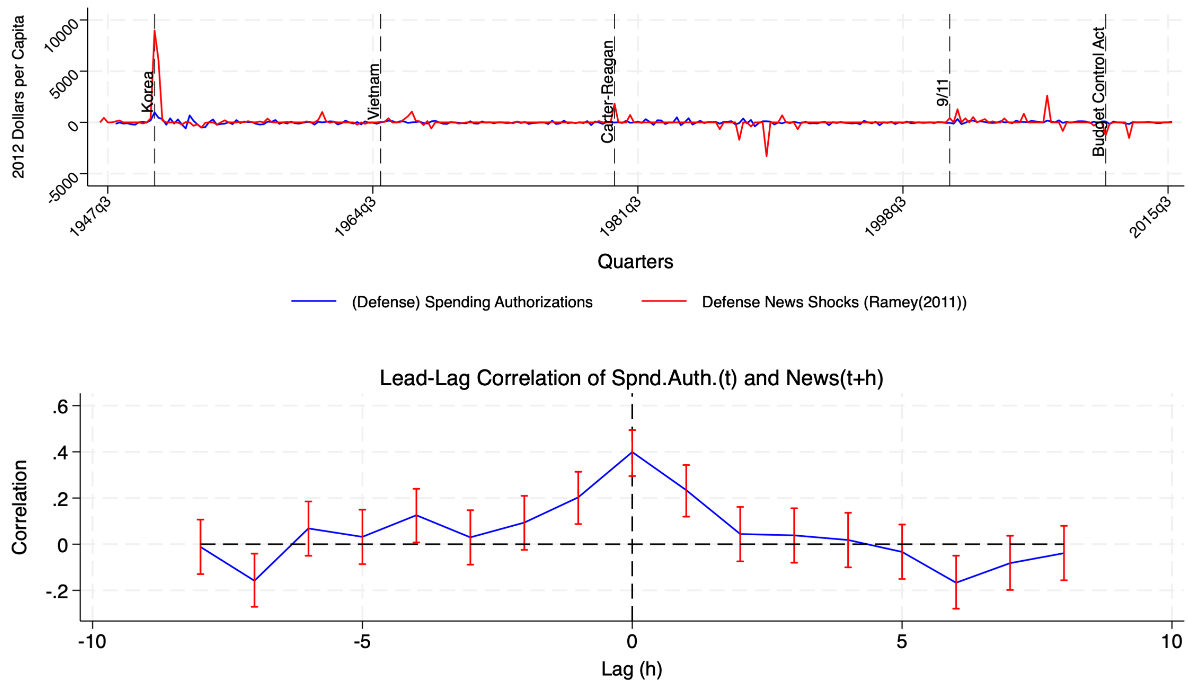


Figure A31: Spending Authorization Shocks and Defense News Shocks Occur at the same Time

Notes: variables are in real per capita values. Price deflator is GDP price deflator (2012=100). Shocks to (defense) spending authorizations are constructed as OLS residuals of a regression of real per capita values of spending authorizations on its four lags. Top panel: plot of shocks. Bottom panel: lead-lag correlation.

unrealized shocks), ordering defense news shocks first in a recursive VAR does not address the underlying problem that shocks to government spending, orthogonal to defense news shocks, are delayed in NIPA due to accounting rules. Using defense spending authorizations in place of NIPA defense spending *does* solve this timing problem—whether or not defense news shocks are used to augment the VAR—because spending authorizations directly correct for the timing of government purchases in NIPA. In other words: the major substantive difference between spending authorizations and NIPA is timing, while there are significantly larger differences between Ramey's defense news shocks and NIPA.

F.3 Cholesky Decompositions

One common approach to identifying government spending shocks is recursive ordering in a structural VAR. This approach uses the Cholesky decomposition to construct shocks uncorrelated with contemporaneous and lagged values of endogenous variables. Estimating the aggregate effect of government spending requires identifying structural shocks, which must be (i) uncorrelated with contemporaneous and lagged values of endogenous variables, (ii) uncorrelated with other shocks and (iii) unanticipated (Ramey, 2016). The Cholesky decomposition meets criterion (i) by construction. This approach was popularized by Blanchard and Perotti (2002), who order NIPA government spending first in a VAR and treat innovations to this variable as structural shocks. In this approach, identification hinges on the fact that quarterly variation in G is uncorrelated with quarterly variation in output due to implementation lags of fiscal policy, thus meeting condition (ii). However, Ramey (2011) highlights that this method fails to account for anticipation effects. In other words, condition (iii) is not satisfied.

More recent papers have attempted to correct for this problem by including Ramey's defense news shocks and ordering them first in the VAR (e.g., Auerbach and Gorodnichenko, 2012). This does not solve the underlying measurement problem: because defense news shocks include pure expectations shocks (e.g. unrealized shocks to expectations), they differ from changes in NIPA for many reasons other than timing. So including Ramey's defense news shocks does not correct for the timing problems in NIPA, spending shocks still appear in NIPA an average of 3-4 quarters after economic activity begins to respond. Thus estimated responses to spending still miss (and so underestimate) the initial response of economic activity.

G Robustness of Baseline Results

This section reports the robustness results for Section V.1 in the paper.

- IRFs:
  - Figure A32: Local Projections using sample 1951:1-2007:4 (post Korean war outbreak).
  - Figure A33: Local Projections using sample 1947:1-2019:4 (without R&R tax shocks).
  - Figure A34: Log-VAR with real per capita values. Baseline sample: 1951:1-2007:4.
- Fiscal Multiplier:
  - Figure A35: Fiscal Multipliers using sample 1951:1-2007:4 (post Korean war outbreak).
  - Figure A37: Fiscal Multipliers using sample 1940:1-2007:4 (full sample with WWII).
  - Figure A40: Fiscal Multipliers using sample 1947:1-2019:4 (full sample without WWII).
- IRFs and Multipliers using *Budget Authority* and *Contracts* as internal instruments for NIPA government spending, G.

Robustness IRFs Calculate impulse response functions using different samples.

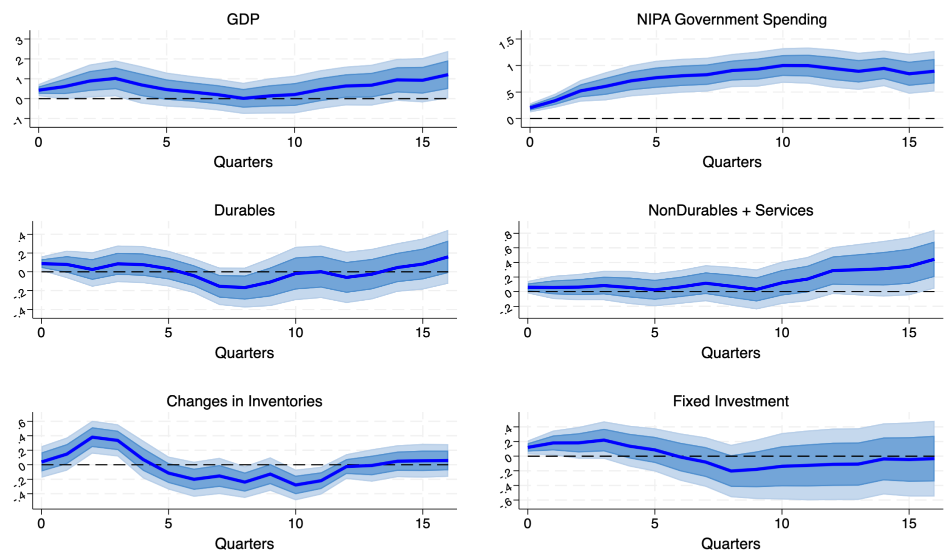


Figure A32: LP Responses of GDP Components to Spending Authorizations (1951Q1 - 2007Q4)

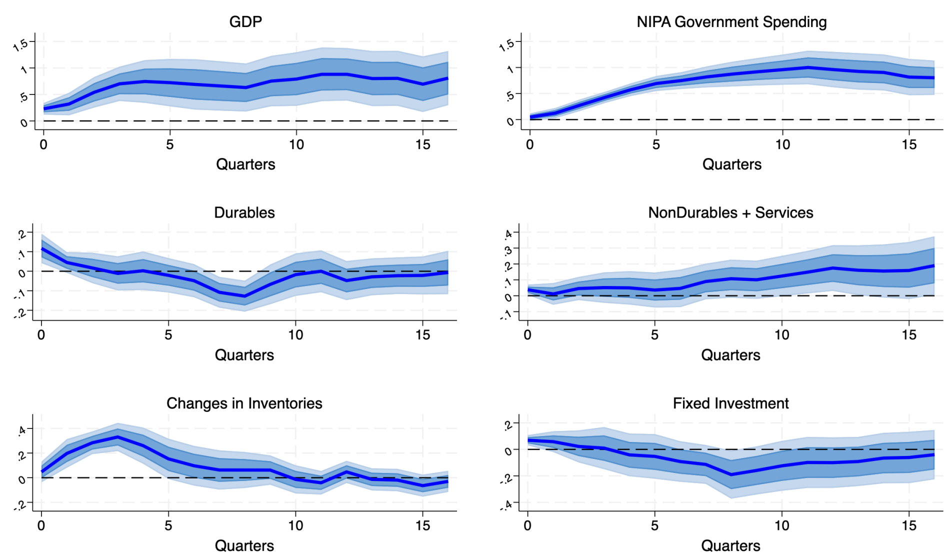


Figure A33: LP Responses of GDP Components to Spending Authorizations (1951Q1 - 2007Q4, excluding RR2010 Tax Shocks)

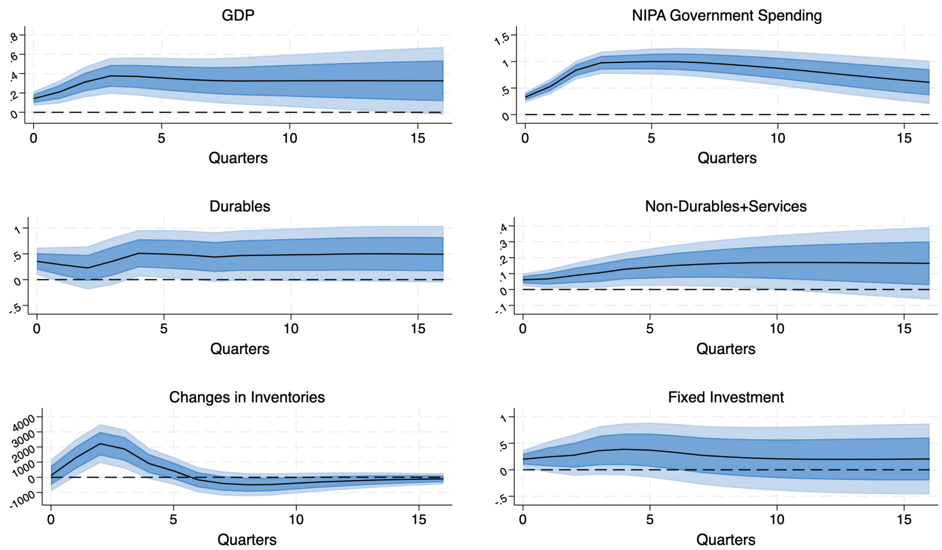


Figure A34: VAR Response of of Real Per Capita GDP Components to Spending Authorizations (1951Q1 - 2007Q4)

Notes: The VAR contains GDP, NIPA's G, spending authorizations, TB3, Romer & Romer (2010)'s exogenous tax shocks and the rotating outcome. Nominal variables are in logs of real per capita values. The Romer & Romer shocks and inventories are simply in real per capita values.

Robustness of Multipliers Calculate fiscal multipliers using different samples. Here fiscal multipliers refer to cumulative GDP multipliers relative to NIPA G, $\mathcal{M}_H$ using the paper notation.

Figure A35: Cum. Fiscal Multiplier - Sample: 1951:1 - 2007:4 (without Korean War Outbreak)

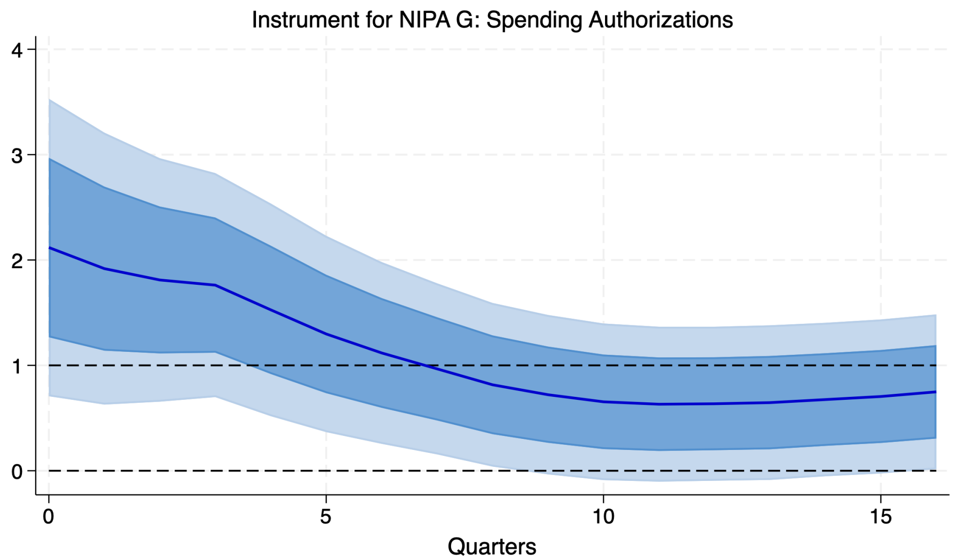


Table A5: Fiscal Multipliers by quarters - Sample 1951:1-2007:4 (Without Korean War Outbreak)

	Local Projections (LP-IV)				Recursive SVAR	
	Spending Authorizations		NIPA Defense Sp.		Spending Authorizations	NIPA Defense Sp.
Horizon	*Multiplier*	*Effective F*	*Multiplier*	*Effective F*	*Multiplier*	*Multiplier*
Impact	2.12	15.24	1.08	348.73	2.37	1.20
	(0.85)		(0.21)			
Quarter 1	1.92	19.88	1.15	180.70	2.39	1.32
	(0.78)		(0.27)			
Quarter 2	1.81	20.58	1.20	118.07	2.37	1.52
	(0.70)		(0.30)			
Quarter 3	1.76	19.35	0.98	97.50	2.44	1.47
	(0.64)		(0.32)			
Quarter 4	1.53	19.28	0.70	83.14	2.43	1.45
	(0.61)		(0.34)			
Quarter 5	1.30	19.45	0.45	69.38	2.37	1.40
	(0.56)		(0.35)			
Quarter 6	1.12	19.99	0.32	61.34	2.31	1.37
	(0.52)		(0.36)			
Quarter 7	0.97	20.61	0.26	54.56	2.25	1.35
	(0.49)		(0.37)			
Quarter 8	0.82	21.61	0.23	50.75	2.21	1.34
	(0.47)		(0.37)			

Figure A36: Cumulative Fiscal Multiplier - Sample: 1955:1 - 2007:4 (After Korea)

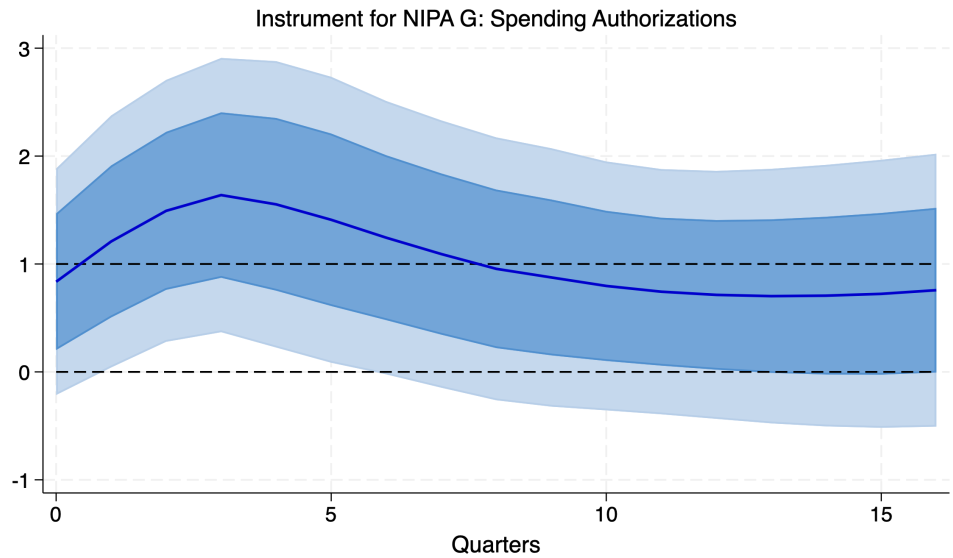


Table A6: Fiscal Multipliers by quarters - Sample 1955:1-2007:4 (After Korea)

	Local Projections (LP-IV)				*VAR Multiplier*	
	Spending Authorizations		*NIPA Defense Sp.*		*Spending Authorizations*	*NIPA Defense Sp.*
Horizon	*Multiplier*	*Effective F*	*Multiplier*	*Effective F*	*Multiplier*	*Multiplier*
Impact	0.84	21.99	1.12	263.73	1.52	1.23
	0.63		0.26			
Quarter 1	1.21	19.94	1.15	120.11	1.95	1.32
	0.71		0.35			
Quarter 2	1.49	24.02	1.23	82.26	2.24	1.51
	0.73		0.40			
Quarter 3	1.64	26.80	1.00	70.30	2.50	1.46
	0.77		0.42			
Quarter 4	1.55	24.77	0.74	62.45	2.55	1.44
	0.80		0.42			
Quarter 5	1.41	22.40	0.52	54.13	2.51	1.40
	0.80		0.43			
Quarter 6	1.24	21.44	0.45	47.40	2.47	1.37
	0.77		0.44			
Quarter 7	1.09	20.31	0.42	43.59	2.40	1.35
	0.75		0.45			
Quarter 8	0.96	19.27	0.40	41.33	2.32	1.34
	0.74		0.46			

Figure A37: Fiscal Multiplier - Full Sample: 1940:1 - 2019:4 (with WWII)

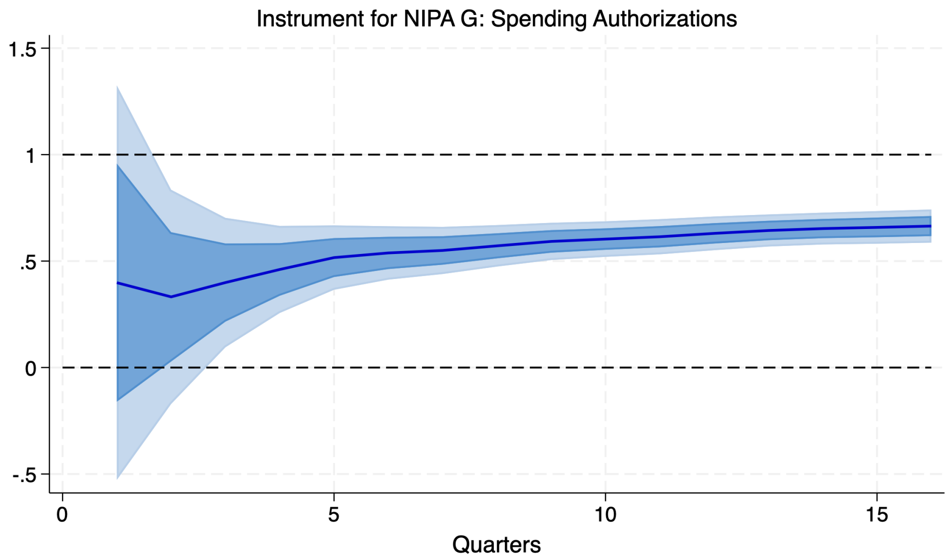
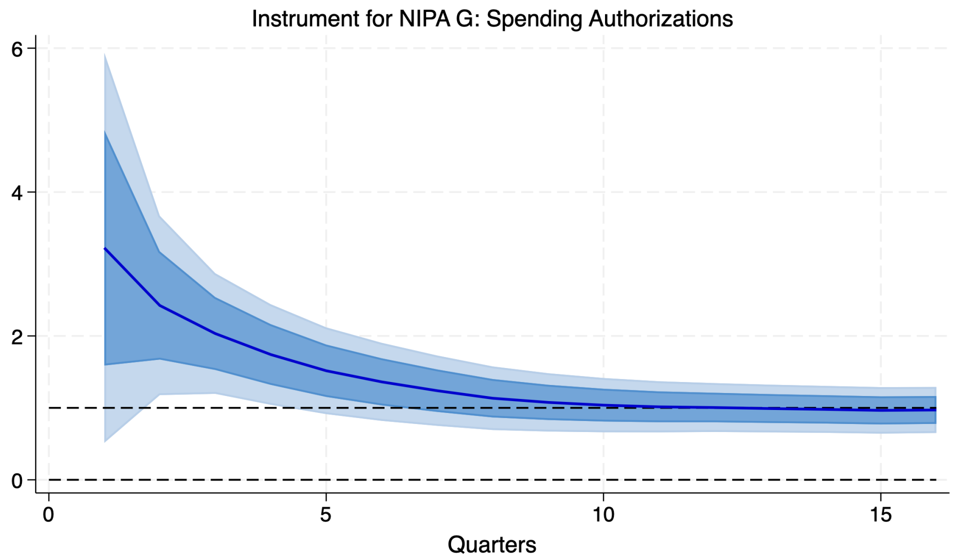


Table A7: Multipliers by quarters - Sample 1940:1-2019:4 (Full sample with WWII)

	Local Projections (LP-IV)				Recursive SVAR	
	Spending Authorizations		NIPA Government Sp.		Spending Authorizations	NIPA Government Sp.
Horizon	*Multiplier*	*Effective F*	*Multiplier*	*Effective F*	*Multiplier*	*Multiplier*
Impact	-4.66	0.01	0.75	-	5.05	1.01
	66.58		0.10			
Quarter 1	0.40	3.96	0.66	342.70	2.02	0.85
	(0.56)		(0.09)			
Quarter 2	0.33	7.38	0.62	142.89	1.55	0.81
	(0.30)		(0.07)			
Quarter 3	0.40	10.90	0.59	112.55	1.44	0.81
	(0.18)		(0.05)			
Quarter 4	0.46	16.40	0.60	92.58	1.32	0.83
	(0.12)		(0.05)			
Quarter 5	0.52	20.10	0.62	79.74	1.22	0.86
	(0.09)		(0.05)			
Quarter 6	0.54	19.99	0.63	71.90	1.16	0.88
	(0.08)		(0.05)			
Quarter 7	0.55	18.44	0.64	66.66	1.13	0.90
	(0.07)		(0.04)			
Quarter 8	0.57	17.46	0.65	61.22	1.11	0.92
	(0.06)		(0.04)			

Figure A38: Fiscal Multiplier - Sample: 1947:1 - 2019:4
(without RR2010 Tax Shocks)



Results with Budget Authority and Contracts Calculate impulse response functions and the corresponding multipliers using (a) defense *contracts* and (b) defense spending authorizations aggregated at fiscal year frequency, i.e, defense *budget authority*. In particular, we use these two variables as internal instruments for NIPA government spending, G. Results refer to the baseline sample 1947:1 to 2007:4.

Figure A39: LP Response of GDP Components to Contracts - Sample: 1947:1 - 2007:4

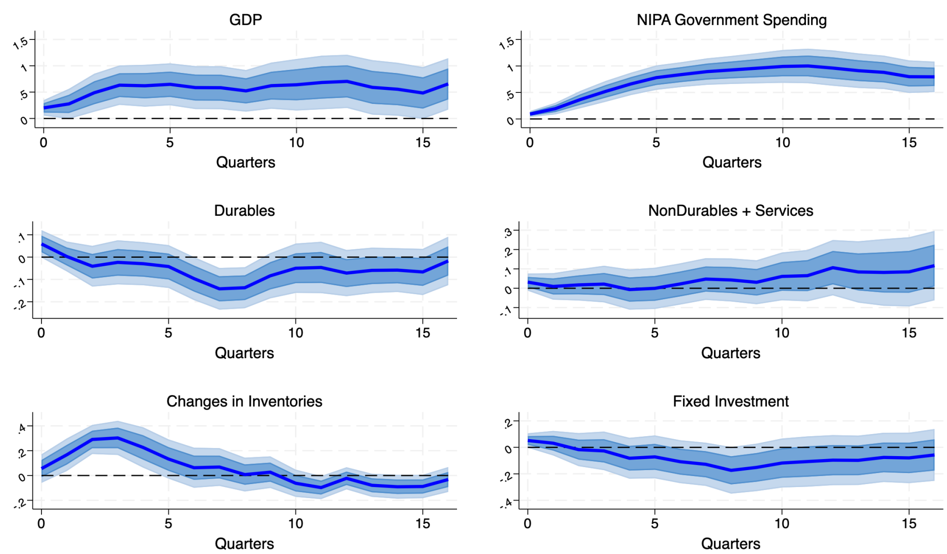


Figure A40: Fiscal Multiplier using **Contracts** - Sample: 1947:1 - 2007:4

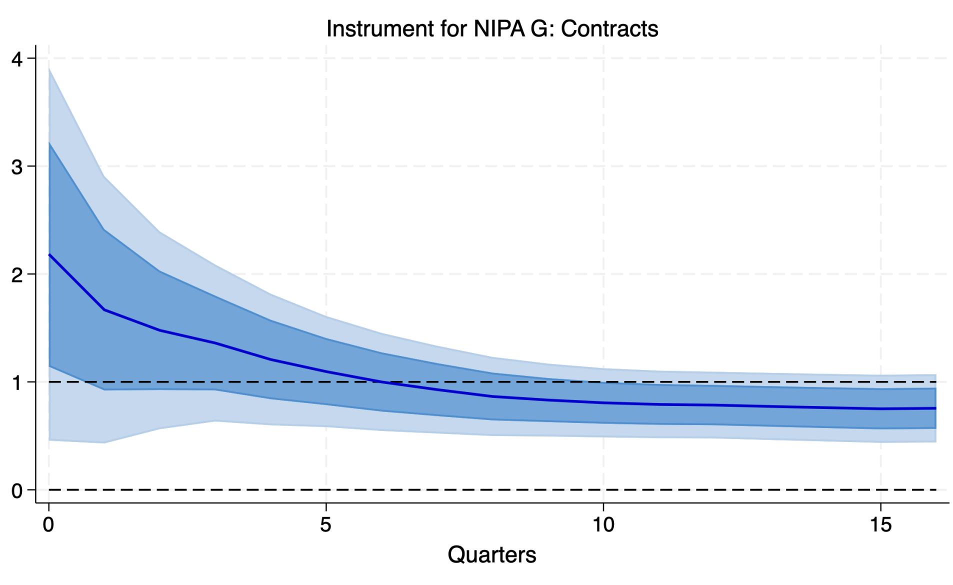


Figure A41: LP Response of GDP Components to Budget Authority - Sample: 1947:1 - 2007:4

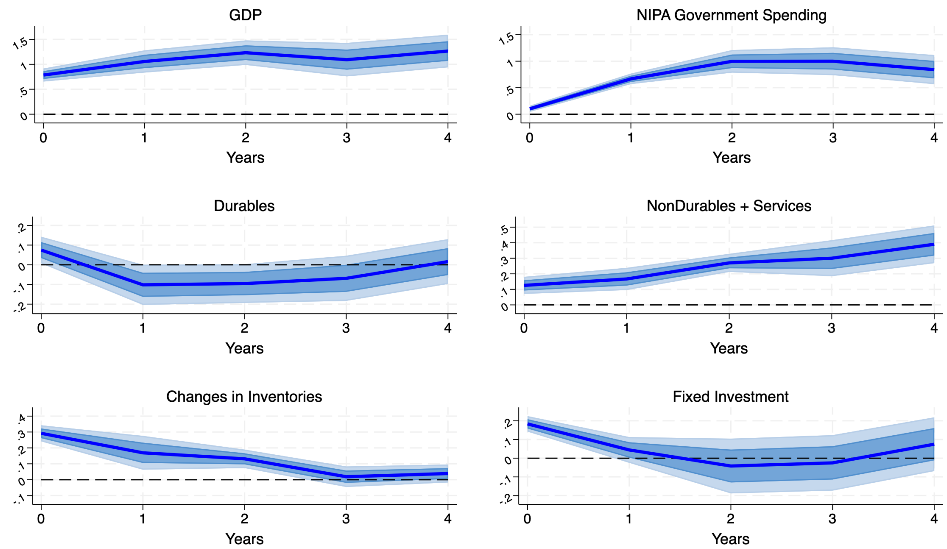
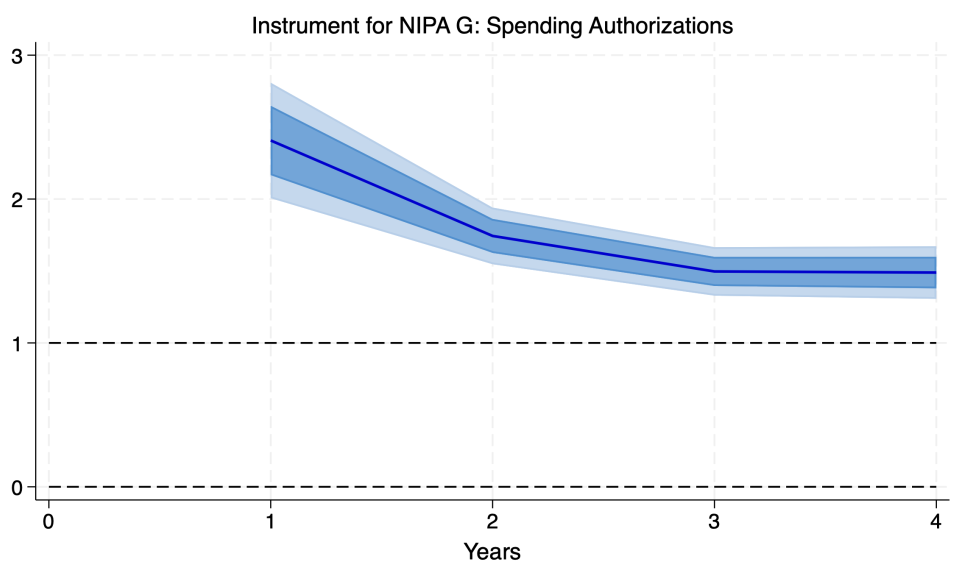


Figure A42: Fiscal Multiplier using **Budget Authority** - Sample: 1947:1 - 2007:4



H Multiplier Breakdown - Mathematical Derivation

First, consider the following set of impulse response functions with respect to a shock to spending authorizations:

$$
\begin{aligned}
\frac{\mathrm{GDP}_{t+h} - \mathrm{GDP}_{t-1}}{\mathrm{GDP}_{t-1}} &= \beta_h \cdot \frac{\mathrm{SA}_t - \mathrm{SA}_{t-1}}{\mathrm{GDP}_{t-1}} + \text{lags} + \nu_{t+h}, \\
\frac{\mathrm{G}_{t+h} - \mathrm{G}_{t-1}}{\mathrm{GDP}_{t-1}} &= \gamma_h \cdot \frac{\mathrm{SA}_t - \mathrm{SA}_{t-1}}{\mathrm{GDP}_{t-1}} + \text{lags} + \epsilon_{t+h} \\
\frac{\mathrm{PRS}_{t+h} - \mathrm{PRS}_{t-1}}{\mathrm{GDP}_{t-1}} &= \phi_h \cdot \frac{\mathrm{SA}_t - \mathrm{SA}_{t-1}}{\mathrm{GDP}_{t-1}} + \text{lags} + \varepsilon_{t+h}, \\
\frac{\mathrm{SA}_{t+h} - \mathrm{SA}_{t-1}}{\mathrm{GDP}_{t-1}} &= \rho_h \cdot \frac{\mathrm{SA}_t - \mathrm{SA}_{t-1}}{\mathrm{GDP}_{t-1}} + \text{lags} + \xi_{t+h},
\end{aligned}
$$

where lags includes four lags of GDP, NIPA's G, defense spending authorizations, TB3 and the Romer and Romer exogenous tax shocks; GDP_t is GDP, G_t is the NIPA measure of government spending, SA_t is defense spending authorizations and PRS_t is private spending: $\mathrm{PRS}_t := \mathrm{GDP}_t - \mathrm{G}_t$. The OLS estimates β_h and γ_h represent the impulse response functions at horizon h of GDP and NIPA G to a shock to spending authorizations, as reported in the paper. Similarly, the OLS estimate ϕ_h is the impulse response functions of private spending to a shock to spending authorizations. Thanks to (i) the linearity of the OLS estimator and (ii) the identical right-hand-side of the above equations, we have:

$$\beta_h = \phi_h + \gamma_h.$$

Lastly, ρ_h is the impulse response function of spending authorizations to its shock; it is equal to one on impact and then it geometrically decays, reflecting serial correlation in spending authorizations in response to a shock.

The standard fiscal multiplier at horizon H is defined as the ratio of the integral of the impulse response function of GDP and NIPA G:

$$\mathcal{M}_H = \frac{\sum_{h=0}^{H} \beta_h}{\sum_{h=0}^{H} \gamma_h}. \tag{9}$$

which can be broken down using the response of private spending and NIPA G:

$$\mathcal{M}_H = \frac{\sum_{h=0}^{H} \phi_h + \gamma_h}{\sum_{h=0}^{H} \gamma_h} = 1 + \frac{\sum_{h=0}^{H} \phi_h}{\sum_{h=0}^{H} \gamma_h}.$$

Now, divide both the numerator and the denominator of the above ratio by the cumulative impulse response function of defense spending authorizations ($\sum_{h=0}^{H} \rho_h$):

$$\mathcal{M}_H = 1 + \frac{\frac{\sum_{h=0}^{H} \phi_h}{\sum_{h=0}^{H} \rho_h}}{\frac{\sum_{h=0}^{H} \gamma_h}{\sum_{h=0}^{H} \rho_h}} = 1 + \frac{\mathcal{P}_H^A}{\mathcal{G}_H^A},$$

where

$$\mathcal{G}_H^A = \frac{\sum_{h=0}^{H} \gamma_h}{\sum_{h=0}^{H} \rho_h}, \quad \mathcal{P}_H^A = \frac{\sum_{h=0}^{H} \phi_h}{\sum_{h=0}^{H} \rho_h}.$$

Define the numerator $\mathcal{P}_H^A$ as the private spending *authorization multiplier*: the effect of \$1 of defense spending authorization on private spending. Similarly, define the denominator $\mathcal{G}_H^A$ the *spending-to-authorization multiplier*: the effect of \$1 of spending authorization on NIPA government spending. In the medium to long run, we expect that every authorized dollar is spent and ultimately flows into NIPA G:

$$\lim_{H \to \infty} \mathcal{G}_H^A = 1$$

In other words, $\mathcal{G}_H^A$ measures the *time-to-spend* authorized funds. The faster $\mathcal{G}_H^A$ converges to one, the faster the government disburses funds. In the short-run, we have that $\mathcal{G}_H^A < 1$, since it takes time for authorized funds to be disbursed due to measurement delays in the defense procurement accounting.

Interpretation In other words, the standard GDP fiscal multiplier $\mathcal{M}_H$ ("*effect of \$1 of NIPA spending on output*") can be broken down as one plus the ratio of the private spending authorization multiplier ("*effect of \$1 of defense spending authorization on private spending*") and the spending-to-authorization multiplier ("*effect of \$1 of defense spending authorization on NIPA spending*").

As shown in the paper, ultimately, every authorized dollar is spent in the medium to long run (i.e., $\lim_{H\to\infty} \mathcal{G}_H^A = 1$), therefore, we can construct a "*heuristic multiplier*" as $1 + \mathcal{P}_H^A$, assuming $\mathcal{G}_H^A = 1$.

The advantage of the heuristic multiplier is that it is not inflated at short horizons thus providing a reliable measure of the true impact of fiscal policy on output. In fact, the delayed response of NIPA G, implies that $\sum_{h=0}^{H} \gamma_h$ is close to zero at short horizons, thus inflating the standard fiscal multiplier, as discussed in the model section of the paper: the denominator of the multiplier is close to zero and the value of the multiplier diverges to infinity on impact.

Estimation A single-step IV framework (LP-IV) is able to provide consistent estimates of the multipliers with the extra benefit of directly estimating the confidence bands (Ramey, 2016).

To estimate the usual fiscal multiplier, $\mathcal{M}_H$:

$$\sum_{h=0}^{H} \frac{\text{GDP}_{t+h} - \text{GDP}_{t-1}}{\text{GDP}_{t-1}} = \mathcal{M}_H \cdot \underbrace{\sum_{h=0}^{H} \frac{\text{G}_{t+h} - \text{G}_{t-1}}{\text{GDP}_{t-1}}}_{\text{Instrument: } Z_t} + \text{lags} + \nu_t, \tag{10}$$

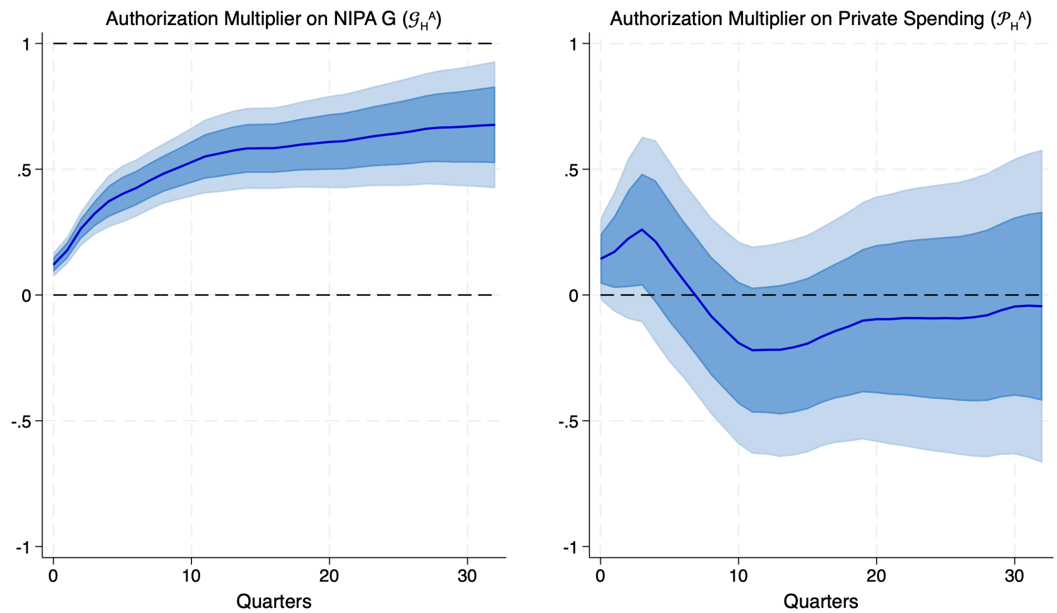
where the cumulative change in NIPA G is instrumented with:

$$Z_t = \frac{\text{SA}_t - \text{SA}_{t-1}}{\text{GDP}_{t-1}}.$$

To estimate the private spending to authorization multiplier, $\mathcal{P}_H^A$:

$$\sum_{h=0}^{H} \frac{\text{PRS}_{t+h} - \text{PRS}_{t-1}}{\text{GDP}_{t-1}} = \mathcal{P}_H^A \cdot \underbrace{\sum_{h=0}^{H} \frac{\text{SA}_{t+h} - \text{SA}_{t-1}}{\text{GDP}_{t-1}}}_{\text{Instrument: } Z_t} + \text{lags} + \nu_t, \tag{11}$$

Figure A43: Analogous of Figure 11 in the Paper: Sample 1951-2007



Lastly, to estimate the spending to authorization multiplier, $\mathcal{G}_H^A$:

$$\sum_{h=0}^{H} \frac{\mathrm{G}_{t+h} - \mathrm{G}_{t-1}}{\mathrm{GDP}_{t-1}} = \mathcal{G}_H^A \cdot \underbrace{\sum_{h=0}^{H} \frac{\mathrm{SA}_{t+h} - \mathrm{SA}_{t-1}}{\mathrm{GDP}_{t-1}}}_{\text{Instrument: } Z_t} + \text{lags} + \nu_t, \tag{12}$$

Thanks to (i) the linearity of the 2SLS, (ii) the fact that $\mathrm{GDP}_t = \mathrm{PRS}_t + \mathrm{G}_t$ by construction and (iii) the identical right-hand-side in the above expressions, we have that the estimated 2SLS multipliers will satisfy the following condition:

$$\mathcal{M}_H = 1 + \frac{\mathcal{P}_H^A}{\mathcal{G}_H^A}.$$

H.1 Robustness

In this section we replicate Figure 11 and Table 5 of the paper using two other samples. First, we look at the sample from 1951:1 which excludes the interpolated contracts data and the outbreak of the Korean war. Second, we look at the whole sample with WWII.

Table A8: Analogous of Table 5 in the Paper for sample 1951-2007

Horizon	*GDP Multiplier*	*Private Spending to Authorization Multiplier*		*Heuristic Multiplier*
		Multiplier	*pvalue*	
Impact	2.12	0.13	0.19	1.13
	(0.85)	(0.10)		(0.10)
Quarter 1	1.92	0.16	0.27	1.16
	(0.78)	(0.14)		(0.14)
Quarter 2	1.81	0.21	0.27	1.25
	(0.70)	(0.19)		(0.19)
Quarter 3	1.76	0.25	0.27	1.12
	(0.64)	(0.22)		(0.22)
Quarter 4	1.53	0.20	0.41	0.98
	(0.61)	(0.23)		(0.23)
Quarter 5	1.30	0.12	0.61	1.12
	(0.56)	(0.23)		(0.23)
Quarter 6	1.12	0.05	0.82	1.05
	(0.52)	(0.22)		(0.22)
Quarter 7	0.97	-0.02	0.95	0.98
	(0.49)	(0.22)		(0.22)
Quarter 8	0.82	-0.09	0.70	0.91
	(0.47)	(0.22)		(0.22)

Table A9: Analogous of Table 5 in the Paper for sample 1940-2019

Horizon	*GDP Multiplier*	*Private Spending to Authorization Multiplier*		*Heuristic Multiplier*
		Multiplier	*pvalue*	
Impact	-4.66	-0.01	0.74	0.99
	(66.58)	(0.02)		(0.02)
Quarter 1	0.40	-0.03	0.29	0.97
	(0.56)	(0.02)		(0.02)
Quarter 2	0.33	-0.06	0.05	0.94
	(0.30)	(0.03)		(0.03)
Quarter 3	0.40	-0.09	0.02	0.91
	(0.18)	(0.03)		(0.03)
Quarter 4	0.46	-0.12	0.01	0.88
	(0.12)	(0.04)		(0.04)
Quarter 5	0.52	-0.15	0.01	0.85
	(0.09)	(0.05)		(0.05)
Quarter 6	0.54	-0.18	0.00	0.82
	(0.08)	(0.05)		(0.05)
Quarter 7	0.55	-0.20	0.00	0.80
	(0.07)	(0.05)		(0.05)
Quarter 8	0.57	-0.21	0.00	0.79
	(0.06)	(0.04)		(0.04)

Figure A44: Analogous of Figure 11 in the Paper: Sample 1955-2007

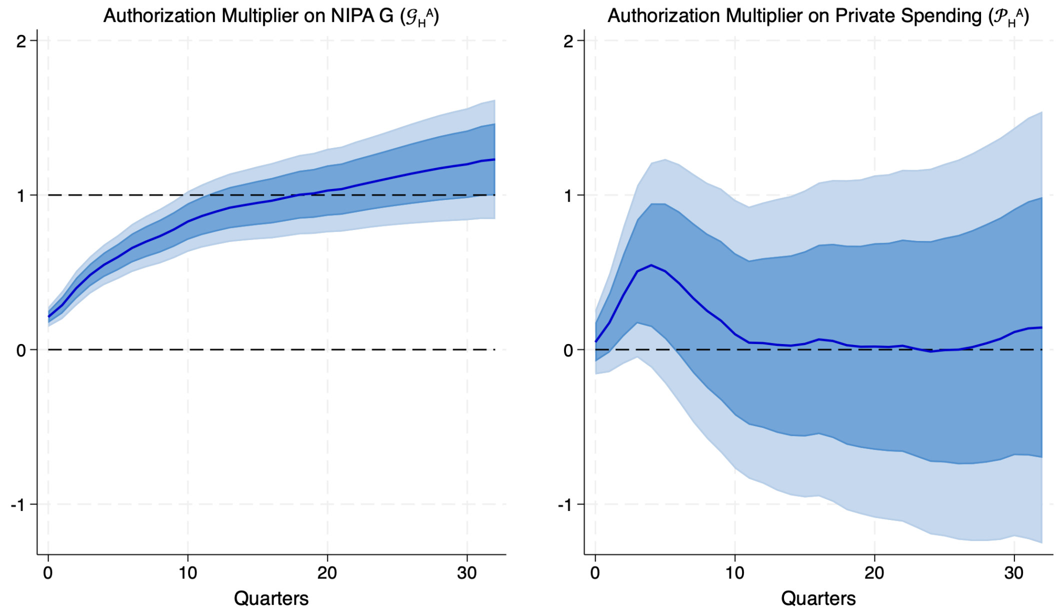


Figure A45: Analogous of Figure 11 in the Paper: Sample 1940-2019

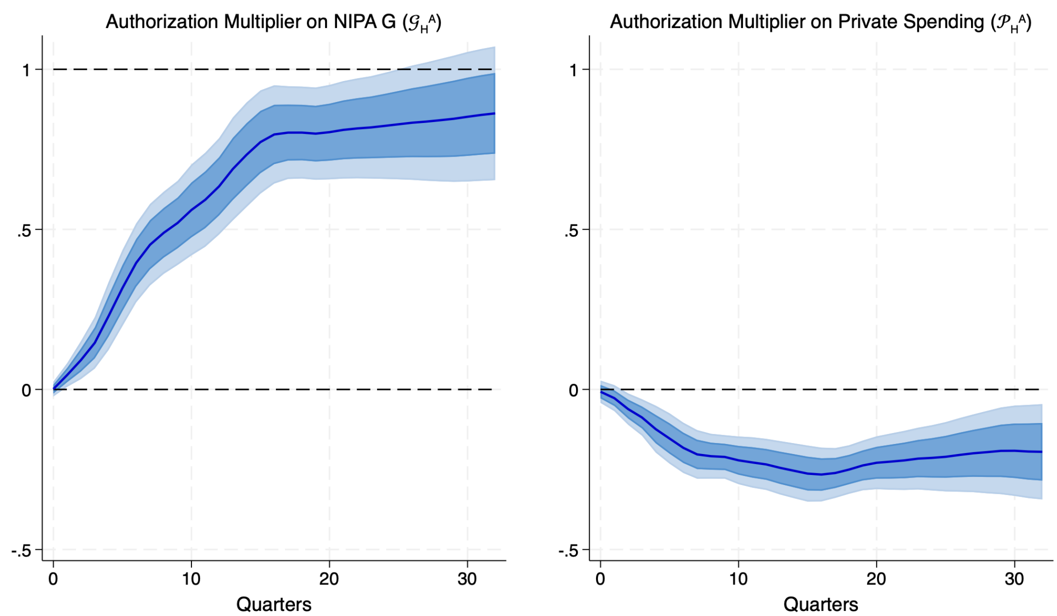


Table A8 breaks down the fiscal multiplier $\mathcal{M}_H$ for the sample starting in 1951:1, and provides a robustness check for Table 5 in the paper. The values of the authorization multiplier on private spending, $\mathcal{P}_H^A$ and the authorization to spending multiplier, $\mathcal{G}_H^A$ are plotted in Figure A43. Notice that $\mathcal{G}_H^A$ does not appear to converge to 1 in this case. This is due to the fact that the sample begins in the middle of the Korean war. Figure A44 shows the same figure when the sample begins in 1955:1, that is, after the end of the Korean war. Notice that in this case $\mathcal{G}_H^A$ does converge to one.

Table A9 and Figure A45 replicate the analysis by letting the sample begin in 1940:1 and end in 2019:4. Recall that in this case there is no Romer and Romer tax control, due to lack of data for this time period.

I Industry Analysis: Inventory Responses Are Driven by Defense Sector

Given the positive and robust aggregate response of inventories in response to fiscal shocks, we examine the heterogeneity in this response across manufacturing industries during periods of military build-ups. We find that industries more closely connected to the government through defense contracts experience relatively larger increases in inventories compared to industries less connected to the government.

Data. We use monthly data from the Bureau of Economic Analysis (BEA) to construct a panel of real inventories for 18 manufacturing industries covering the period from January 1959 to December 1997, which we de-trend using Hamilton (2018).[13]

The data is reported in real chained dollars, which limits our ability to make quantitative statements regarding the relative response of manufacturing inventories compared to their aggregate response analyzed in the main body of the paper. However, we document that the response of manufacturing inventories is proportional to the degree of connection with the government, providing compelling evidence for our proposed mechanism. Specifically, during a military build-up, inventories increase as they capture the work-in-progress of defense contractors, which is not recorded in NIPA government spending until the items are completed and delivered to the government. Conversely, if inventories were increasing for other reasons, we would expect to observe a uniform response across sectors, regardless of their degree of connection with the government.

I.1 Construction of Industry Weights: Reliance on Government Purchases

To construct industry weights, we combine information from the Make and Use tables with more than 60 non-government sectors between 1963 and 1996. Following Horowitz and Planting (2009),

[13] We thank Valerie Ramey for providing this data. While our data ends in 1997, most of the variation in defense spending originates from earlier periods, such as the Vietnam War and the Soviet invasion of Afghanistan.

we derive direct requirement industry-by-industry matrices A_t and direct sales from the private sectors to the government.

Government Direct Purchases. We construct a vector of government purchases (i.e., direct requirements) relative to industry output:

$$\underset{n\times 1}{\boldsymbol{\gamma}_{0,t}} = \begin{bmatrix} \frac{\text{SALES}_{1\to G,t}}{\text{SALES}_{1,t}} \\ \vdots \\ \frac{\text{SALES}_{n\to G,t}}{\text{SALES}_{n,t}}, \end{bmatrix}$$

where t denotes the year, n is the number of manufacturing sub-industries, G denotes the federal general government, and the 0 subscript refers to the order of included input-output connections, that is, it only accounts for direct sales to the government. $\text{SALES}_{i\to G,t}$ includes both government gross investments, which appear as final uses in the Use tables, and direct requirements. We report the time-average values of $\boldsymbol{\gamma}_{0,t}$ in the third column of Table A10.

Government Indirect Purchases Following Nekarda and Ramey (2011), we also include downstream input-output linkages to account for indirect sales to the government. In order to do so, we construct yearly $n \times n$ input-output matrices A_t in which (i, j)th element of matrix A_t is given by:

$$\frac{\text{SALES}_{i\to j,t}}{\text{SALES}_{i,t}}.$$

We then construct a vector of direct and first-order indirect sales shares as follows:

$$\boldsymbol{\gamma}_{1,t} = (I_n + A_t) \cdot \boldsymbol{\gamma}_{0,t}.$$

Sector	*Commodity Description*:	$\gamma_{0,i}$	$\gamma_{1,i}$	$\gamma_{2,i}$	θ_i
3364	Other transportation equipment	34.43%	42.00%	43.94%	1.00
334	Computer and electronic products	13.09%	17.04%	18.38%	0.42
323	Printing and related support activities	7.98%	9.35%	9.95%	0.23
332	Fabricated metal products	3.73%	4.78%	5.37%	0.12
3361	Motor vehicles, bodies and trailers, and parts	2.09%	3.70%	4.64%	0.11
339	Miscellaneous manufacturing	2.31%	3.80%	4.49%	0.10
333	Machinery	2.65%	3.84%	4.44%	0.10
335	Electrical equipment, appliances, and components	2.37%	3.66%	4.31%	0.10
325	Chemical products	1.91%	3.50%	4.27%	0.10
324	Petroleum and coal products	2.71%	3.50%	4.17%	0.09
326	Plastics and rubber products	1.13%	2.20%	2.89%	0.07
337	Furniture and related products	0.66%	1.63%	2.19%	0.05
331	Primary metals	0.54%	1.44%	2.06%	0.05
313	Textile mills and textile product mills	0.48%	1.31%	2.01%	0.05
315	Apparel and leather and allied products	0.57%	1.37%	1.98%	0.05
327	Nonmetallic mineral products	0.49%	1.35%	1.91%	0.04
322	Paper products	0.51%	1.25%	1.83%	0.04
311	Food and beverage and tobacco products	0.38%	1.16%	1.77%	0.04
321	Wood products	0.19%	0.91%	1.53%	0.03

Table A10: *Notes:* The last column divides $\theta_{2,i}$ by the maximum value (i.e., the one for Other Transportation Equipment Manufacturing). The weights θ_i used in this paper include second-order connections, normalized as shown in the last column.

Notice that the ith element of $\boldsymbol{\gamma}_{1,t}$ is given by:

$$\gamma_{1,i,t} = \underbrace{\frac{\text{SALES}_{i\to G,t}}{\text{SALES}_{i,t}}}_{\text{Direct Sales}} + \underbrace{\sum_{j=1}^{n} \frac{\text{SALES}_{i\to j,t}}{\text{SALES}_{i,t}} \cdot \frac{\text{SALES}_{j\to G,t}}{\text{SALES}_{j,t}}}_{\text{Indirect Sales.}}$$

We report the time-average of $\boldsymbol{\gamma}_{1,t}$ in the fourth column of Table A10. Similarly, we construct direct, first and second order indirect sales to the government, shares of total output as:

$$\boldsymbol{\gamma}_{2,t} = \left(I_n + A_t + A_t^2\right) \cdot \boldsymbol{\gamma}_{0,t}.$$

We report the time-average values of $\boldsymbol{\gamma}_{2,t}$ in the fifth column of Table A10. As you can see, second-order downstream connections do not add up much to first-order connections, i.e. $\gamma_1 \approx \gamma_2$. Therefore, we only include downstream connections up to the second order.

Finally, we construct our industry weights θ_i as:

$$\theta_i := \frac{\mathbb{E}\left[\gamma_{2,i,t}\right]}{\max_i \mathbb{E}\left[\gamma_{2,i,t}\right]}$$

We report the weights in the last column of Table A10. Without loss of generality, we normalize the weight by the maximum value of γ_2, which corresponds to the value of Other Transportation Equipment manufacturing.

I.2 The Heterogeneous Response of Manufacturing Industries

The production of defense goods is concentrated in the manufacturing sector (see e.g., Ramey and Shapiro (1998), Nekarda and Ramey (2011) and Cox, Muller, et al. (2023)). However, the level of government involvement varies greatly among manufacturing sub-industries. For example, the 'Other Transportation Equipment" sector has 34% of its sales directly from the government. Accounting for indirect sales via input-output connections, the sector's dependence on government purchases rises to 42% and 44% with first and second order downstream connections included (as done in Nekarda and Ramey (2011)). This heavy reliance on government purchases is unsurprising given that the sector includes sub-industries like Aircraft, Ship Building, Guided Missiles, and Space Vehicles. Conversely, the Wood Products" sector has no sales to the government as it does not include any defense item producers.

Therefore, we construct a weight θ_i for each industry which captures the long-run average share of industry sales coming from government purchases (i.e., last column of Table A10). Using industry-by-industry input-output matrices, our weights include up to second-order downstream connections. Then we estimate the following equation:

$$\text{Invt}_{i,t+h} = \lambda_{ih} + \alpha_h \cdot \text{War}_t + \beta_h \cdot \text{War}_t \cdot \theta_i + \sum_{p=1}^{12} \varphi_{ph} \cdot \text{Invt}_{i,t-p} + \varepsilon_{i,t+h} \tag{13}$$

where $h = 0, 1, ..., 24$, Invt_{it} is total real inventories of industry i in month t, λ_{ih} is an industry fixed-effect, and War_t is war dates.[14] Consistent with Ramey and Shapiro (1998) and Eichenbaum and Fisher (2005), our war event variable is a weighted dummy with value 1 on March 1965 and 0.3 on January 1984 to indicate the start of the Vietnam War and Soviet invasion of Afghanistan, respectively. Results using unweighted war dates are similar and are reported in the robustness section, below.

To build intuition on the interpretation of the coefficients of interest, we can collect the war-date shock form the above equation to get the following expression:

$$... + (\alpha_h + \beta_h \cdot \theta_i) \cdot \text{War}_t + ...$$

α_h captures the average effect of a military build-up on sectoral inventories that would have occurred, had not the sectors be connected to the government, i.e., $\theta_i = 0$. On the contrary, β_h captures the heterogeneous response of sectors to war dates, proportional to their degree of connection to the government. As long as estimates of β_h are positive and significant, it means that the sectoral response of inventories to a military build-up is a positive function of their degree of connection to the government.

The right panel of Figure A46 shows the OLS estimates of β_h.

The estimates are positive and significant, indicating that the effect of war dates on sectoral inventories is a positive function of the degree of connection of a sector to the government. The left panel of the same figure shows the estimates of α_h, which is negative and very small. The sum of α_h and β_h corresponds to the total effect of a war date on a sector with weight equal to one, i.e. $\theta_i = 1$, which corresponds to the sector with the largest reliance. Notice that the positive effect of a military build-up on inventories of the sector with $\theta_i = 1$ is all driven by the β_h. On the contrary,

[14] We use war dates instead of defense news shocks since the former can easily be converted into monthly frequency to match our inventories data.

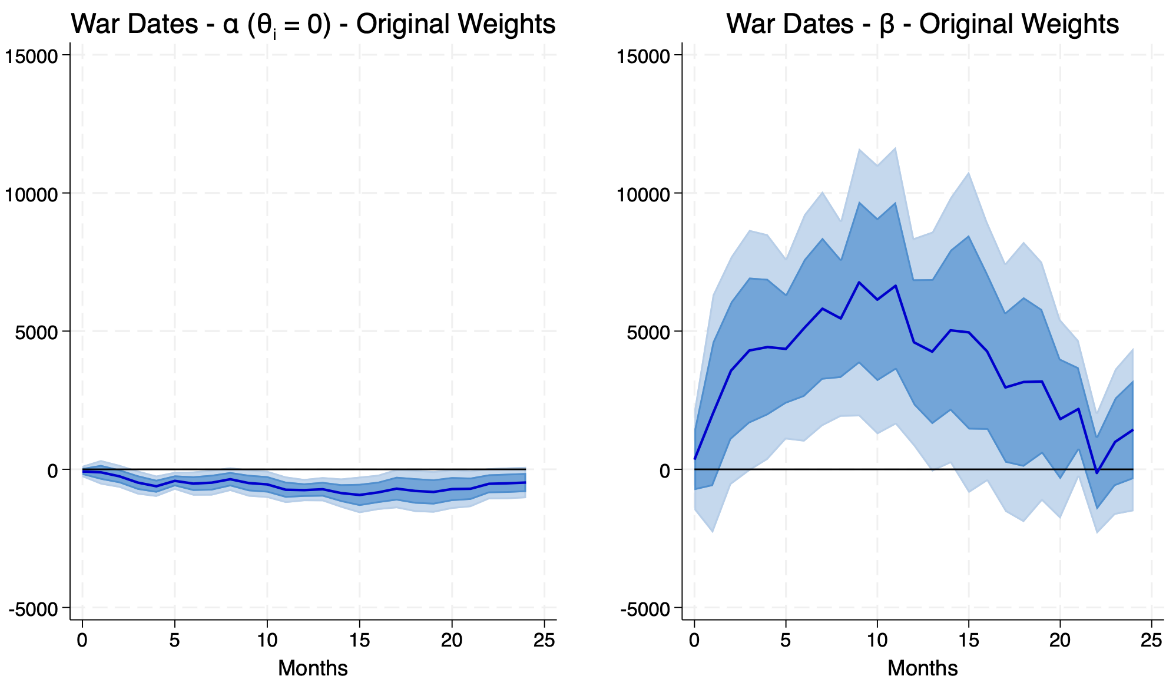


Figure A46: Response of Sectoral Inventories to War Events.

Notes: Left panel shows estimates of α_h (response when $\theta_i = 0$), right panel reports estimates of $\alpha_h + \beta_h$ (response when $\theta_i = 1$). Weights are normalized by maximum weight (i.e. the one of Other Transportation Equipment Manufacturing). Since Real Inventories are trending, data is filtered using Hamilton (2018)'s filter (we set $h = 24$ and $p = 12$, that is two years lag plus one more year of lags). The unit of real inventories is millions of 2005 chained dollars. Sample goes from 1959-Jan to 1997-Dec and uses 18 sectors breakdown of Manufacturing. Confidence bands are 68% and 90%. Standard errors are obtained via Bootstrap (standard Stata routine for `xtreg`: we use `vce(boot)` and set the seed for replicability of results; Stata uses a non-parametric type of bootstrap which resamples data with replacement).

sectors with values of θ_i close to zero will have a zero net effect of military build-up on inventories:

$$(\text{Total Effect})_{i,h} = \alpha_h + \beta_h \cdot \underbrace{\theta_i}_{\approx 0} \approx 0.$$

Therefore, all of the effect of war dates on manufacturing sectoral inventories is explained by the degree of connection of each sector to the government.

I.3 Robustness

We carry out three robustness checks.

Business Cycle Sensitivity. We verify that the heterogeneous response of manufacturing industries' inventories – the positive, significant and large estimate of β_h – is not driven by a different sectoral sensitivity to the business-cycle. In particular, we replace war dates with a monetary policy shock, as constructed narratively by Romer and Romer (2004) and updated by Wieland and Yang (2020) and then we re-estimate α_h and β_h in this case.

The left column of Figure A47 reports the baseline estimates of α_h (top-panel) and β_h (bottom-panel) reported in Figure A46. The right column of Figure A47 shows the estimates of α_h and β_h when war dates are replaced with the monetary policy shock.

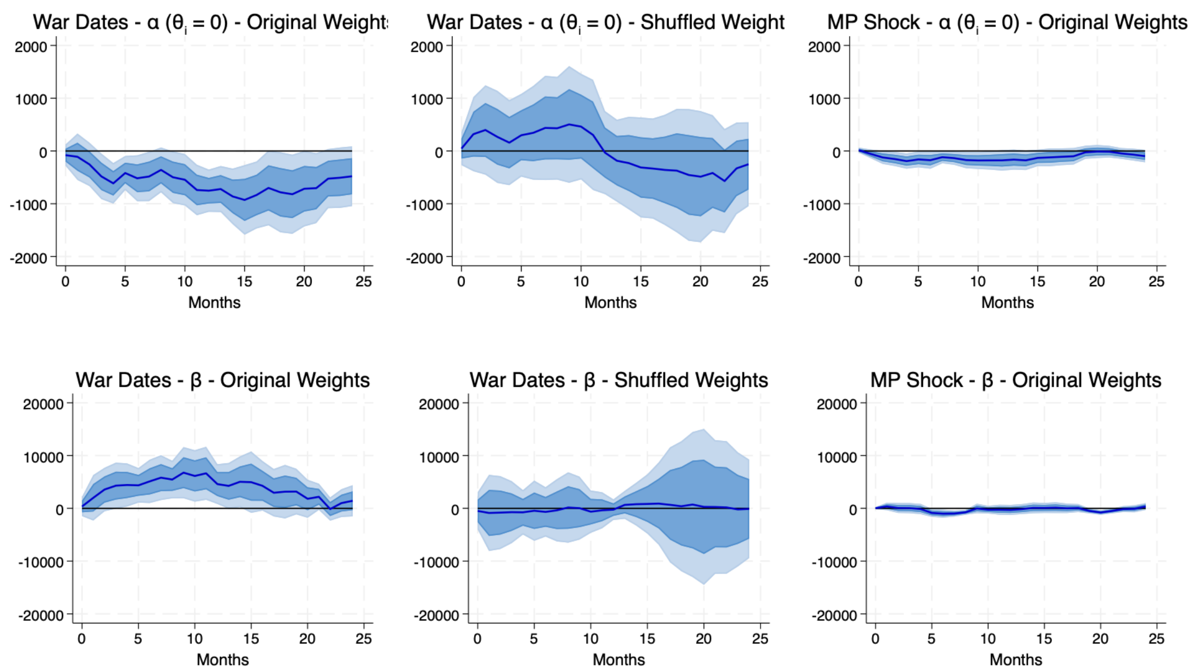


Figure A47: Response of Sectoral Inventories to War Events (Robustness)

Notes: Same as in Figure A46.

Notice from the top and bottom panels of the right columns that the estimates of α_h and β_h are not statistically different from zero. This confirms that the reaction of federal contractors to defense news shocks is driven by war-related forces and not the associated business-cycle fluctuations.[15]

[15] We thank Juan Herreño for suggesting this test.

Placebo. Furthermore, we make sure that the differential response of defense industries during a military build-up is not driven by spurious correlation. In particular, we re-estimate Equation (13) using randomly re-shuffled weights, θ_i. Results are reported in the middle panel of Figure A47. Again, we estimate the heterogeneous response β_h to be statistically indistinguishable from zero.

Un-weighted War-dates. Lastly, since we only have two military build-ups in our sample, i.e. the Vietnam war and the Carter-Reagan build-up, we repeat our baseline analysis without weighting the two events. The unweighted results are reported in Figure A48 and they are consistent with the baseline results shown in Figure A46.

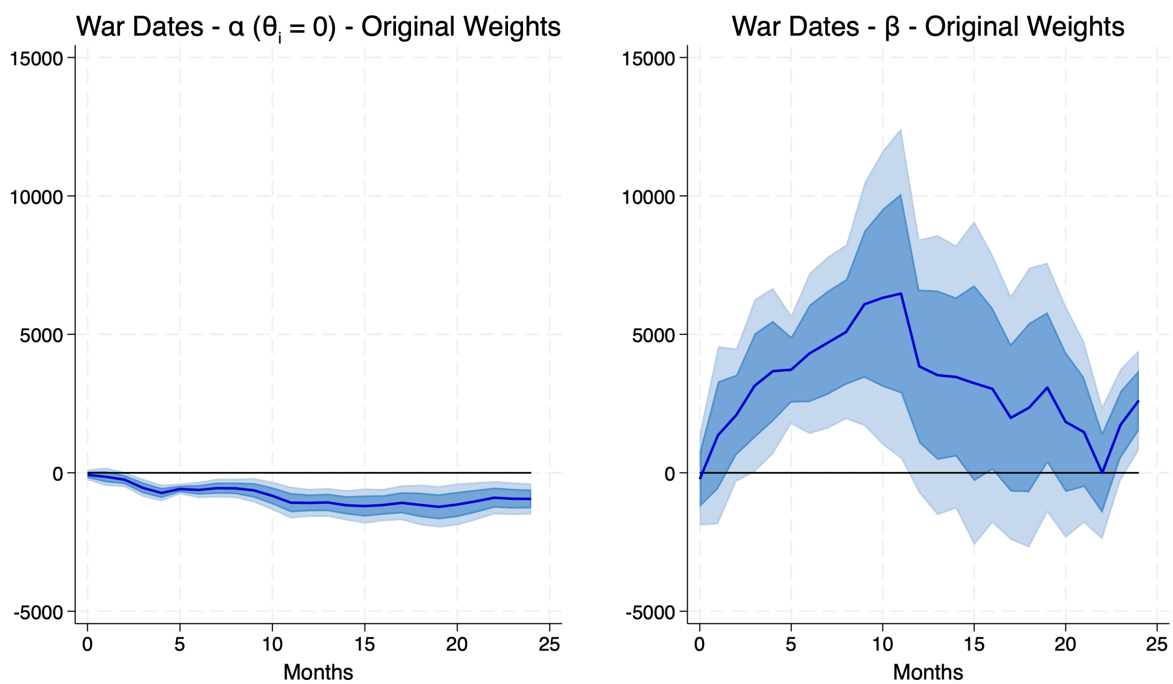


Figure A48: Response of Sectoral Inventories to War Events (Unweighted)

Notes: Same as in Figure A46.

References

Auerbach, **Alan J** and **Gorodnichenko**, **Yuriy** (May 2012). Measuring the Output Responses to Fiscal Policy. *American Economic Journal: Economic Policy*. 2012. 4, pp. 1–27.

Bils, **Mark** and **Cho**, **Jang-Ok** (1994). Cyclical Factor Utilization. *Journal of Monetary Economics*. 1994. 33, pp. 319–354.

Blanchard, **Olivier** and **Perotti**, **Roberto** (Nov. 2002). An empirical characterization of the dynamic effects of changes in government spending and taxes on output. *The Quarterly Journal of Economics*. 2002. 117, pp. 1329–1368.

Bouakez, **Hafedh**, **Rachedi**, **Omar**, and **Santoro**, **Emiliano** (Jan. 2023). The Government Spending Multiplier in a Multisector Economy. *American Economic Journal: Macroeconomics*. 2023. 15, pp. 209–239.

Brunet, **Gillian** (Mar. 2024). Stimulus on the Home Front: The State-Level Effects of WWII Spending. *Review of Economics and Statistics*. 2024, pp. 1–46.

Cox, **Lydia**, **Muller**, **Jernot J.**, **Pasten**, **Ernesto**, **Schoenle**, **Raphael**, and **Weber**, **Michael** (Sept. 2023). Big G. *Working Paper*. 2023.

Cox, **Lydia**, **Müller**, **Gernot J.**, **Pastén**, **Ernesto**, **Schoenle**, **Raphael**, and **Weber**, **Michael** (Oct. 2024). Big *G*. *Journal of Political Economy*. 2024. 132, pp. 3260–3297.

Dupor, **Bill** and **Guerrero**, **Rodrigo** (Dec. 2017). Local and aggregate fiscal policy multipliers. *Journal of Monetary Economics*. 2017. 92, pp. 16–30.

Eichenbaum, **Martin** and **Fisher**, **Jonas** (2005). Fiscal Policy in the Aftermath of 9/11. *Journal of Money, Credit, and Banking*. 2005. 37, pp. 1–22.

Fernald, **John** (Apr. 2012). A Quarterly, Utilization-Adjusted Series on Total Factor Productivity. *Federal Reserve Bank of San Francisco, Working Paper Series*. 2012, pp. 01–28.

Fisher, **Jonas** and **Peters**, **Ryan** (May 2010). Using Stock Returns to Identify Government Spending Shocks. *The Economic Journal*. 2010. 120, pp. 414–436.

Hall, **Robert** (1990). Invariance Properties of Solow's Productivity Residual. *NBER working paper*. 1990. 3034.

Hamilton, **James D.** (1994). *Time series analysis*. en. Princeton, N.J: Princeton University Press. isbn: 978-0-691-04289-3.

— (Dec. 2018). Why You Should Never Use the Hodrick-Prescott Filter. *The Review of Economics and Statistics*. 2018. 100, pp. 831–843.

Horowitz, **Karen** and **Planting**, **Mark** (Apr. 2009). *Concepts and Methods of the U.S. Input-Output Accounts*. en.

Ilzetzki, **Ethan** (Jan. 2023). Learning by Necessity: Government Demand, Capacity Constraints, and Productivity Growth. *Working Paper*. 2023.

McGrattan, **Ellen R.** and **Ohanian**, **Lee** (May 2010). Does Neoclassical Theory Account for the Effects of Big Fiscal Shocks? Evidence from World War II. *International Economic Review*. 2010. 51, pp. 509–532.

Nakamura, **Emi** and **Steinsson**, **Jón** (Mar. 2014). Fiscal Stimulus in a Monetary Union: Evidence from US Regions. *American Economic Review*. 2014. 104, pp. 753–792.

Nekarda, **Christopher** and **Ramey**, **Valerie** (Jan. 2011). Industry Evidence on the Effects of Government Spending. *American Economic Journal: Macroeconomics*. 2011. 3, pp. 36–59.

Perotti, **Roberto** (Jan. 2007). In Search of the Transmission Mechanism of Fiscal Policy [with Comments and Discussion]. *NBER Macroeconomics Annual*. 2007. 22, pp. 169–249.

— (May 2014). Defense Government Spending Is Contractionary, Civilian Government Spending Is Expansionary. *NBER Working Paper*. 2014, w20179.

Ramey, **Valerie** (June 1989). Inventories as Factors of Production and Economic Fluctuations. *American Economic Review*. 1989. 79, pp. 338–354.

— (Feb. 2011). Identifying Government Spending Shocks: It's all in the Timing*. *The Quarterly Journal of Economics*. 2011. 126, pp. 1–50.

— (2016). Macroeconomic Shocks and Their Propagation. In: *Handbook of Macroeconomics*. Vol. 2. Elsevier, pp. 71–162.

Ramey, **Valerie** and **Shapiro**, **Matthew** (1998). Costly capital reallocation and the effects of government spending. *Carnegie-Rochester Conference on Public Policy*. 1998. 48, pp. 145–194.

Ramey, **Valerie** and **Zubairy**, **Sarah** (Mar. 2018). Government Spending Multipliers in Good Times and in Bad: Evidence from US Historical Data. *Journal of Political Economy*. 2018. 126, pp. 850–901.

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

— (2004). A New Measure of Monetary Shocks: Derivation and Implications. *The American Economic Review*. 2004. 94, p. 64.

Wieland, **Johannes F.** and **Yang**, **Mu□Jeung** (Feb. 2020). Financial Dampening. *Journal of Money, Credit and Banking*. 2020. 52, pp. 79–113.