Why Does GDP Move Before Government Spending? It's All in the Measurement

Edoardo Briganti   Gillian Brunet *   Victor Sellemi

April 9, 2026

Abstract

Government spending is a long, complex process and not an event that occurs at a single point in time. In this paper, we argue that national income accounting (NIPA) measures government spending too late in the process to fully capture its economic effects, resulting in a systematic downward bias in aggregate time series estimates of the fiscal multiplier. This bias occurs because new government purchases initially show up in GDP as inventory investment rather than government purchases. We combine budget and contract data to produce an alternate measure of government spending based on authorizations. This new measure anticipates NIPA G by 3-4 quarters and allows for cleaner, more precise identification of fiscal shocks. We show that our new measure produces time series multiplier estimates of approximately 1 or higher at all time horizons.

(JEL E60, E62)

*Contacts, acknowledgments and disclaimer.

Briganti: Bank of Canada (email:`ebriganti@bankofcanada.ca`). Brunet: Smith College (`gbrunet@smith.edu`). Sellemi: University of California San Diego (email: `vsellemi@ucsd.edu`). *Disclaimer*: The views expressed here are those of the authors and do not necessarily reflect the views of the Bank of Canada or other members of its staff. We are particularly thankful to our advisors, Valerie Ramey and Christina Romer, for their help and support. We also acknowledge helpful comments from Karl Boulware, J. Bradford DeLong, Barry Eichengreen, Yuriy Gorodnichenko, James Hamilton, Juan Herreño, Nir Jaimovich, Marc Muendler, Emi Nakamura, Martha Olney, Sarah Quincy, David Romer, Allan Timmermann, Fabian Trottner, Johannes Wieland and anonymous referees. We are further indebted to Richard Kogan and Paul Van de Water for sharing invaluable expertise on the technical details of the federal budget process. We would also like to thank seminar participants at UC Berkeley, UC San Diego, the Federal Reserve Board of Governors, the Bank of Canada, the Banco de España, the University of Colorado Boulder, ISER Osaka University, the UC Davis Macrohistory Workshop, Smith College, the University of Naples Federico II, the Frankfurt School of Finance and the Kiel Institute for the World Economy. All errors are our own.

"We shall not, however, succeed in spending the £1,250 million this year, and some of the late Government's programme must necessarily roll forward into a future year. [...] I have never yet seen a munitions programme—and I have seen several—which did not lag behind the plans."

— Winston Churchill (December 6, 1951)[1]

I. Introduction

More than 15 years into the renaissance of macroeconomic research on fiscal policy, estimating the aggregate fiscal multiplier remains extraordinarily difficult. Reliable estimates are valuable for both policymakers and researchers. Yet the problem is inherently challenging because government spending is endogenous, and identification depends crucially on timing. Despite many excellent papers on the topic, firm conclusions about the size of the aggregate multiplier remain elusive.

We argue that aggregate fiscal multiplier estimates suffer from a fundamental measurement problem and show that correcting this problem has important implications for the size of the multiplier. The National Income and Product Accounts (NIPA) record government spending so that its timing is more closely aligned with the disbursement of government funds. For some applications, such as national income accounting, this treatment is appropriate. But if the goal is to measure the effects of government spending on economic activity, NIPA measures spending too late in the process. A significant share of government payments occurs only after firms have received a contract, hired workers, purchased materials, and produced goods. In the meantime, the associated activity appears in NIPA as inventory investment by firms producing for the government. When the firm eventually delivers the finished good, NIPA records a decline in inventories and an increase in government spending. As a result, NIPA can record government spending only after its direct effects on economic activity have already begun, and in some cases after much of the effect has already occurred.

The idea that government spending may be anticipated is, of course, not new. A large literature (including Ramey, 2011; Mertens and Ravn, 2010; Leeper, Walker, and Yang, 2013) studies anticipation effects and emphasizes that economic agents respond to fiscal news, with important implications for the identification of macroeconomic shocks. Our argument is different but complementary. We show that the timing delay built into NIPA provides a second source of anticipation effects. In this channel, GDP can respond before measured government spending not only because agents foresee future fiscal policy, but also because the underlying production process begins before NIPA records the spending itself.

This is not a minor accounting detail. It matters for the identification of fiscal shocks in U.S. aggregate data because a substantial share of variation in government spending is tied to defense spending, especially defense procurement spending (Cox et al., 2024). Moreover, the issue is likely relevant outside the United States as well, since countries that follow the System of National Accounts (SNA) apply a similar accounting treatment to defense procurement. Indeed, recent work by Alloza et al. (2026) documents a similar mechanism in Spanish defense procurement data. Therefore, the possibility that government spending is recorded with economically meaningful delays has implications that extend

[1] We thank Edward Nelson and Valerie Ramey for bringing this quote to our attention.

beyond the particular U.S. setting.

We start from the observation that government spending is a long and complex process rather than a single event. Economic activity associated with government spending begins relatively early in that process, when materials are purchased and goods are produced, not only when goods are delivered to the government and payments are disbursed. The time between contract award and completion can be long, and it can vary substantially across contracts depending on the scale of the project, the complexity of the goods involved, and production schedules and delays. The lag between authorizations and government purchases as recorded in NIPA is therefore both long and variable.

A new measure of military spending. We address this timing problem by introducing a timing-corrected measure of military spending which combines defense budget authorizations—budget authority—with military contract data. Budget authority has the advantage of covering all defense spending, including direct government spending such as military pay and production at government facilities. Its drawback is that it is available only at annual frequency. Military contracts are narrower than budget authority, but they are available at quarterly frequency. We use the quarterly variation in military contracts to interpolate quarterly variation within the annual totals implied by budget authority. We call the resulting measure *spending authorizations*. A major advantage of this measure is its flexibility: it can be combined with a wide range of existing approaches to fiscal multiplier estimation.

NIPA government spending can be understood as a delayed moving average of spending authorizations. We show that when the delay is severe enough, the delayed measure does not contain enough information to recover the underlying fiscal shocks. We then show that the delays in NIPA are severe enough to cross that threshold using annual data. At the quarterly frequency, we provide evidence that spending authorizations Granger-cause NIPA measures of spending, even after controlling for defense news shocks (from Ramey and Zubairy, 2018). By contrast, defense news shocks and spending authorizations appear to Granger-cause each other. Spending authorizations therefore appear to contain information that is distinct from existing measures of fiscal foresight.

We therefore use spending authorizations to construct unanticipated government spending shocks following the criteria in Ramey (2016). Crucially, we show that shocks to spending authorizations Granger-cause shocks to NIPA G, i.e., Blanchard and Perotti (2002) shocks (BP shocks), even when the NIPA shocks are orthogonalized by defense news shocks. This corroborates the view that anticipation effects have two mechanisms at work: one working through expectations and a distinct one working through delayed accounting of NIPA government spending.

The inventory channel. We further show that inventory investment is the main mechanism through which these measurement delays generate early GDP responses. Inventory investment responds quickly and positively to a shock to spending authorizations. The reason is that NIPA records the accumulation of private inventories while military contractors are producing goods that have not yet been completed and delivered to the government. In other words, the economic activity associated with the production of military goods initially enters GDP through inventories rather than through government purchases.

We show that estimating multipliers using BP shocks biases estimates downward because of these delays. Using BP shocks produces multiplier estimates between 0.5 and 0.8, consistent with the prior literature (e.g. Hall, 2009; Barro and Redlick, 2011). The response of inventories to spending-authorization shocks is much larger and more persistent than the response of inventories to BP shocks. We interpret this difference as evidence that BP shocks understate the fiscal multiplier because they miss the early inventory response to government spending and therefore miss part of the underlying spending process. Using spending authorizations solves these problems and allows for more accurate and more precise estimates of the fiscal multiplier than estimates based on NIPA.

A new decomposition of the multiplier. The timing mismatch between spending authorizations and NIPA also raises a conceptual issue for the definition of the multiplier itself. The cumulative multiplier is constructed by dividing the cumulative response of GDP by the cumulative response of NIPA G. This has important implications for multipliers estimated using spending authorizations. At short horizons, new authorizations rise much faster than NIPA G, so the denominator adjusts only slowly. This generates extremely large and imprecise multiplier estimates at horizons shorter than one year, a common problem affecting all timely measures of fiscal shocks (see Ramey, 2016). Over longer horizons, spending authorizations and NIPA gradually converge, so the problem becomes less severe.

In other words, using spending authorizations improves the identification of unanticipated shocks by accounting for measurement delays, but it does not by itself solve the denominator problem in the standard cumulative multiplier. Output is still divided by NIPA G, and the denominator remains affected by the same timing problem even when the shock is not. Under the textbook definition of the multiplier as the response of output *per dollar of spending as recorded in NIPA*, the large short-run multipliers are not conceptually wrong, but in practice they are extremely imprecise. However, if policymakers and researchers care about the response of output *per dollar authorized*, then the standard multiplier based on spending authorizations is biased upward at short horizons, because much of the authorized spending has not yet appeared in NIPA G.

To address this issue, we introduce an *accounting-corrected multiplier*, which measures the fiscal multiplier *per dollar authorized*. The accounting-corrected multiplier is not affected by timing delays. Under the assumption that every authorized dollar is ultimately spent, we show that the accounting-corrected multiplier can be written as a function of the multiplier on private spending (essentially GDP net of G) and the response of the residual component of NIPA G (the portion of G not accounted for by spending authorizations, including both non-defense government spending and consumption of fixed capital) to spending authorizations. Unlike the traditional fiscal multiplier, the accounting-corrected multiplier is well defined even at very short horizons. We estimate an accounting-corrected multiplier of 1.2 on impact. It peaks at roughly 1.25 after three to four quarters and then gradually declines toward 1. The accounting-corrected multiplier is precisely estimated at all horizons and avoids the bias created by timing delays.

Organization. The rest of the paper proceeds as follows. Section II explores the sources of anticipation effects and shows why inventories provide a second, complementary channel for understanding why GDP responds before NIPA G. Section III describes the government spending process and the construction of our military-contract and defense-authorization measures. Section IV uses (defense) spending authorizations to identify unanticipated government spending shocks and presents the corresponding impulse responses. Section V shows that failing to account for the timing delay in NIPA government spending understates the fiscal multiplier. It then introduces the *accounting-corrected fiscal multiplier*. Section VI concludes.

II. Motivation: Why Does GDP Rise Before G?

The seminal contributions of Ramey and Shapiro (1998) and Ramey (2011) show that government spending is often anticipated by private agents. This, in turn, has important implications for identification. Mertens and Ravn (2010) show that anticipation can render standard SVARs misspecified, because measured spending innovations may combine unanticipated and anticipated structural shocks. Related work by Leeper, Walker, and Yang (2013) shows that, in the presence of fiscal foresight, structural fiscal shocks need not be recoverable from current and past observed macroeconomic variables alone.

Identifying unanticipated fiscal shocks therefore requires separating realized movements in government spending from the expectations that preceded them. One way to do so is to measure expectations directly (Ramey, 2016). For example, Ramey (2011) constructs a narrative series of defense news shocks that measures expected changes in defense spending. Figure 1 replicates Ramey (2016) and plots the first eight quarters of the impulse response functions (IRFs) of GDP and government spending, G, in response to a defense news shock. The figure shows that GDP rises before G: output responds to defense news before measured government spending does. In other words, the fiscal foresight problem is directly visible in the data.

Why does GDP rise before measured government spending in response to fiscal news? Existing work has mainly interpreted this pattern through economic theory. In that view, forward-looking agents anticipate higher future government spending and higher future taxes, adjust their behavior in advance, and generate the familiar negative wealth effect before the spending itself is recorded in the national accounts.

This explanation is important, but it is not the whole story. We argue that the timing gap between GDP and G also reflects the way defense spending moves through production and is recorded in the national accounts. To motivate this argument, we show two facts. First, the early response of GDP to defense news is concentrated in inventories. Second, defense news is followed immediately by new spending authorizations, new military contracts, and a prompt production response. Taken together, these facts point to a complementary channel that runs through the defense spending process itself and through the accounting treatment of work in progress.

To shed light on the first fact, Figure 2 decomposes the response of GDP to a defense news shock into

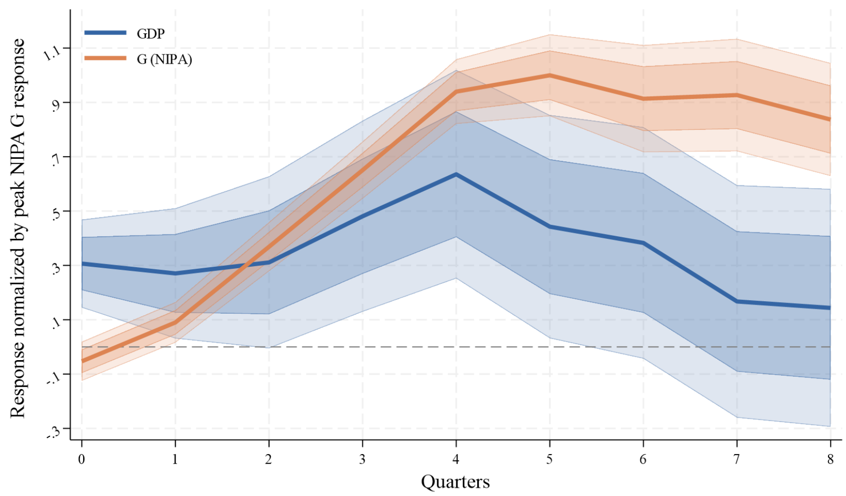

FIGURE 1 — GDP MOVES FASTER THAN G IN RESPONSE TO DEFENSE NEWS
EVIDENCE OF FISCAL FORESIGHT

Notes: IRFs are estimated using lag-augmented LPs. Variables include defense news shocks, GDP, G, TB3, and R&R tax shocks updated using information from Liu and Williams (2019). Nominal variables are transformed using the Gordon-Krenn transformation. The specification includes four lags of each variable and a quadratic time trend. The price deflator is the GDP price deflator. The sample runs from 1947Q1 to 2015Q4, reflecting the availability of the defense news shock series in Ramey and Zubairy (2018). Confidence bands are 68% and 90% and are based on robust standard errors (Montiel Olea and Plagborg-Møller, 2021).

its major components. Each series is normalized by the peak response of NIPA G to the news shock, so all responses are expressed in comparable units.

The figure shows that the early positive response of GDP, shown in blue, is concentrated in inventories, shown in orange in the top-left panel. This is a striking pattern. Inventories are a small and volatile component of investment, and they have received relatively little attention in the fiscal-policy literature. Yet in the data they account for most of the initial rise in GDP after a defense news shock.

To motivate the second fact, Figure 3 traces variables that lie closer to the underlying defense spending process.

The left panel reports the IRFs of military contracts and (defense) spending authorizations from the federal budget, whose construction is described in detail in the next section. On impact, a defense news shock triggers both new military contract awards and new spending authorizations, and both series respond faster than NIPA G. The right panel shows the response of military production, proxied by average weekly hours of production workers in the aircraft industry, which provides a timely measure of defense production dynamics.[2] This measure also responds promptly.

The empirical evidence suggests that at least part of the early response of GDP is not necessarily a forward-looking reaction to future spending. Defense news initiates a new cycle of authorization, contracting, and production. Because military procurement often involves complex goods such as aircraft and missiles, the production process is lengthy. During that interval, the national accounts record

[2] See Briganti (2023).

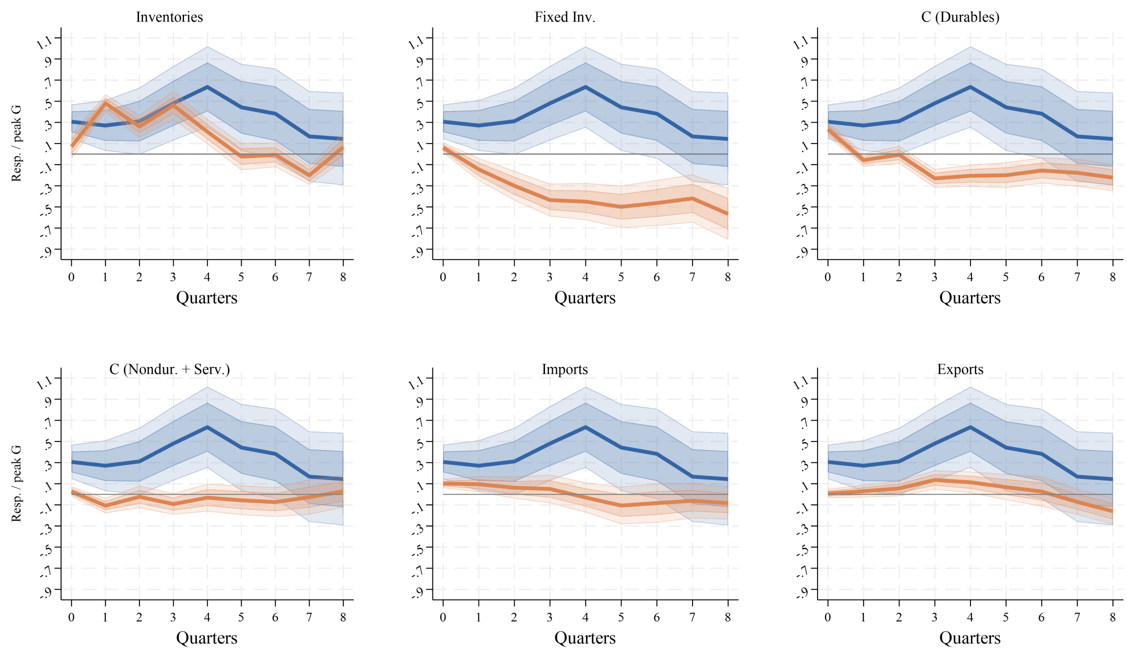

FIGURE 2 — INVENTORIES DRIVE THE GDP RESPONSE TO DEFENSE NEWS

Notes: The blue IRF is GDP. The orange IRF is the indicated GDP component. Each component enters the specification with four lags as a rotating outcome variable. Everything else is identical to Figure 1.

the associated activity as private inventory accumulation, which is part of GDP. Only when the good is completed and delivered to the government do the national accounts record a decline in inventories and a corresponding increase in defense procurement spending, a component of NIPA G. As a result, GDP can rise before measured G even when both movements are generated by the same underlying spending process.

Our argument is therefore that there are two complementary channels at work. Defense news can affect activity through expectations, and it can also set in motion a defense production process that appears first in inventories and only later in NIPA G. The next section develops this second channel in detail by describing how defense procurement is recorded in the national accounts and by constructing the military-contract and spending-authorization series used in the rest of the paper.

III. Why Inventories? It Is in the Measurement

Under the *System of National Accounts* (SNA 2008), output is recorded when production occurs rather than when a good is sold. For goods whose production spans multiple periods, the unfinished portion is recorded as *work-in-progress* (WIP) and enters GDP through *changes in inventories* (SNA 2008).[3] This principle ensures that value added is allocated to the period in which production takes place, indepen-

[3] See United Nations (2009), *System of National Accounts 2008*, Chapter 6, paras. 6.110–6.113, esp. pp. 108–109.

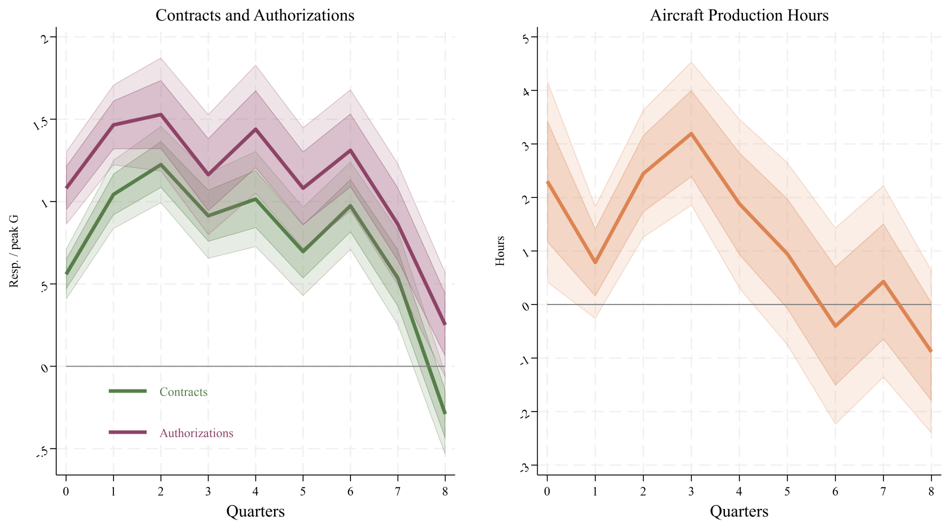

FIGURE 3 — MILITARY CONTRACTS, AUTHORIZATIONS, AND PRODUCTION IMMEDIATELY REACT TO DEFENSE NEWS

Notes: For average weekly hours of aircraft production workers, the sample ends in 2001Q4 because data from the discontinued BLS hours-worked database are unavailable after that date. Everything else is identical to Figure 1.

dently of the timing of delivery or payment.

In the United States, the *National Income and Product Accounts* (NIPAs) follow this production-based principle. As stated in Chapter 7 of the NIPA Handbook:

> "A general principle underlying NIPA accounting is that production should be recorded at the time it occurs. [...] The recording of movements of goods in inventory — materials and supplies, work-in-process, and finished goods — and from inventories to final sales provides the means to allocate production to the period in which it occurred."

When ownership does not transfer until completion—as is often the case for complex military equipment—the output generated during the production phase is recorded as additions to the producer's inventories. Upon delivery and transfer of ownership, the withdrawal from inventories is offset by the corresponding sale, so that total output remains distributed over the production period (SNA 2008, paras. 6.112–6.113). Since SNA 2008 capitalizes military weapons systems as fixed assets of government (Chapter 10, para. 10.87), the delivery of a completed aircraft or missile system is recorded as government gross fixed capital formation (GFCF). In the United States, this corresponds to NIPA federal government gross investment.

In expenditure terms, this accounting treatment implies the following composition effect. During the production phase, GDP may increase through private inventory investment. At delivery, inventories decline and government investment rises by a corresponding amount. Aggregate GDP is unaffected by this reclassification at the time of delivery, but its composition shifts from private inventory accumulation to government investment.

This mechanism was already emphasized by Hickman (1955) in his analysis of the Korean War buildup (pp. 18–19):[4]

> "It is apparent that a defense mobilization will provide a stimulus to economic expansion if government expenditures increase the aggregate demand for goods and services. However, the expansion need not await the actual growth of government expenditures. In the first place, some government expenditures for defense will lag behind the placement of orders. For some time, the increase in production that follows orders will be reflected in private inventory investment rather than in government expenditures."

Consistent with this description, the BEA defines private inventories to include work-in-process (NIPA Handbook, Chapter 2) and records changes in private inventories as part of gross private domestic investment. For federal defense equipment, the BEA's government transactions methodology specifies that long-term production items—such as aircraft and missiles—are treated as business work-in-progress until delivery, at which point government investment is recorded (BEA, *Concepts and Methods of the U.S. NIPAs*, Chapter 9; BEA MP-5, pp. II-11):[5]

> "The largest timing difference is for national defense gross investment for relatively long-term production items, such as aircraft and missiles, for which the work in progress is considered as part of business inventories until the item is completed and delivered to the Government."

Consequently, during periods of rapid military buildup, GDP can increase before measured government purchases or investment rise, reflecting the accumulation of work-in-progress by domestic producers.

This accounting mechanism is consistent with the empirical evidence presented in the previous section. Following a defense news shock, contract awards increase, initiating a new wave of multi-period production. This production phase is reflected in the strong response of inventories, even before government spending rises.

If the time lag between contract awards and recorded government expenditure is economically meaningful, it should become visible during large swings in defense activity. In what follows, we test empirically for the existence and magnitude of this delay. Specifically, we construct a measure of military contract awards and one of NIPA defense procurement spending to document directly the time gap between the initiation of production (contract award) and the recording of government investment in NIPA (i.e., at delivery and transfer of ownership).

[4]See Hickman (1955), *The Korean War and United States Economic Activity, 1950–1952*.
[5]See BEA MP-5, Government Transactions.

NIPA Defense Procurement Spending. In the defense procurement process, contract awards and spending represent two distinct stages. The process begins with the award of a contract, at which point the government creates a legal obligation for the contractor to deliver a good or provide a service, and commits to pay for it. Following the award, contractors initiate a production process that may span several years.[6]

In the NIPAs, defense procurement spending is not reported as a standalone series. Instead, it is embedded within government consumption expenditures and gross investment (G), one of the components of GDP measured from the expenditure side. To isolate defense procurement spending, we follow Cox et al. (2024) and construct it by summing (i) NIPA defense consumption expenditures on intermediate goods and services and (ii) NIPA defense gross investment in structures, equipment, and software.[7] The resulting series captures the flow of government expenditures associated with defense procurement and allows us to compare directly the timing of recorded spending in the NIPAs with the timing of contract awards.

III.1. A New Series of Defense Contracts

We construct a new quarterly series of military contracts beginning in 1947. Contracts measure the dollar value of prime contracts awarded by the Department of Defense (DoD) to private firms. Contract values are assigned to the award date (typically at the monthly or quarterly frequency). As such, contracts capture the initiation of production and therefore precede the recording of government expenditures in the NIPAs.

To assemble a continuous quarterly time series, we combine military contract data from multiple sources, each covering different time periods. A brief overview of the sources and procedures is provided below; additional details are reported in Appendix A.

(i) 1947Q1–1950Q4. We do not observe direct data on military contract awards for 1947Q1–1950Q4. To fill this gap, we estimate contracts over this period using two variables that are strongly correlated with military contract activity: (i) average weekly hours of production workers in aircraft manufacturing and (ii) NIPA defense procurement spending, as constructed above. Average weekly hours in aircraft manufacturing are available at the monthly frequency from the BLS (archived series). Aircraft production hours closely comove with military contracts because aircraft and related components account for a substantial share of defense procurement.[8]

Using observed military contracts from 1951Q1 to 1980Q4 (described below), we estimate the fol-

[6] Contract-level data from the post-2000 period indicate that the mean and median duration of $1 defense procurement contracts are 4.2 and 5.4 years, respectively. We measure duration as the period of performance, defined as the number of days between the award date and the contract end (i.e., full delivery) date. Total defense procurement spending is dominated by a small number of very large, long-duration contracts, consistent with Cox et al. (2024).

[7] Further details on the accounting construction of procurement spending are provided in Appendix C.1.

[8] See Appendix A.1 for details.

lowing equation by OLS:

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

which delivers an in-sample R^2 of 64.8%.

We then extrapolate contracts for 1947Q1–1950Q4 using the estimated coefficients and the observed values of aircraft production hours and NIPA defense procurement. The extrapolated series rises sharply at the onset of the Korean War in 1950. The timing of this increase is driven primarily by aircraft production hours and is consistent with contemporaneous narratives in the *Survey of Current Business* (August 1950). Additional details are provided in Appendix A.1. Because the post-1951 estimation sample relies on military contracts data from Business Conditions Digest (BCD), we refer to this extrapolated segment as *BCD Extrapolated*.

(ii) 1951Q1–1982Q4. For 1951Q1 through 1982Q4, we use military contract data from Ramey (1989). The original sources are the *Business Conditions Digest* and, prior to 1961, *Business Cycle Developments* (collectively referred to as "BCD").[9] BCD reports seasonally adjusted monthly military contracts from January 1951 through November 1988.[10]

(iii) 1983Q1–2003:3. For 1983Q1 to 2003:3, we rely on the annual *Federal Procurement Summary Report* (FPSR), produced by the Directorate for Information Operations and Reports (DIOR). FPSR reports the annual value of military contract awards by fiscal year. It also provides bar charts with quarterly values of total federal procurement contracts (defense and non-defense combined). On average, approximately 80% of total federal procurement consists of military contracts. We interpolate the annual military procurement series to quarterly frequency using the quarterly distribution of total federal contracts.

(iv) 2003Q4 onward. Beginning in 2003Q4, all federal procurement transactions are reported at the daily level in the Federal Procurement Data System–Next Generation (FPDS-NG). We aggregate military contracts to the fiscal-year level and then interpolate to quarterly frequency using the within-year variation in newly awarded military contracts, following a methodology analogous to that used for the FPSR data.

We adopt this approach rather than aggregating directly at the quarterly frequency because contracts are occasionally modified or canceled in subsequent quarters, which can artificially inflate high-frequency volatility. Aggregating at the fiscal-year level and interpolating reduces this noise while preserving the annual totals.

[9] Military contracts became part of the set of instruments known as the Hall–Ramey instruments; see Hall (1990) and Ramey (1989).

[10] The series was discontinued in 1988 and partially migrated to the *Survey of Current Business* (SCB) in 1990. SCB data are available from January 1990 through September 1995 but systematically omit the fourth quarter of each year. Because of these gaps, SCB data are less reliable. To maximize data quality, we use BCD data through 1982Q4 and switch to FPSR data when available. Using BCD data through 1988 yields similar but less precise results.

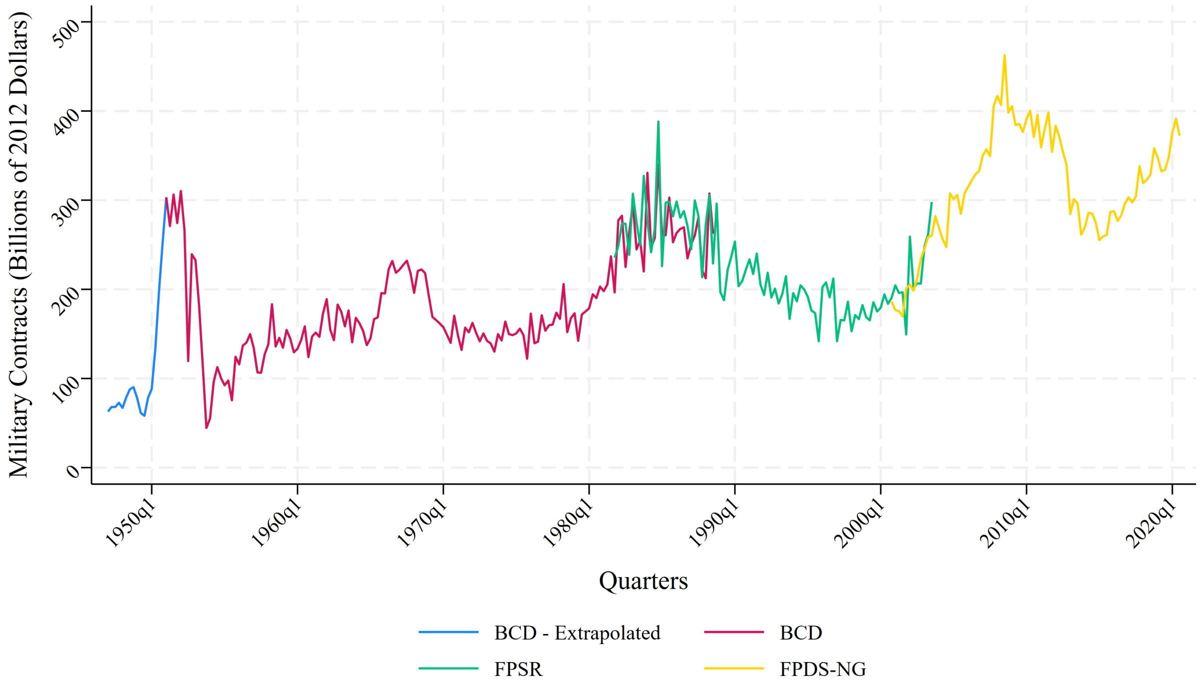

FIGURE 4 — MEASURES OF MILITARY CONTRACTS

Notes: The price deflator is the GDP price deflator. FPSR and FPDS-NG data are seasonally adjusted using the Census Bureau's X-13 ARIMA-SEATS procedure. Data are annualized.

Assembling the Time Series: Military Contracts, 1947–2019. Figure 4 displays real military contracts (in billions of 2012 dollars) constructed from the sources described above. Quarterly amounts are annualized.

When two measures are available for the same quarter (e.g., during the 1980s and early 2000s), we display both. Discrepancies across measures are minimal, indicating strong consistency across sources. Because the alternative measures are nearly identical, the choice of source does not materially affect our results. In overlapping periods, we rely on the source we consider most accurate and complete. Throughout the remainder of the paper, we refer to this constructed series as *military contracts*, or simply *contracts*. Figure 5 plots our newly constructed series of military contracts (red line) alongside NIPA defense procurement spending (blue line). Vertical dashed lines refer to events that shifted the trajectory of U.S. defense policy and are discussed in more detail in the following subsection.

We recognize some disparities in the two series during and after the Korean War period, likely due to the broad awarding of contracts, subsequent cancellations, and the accounting methods utilized by the DoD before McNamara's term.[11] Nonetheless, the two series appear to track each other well, reconciling the micro contract data with national accounts. In Appendix A.4, we show that our contract series, aggregated to the fiscal-year level, closely tracks the data used by Dupor and Guerrero (2017), which

[11]We appreciate Emi Nakamura for pointing out this issue, initially identified during the drafting of Nakamura and Steinsson (2014).

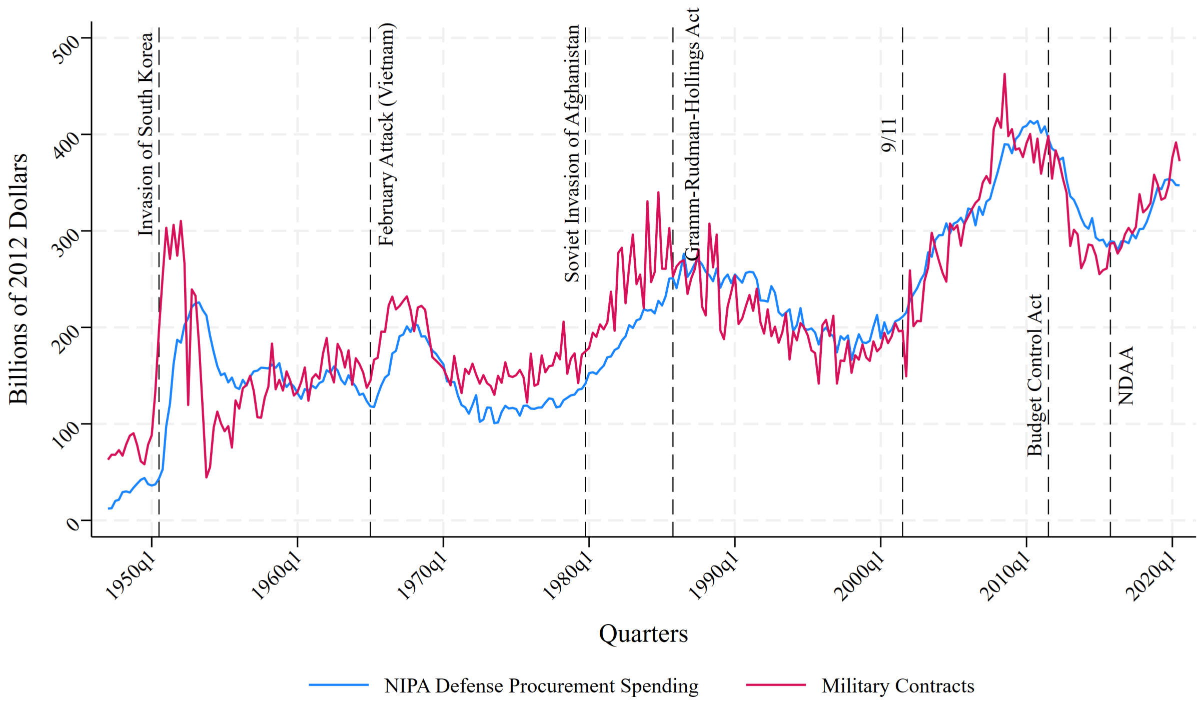

FIGURE 5 — TWO NOVEL SERIES OF DEFENSE PROCUREMENT

Notes: The price deflator is the GDP price deflator. The price deflator before 1947Q1 is from Ramey and Zubairy (2018), adjusted to use 2012 as the base year. Data for calendar year 1946 are linearly interpolated.

extend that used by Nakamura and Steinsson (2014).

III.2. Contracts Lead NIPA

We now document the existence of a time delay in NIPA defense procurement spending relative to military contracts using our newly constructed data. We begin with a graphical analysis around major defense build-ups and drawdowns to provide direct and intuitive evidence of lead–lag patterns between the two series.

Korean War. The top-left panel of Figure 6 shows real per capita NIPA defense procurement spending (blue) and defense contracts (red) around the outbreak of the Korean War. The vertical dashed line marks the invasion of South Korea in June 1950, which triggered the U.S. military escalation (see Ramey and Shapiro, 1998). Thick lines denote fiscal-year averages, while thin lines represent quarterly values.

Contracts rise sharply in 1950:Q3 (fiscal year 1951), whereas NIPA defense procurement responds more gradually and smoothly. The subsequent drawdown also occurs earlier and more abruptly in contracts than in NIPA.[12] This visual evidence is consistent with the accounting mechanism described at

[12]We acknowledge the sizable discrepancies between the two series during and immediately after the Korean War. These differences likely reflect widespread contract awards, subsequent cancellations, and specific DoD accounting practices prior

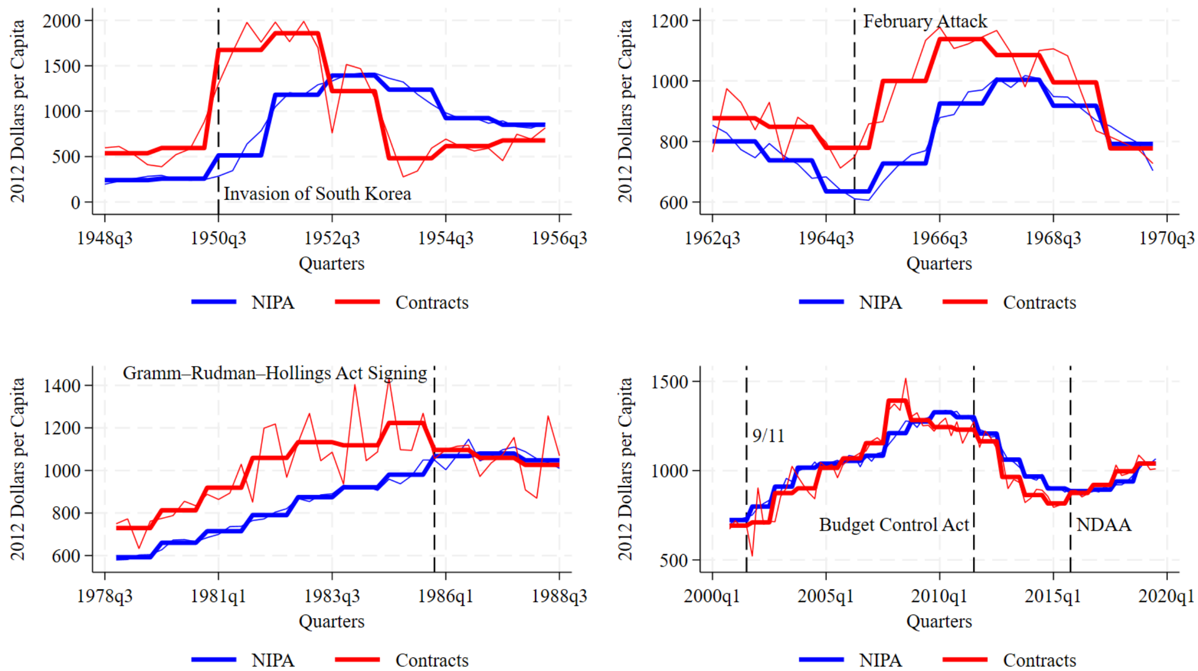

FIGURE 6 — MILITARY CONTRACTS LEAD NIPA DEFENSE PROCUREMENT SPENDING GRAPHICAL ANALYSIS

the beginning of the section: contracts mark the initiation of production, while NIPA records expenditures as production is completed and ownership is transferred.

Vietnam War. The top-right panel presents the same comparison during the Vietnam War buildup. The vertical line marks February 1965, when the attack on U.S. barracks in Pleiku preceded a major escalation of U.S. involvement (see Ramey and Shapiro, 1998). As in the Korean episode, contracts increase more rapidly, while NIPA defense procurement exhibits a slower adjustment, a pattern that remains visible even at the fiscal-year frequency.

End of the Carter–Reagan Buildup. The bottom-left panel illustrates the late Carter and Reagan defense buildup and its subsequent reversal. Because this episode involved a more gradual expansion, the measurement delay is less visually striking during the buildup phase, although contracts grow somewhat faster than NIPA, widening the gap between the two series.

The reversal phase, however, is sharper in contracts. Contracts begin to decline earlier and more abruptly, coinciding with the enactment of the Balanced Budget and Emergency Deficit Control Act of 1985 (commonly known as the Gramm–Rudman–Hollings Act), signed in December 1985. The Act

to the McNamara reforms. We thank Emi Nakamura for highlighting this issue, originally noted in the context of Nakamura and Steinsson (2014).

introduced binding deficit-reduction targets and automatic spending cuts (sequestration), which immediately affected defense programs. NIPA defense procurement declines more gradually thereafter.

Post-2000 Period. The bottom-right panel plots the evolution of the two series following the September 11, 2001 terrorist attacks. Both contracts and NIPA defense procurement increase steadily, but contracts peak in fiscal year 2008, whereas NIPA peaks roughly two years later.

Following the Budget Control Act of 2011 and the implementation of sequestration in 2013, both series decline. Subsequently, the rise of ISIS and Russia's annexation of Crimea in 2014 intensified concerns in Washington regarding U.S. military readiness after several years of constrained budgets. These developments were frequently cited in congressional debates as evidence of a deteriorating strategic environment. They contributed to bipartisan support for the Bipartisan Budget Act of 2015, which raised discretionary spending caps, and for the National Defense Authorization Act (NDAA) for Fiscal Year 2016. Defense spending increased beginning in FY2016 (2015:Q4). As in previous episodes, contracts responded promptly, whereas NIPA defense procurement adjusted more gradually.

Lead-Lag Correlation. To formalize the graphical evidence, Figure 7 reports a lead-lag correlation map between year-over-year quarterly changes in real per capita contracts and NIPA defense procurement spending over the full sample (1947:Q1-2019:Q4).

The correlation is positive in the quadrant where contracts lead NIPA, and it reaches its maximum when NIPA is lagged by approximately three quarters. This pattern suggests an average measurement delay of about three quarters between contract awards and recorded defense procurement spending in the NIPAs.

We can summarize the findings of this section as follows:

STYLIZED FACT 1: *Military contracts lead NIPA defense procurement spending, consistent with substantial time-to-build in military production.*

III.3. The Role of Budget Authorizations

Before the federal government actually spends money, Congress and the executive branch partake in a formal budgeting process. While the federal budget process has changed over time, some aspects of it (e.g., appropriations, the President's budget) have remained the same since the 1920s. As part of this process, Congress establishes *budget authority* for each account in the federal budget. Specifically, budget authority is the legal authority for a government entity (e.g., the Department of Defense) to make contracts to spend money on a specific program or purpose.[13]

Establishing budget authority is the first step in the government spending process. Therefore, budget authority may anticipate new contracts awards. We test this hypothesis empirically, by constructing a novel time series of budget authority.

[13] Budget authority is granted for the federal fiscal year; money does not need to be spent by the end of the year, but it must be obligated (legally committed) by the end of the year or it is returned to the Treasury's general fund.

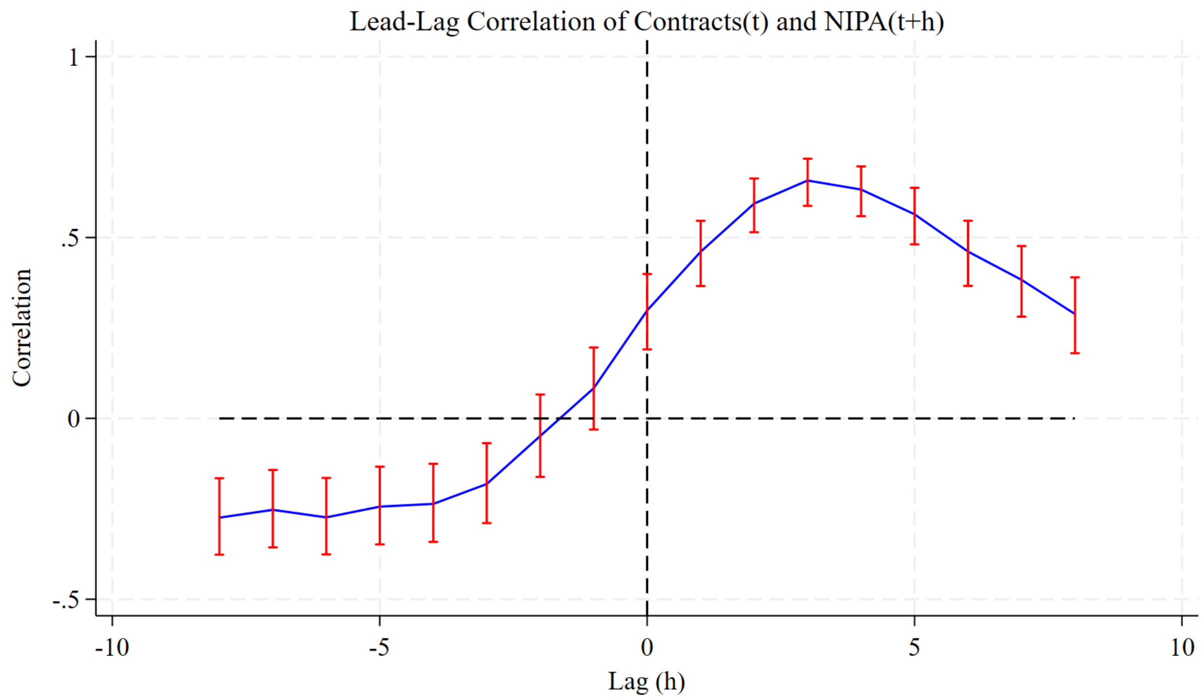

FIGURE 7 — MILITARY CONTRACTS LEAD NIPA DEFENSE PROCUREMENT SPENDING
LEAD-LAG CORRELATION MAP

Notes: Lead-lag correlation map between year-over-year quarterly changes in real per capita military contracts and NIPA defense procurement spending, computed over the full sample 1947:Q1-2019:Q4.

Data. Annual budget authority totals are available from the Office of Management and Budget (OMB) for 1976 onward, disaggregated by category of spending. We construct a historical series of defense budget authority for 1938-1975 using the *Budgets of the United States*, which have been issued annually since 1923.[14] Figure 8 shows in red the time series of (defense) budget authority by fiscal year, in real dollars, starting from 1947, when standardized NIPA data becomes available.

Budget authority is similar to but broader than appropriations. $1 of appropriations directly creates $1 of budget authority, but budget authority also includes government spending not funded through the annual appropriations process (i.e., "mandatory" or "entitlement" spending on programs like Social Security).[15] Over time, however, $1 of budget authority translates fairly directly into $1 of spending in NIPA.[16] The NIPA counterparts of budget authority is NIPA defense spending, net of cost of fixed

[14]We begin our series in 1938 because prior to the New Deal budget totals were regarded as upper bounds rather than as targets, so budgetary information for earlier years cannot be interpreted in the same way. See Schick (1990) for the institutional history of this change.

[15]In addition, in the (relatively rare) instances when money is appropriated but never obligated (committed) or spent before the appropriation expires, budget authority will show a (negative) adjustment for the unspent funds, formally called a rescission. Unlike appropriations, budget authority incorporates rescissions, which can sometimes be important (e.g., the billions of dollars in canceled military contracts at the end of WWII are recorded as a reduction in budget authority, but would not be accounted for in appropriations).

[16]There are exceptions, primarily involving things like funding for revolving loan accounts. However, these exceptions are

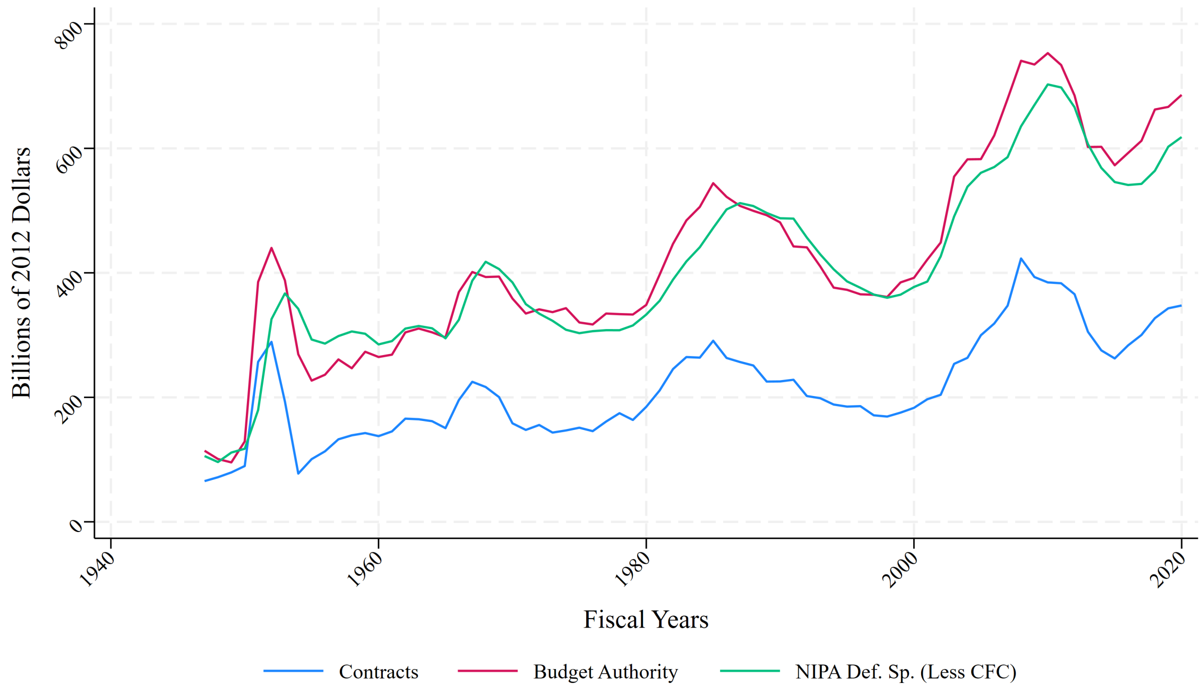

FIGURE 8 — BUDGET AUTHORITY, MILITARY CONTRACTS AND NIPA DEFENSE SP. (NET OF CFC)

capital (CFC), a small fraction of spending which accounts for public capital depreciation. Figure 8 shows in green the time series of real NIPA defense spending net of CFC. The two time series coincide, confirming the reconciliation between budget to national accounts data. Figure 8 also displays in blue the fiscal year values of our military contracts series.

It is possible to grasp from the graph that contracts and budget authority comove, while NIPA defense spending (less CFC) appears to move with delay. This is confirmed in Figure 9: the top panel of the figure displays the lead-lag correlation map of real per capita changes in military contracts and budget authority. The correlation is almost one when the two time series are aligned ($h = 0$), and decays when the two time series are mis-aligned.

On the contrary, the bottom panel of Figure 9 shows the lead-lag correlation map between budget authority and its NIPA counterpart, both measured in real per capita changes, and the correlation in this case spikes when NIPA is delayed by one year. Therefore, the measurement delay between NIPA and contracts appear to persist even at annual frequency and at higher aggregation levels, that is, when looking at budget authority, which is a more comprehensive measure of defense spending, instead of contracts.

STYLIZED FACT 2: *Budget authority co-moves tightly with military contracts and leads NIPA defense spending (net of CFC) at fiscal-year frequency, indicating that the measurement delay persists even in broader annual aggregates.*

almost exclusively concentrated in non-defense spending.

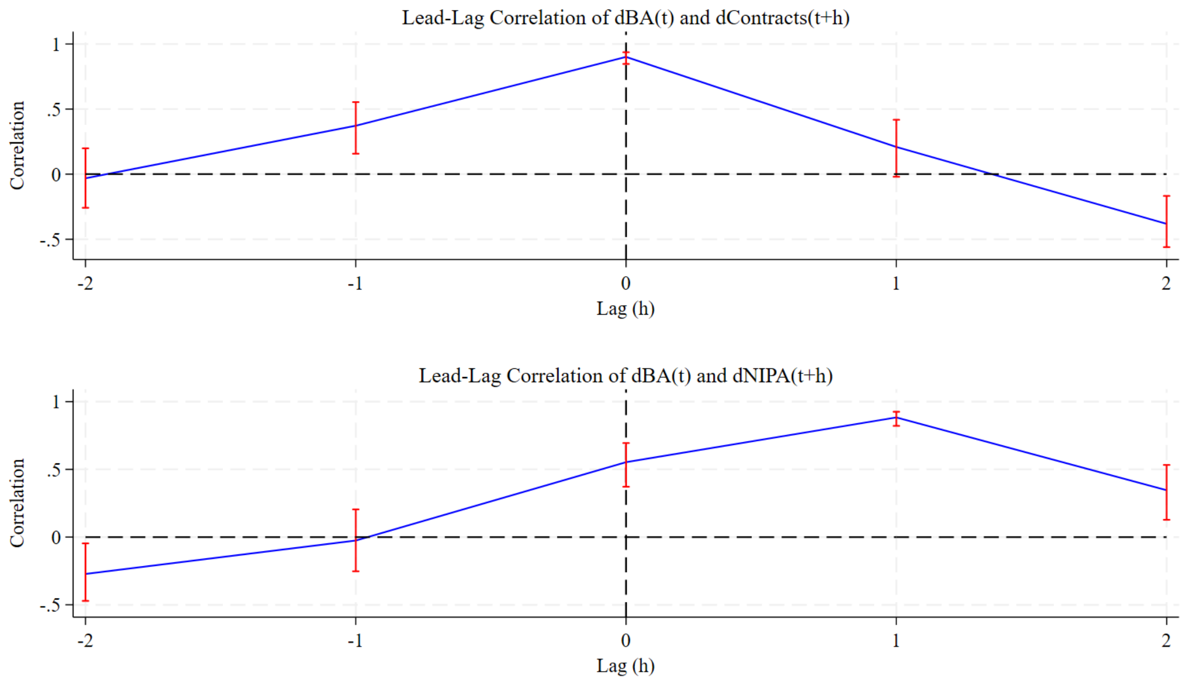

FIGURE 9 — BUDGET AUTHORITY LEADS NIPA DEF.SP. AND CO-MOVES WITH CONTRACTS LEAD-LAG CORRELATION MAP

Notes: Lead-lag correlation map between year-over-year changes in real per capita military contracts, budget authority and NIPA defense spending (net of CFC), computed over the full sample FY1947-FY2019.

IV. Identification of Unanticipated Shocks via Contracts and Authorizations

The previous sections establish that NIPA defense procurement spending adjusts with a delay relative to military contracts. We also show that this delay is visible at higher levels of aggregation: NIPA defense spending (net of CFC) lags behind defense budget authority. Moreover, one dollar of contracts and one dollar of budget authority each eventually translates into one dollar of NIPA spending. These results imply that NIPA spending behaves as a delayed moving average of contracts and budget authority.

In this section, we show that when the delay is sufficiently persistent, current and lagged NIPA spending do not contain enough information to recover the underlying structural fiscal shocks at annual frequency. If the measurement delay is already visible in annual data, then it is likely to be economically meaningful at higher frequencies as well, since short delays tend to be attenuated by annual aggregation. Indeed, we also show that, at quarterly frequency, military contracts and a new measure of spending authorizations—–constructed from contracts and budget authority–—Granger-cause the corresponding NIPA measures of spending. Therefore, the measurement delay has direct implications for the identification of unanticipated government spending shocks, which are typically extracted from NIPA spending series.

Illustrative model. To clarify the identification problem, we introduce a purely illustrative stylized model and then test its key implication using fiscal-year data.

Assume that the government spending process begins with newly awarded military prime contracts, denoted by MPC_t, which follow an AR(1) with white-noise innovations:

$$\text{MPC}_t = \rho \cdot \text{MPC}_{t-1} + \eta_t, \qquad \eta_t \sim_{\text{iid}} WN(0, \sigma_\eta^2). \tag{2}$$

For simplicity, we assume that all variables have already been orthogonalized with respect to news, so that η_t captures only unanticipated innovations to contract awards.

To capture delayed recording in the national accounts, suppose that NIPA defense procurement spending, denoted by G_t, reflects deliveries (i.e., transfers of ownership) such that a fraction $1-\lambda$ of contract activity is recorded contemporaneously and a fraction λ is recorded with a one-period lag:

$$G_t = (1-\lambda) \cdot \text{MPC}_t + \lambda \cdot \text{MPC}_{t-1}, \qquad \lambda \in [0, 1]. \tag{3}$$

When $\lambda \approx 0$, there is little measurement delay and $G_t \approx \text{MPC}_t$. When $\lambda \approx 1$, NIPA spending is approximately a one-period lagged transformation of contracts, $G_t \approx \text{MPC}_{t-1}$.[17]

Combining (2) and (3), G_t follows an ARMA(1,1) process:[18]

$$G_t = \rho \cdot G_{t-1} + \underbrace{(1-\lambda)\eta_t + \lambda\eta_{t-1}}_{:=\xi_t}. \tag{4}$$

Thus, the innovation in NIPA spending, ξ_t, is a weighted average of current and past contract-award shocks. In this setup, the moving-average component is invertible if and only if $\lambda < 1/2$. To see this, rewrite (4) as

$$(1-\rho L)G_t = (1-\lambda)\left(1 + \frac{\lambda}{1-\lambda}L\right)\eta_t.$$

Therefore,

$$\eta_t = \frac{1}{1-\lambda}\left(1 + \frac{\lambda}{1-\lambda}L\right)^{-1}(1-\rho L)G_t.$$

The inverse $\left(1 + \frac{\lambda}{1-\lambda}L\right)^{-1}$ exists only if $\left|\frac{\lambda}{1-\lambda}\right| < 1$, i.e., if $\lambda < 1/2$.

Hence, when measurement delays are severe ($\lambda > 1/2$), lags of NIPA G do not contain sufficient information to recover the underlying contract shocks η_t. This is the familiar *non-fundamentalness* problem emphasized in the fiscal policy literature when identifying government spending shocks using the standard NIPA measure: in presence of fiscal foresight, innovations to NIPA G are a combinations of anticipated and unanticipated government spending shocks (Leeper, Walker, and Yang, 2013; Ramey, 2011; Mertens and Ravn, 2010; Forni and Gambetti, 2016; Ascari et al., 2023). Here the non-

[17] This distributed-lag representation is motivated by the *time-to-spend* mechanism discussed by Leeper, Walker, and Yang (2010) and Ramey (2021) in the context of infrastructure spending.

[18] From (2), $\text{MPC}_t = (1-\rho L)^{-1}\eta_t$. From (3), $G_t = ((1-\lambda) + \lambda L)\text{MPC}_t$. Substituting and rearranging yields (4).

fundamentalness problem arises by construction of NIPA government spending, as a consequence of accounting practices which generates *measurement delays*.

We test empirically the degree of measurement delays by estimating Equation (3) using fiscal year data, whose dynamics can be decently approximated by AR(1) models.[19] Table 1 reports the estimation results.

TABLE 1 — ESTIMATION OF EQUATION (3) - MEASUREMENT DELAYS ARE SEVERE: $\lambda > \frac{1}{2}$

Dependent Variable: G_t (NIPA defense procurement spending)								
	(A) Full Sample (Baseline)				(B) No Extrapolated Data and Korean War			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
MPC_{t+1}			0.009				-0.023	
			(0.109)				(0.176)	
MPC_t $(1-\lambda)$	0.028	0.133	0.019	0.108	0.384	0.431	0.412	0.411
	(0.096)	(0.092)	(0.148)	(0.101)	(0.163)	(0.158)	(0.272)	(0.176)
MPC_{t-1} (λ)	0.774	0.867	0.778	0.777	0.532	0.569	0.523	0.505
	(0.094)	(0.092)	(0.107)	(0.093)	(0.161)	(0.158)	(0.175)	(0.172)
NEWS_{t-1}				-0.099				-0.077
				(0.039)				(0.071)
T	72	72	72	68	64	64	64	60
Sample	1947-2019	1947-2019	1947-2019	1947-2015	1956-2019	1956-2019	1956-2019	1956-2015
Adjusted R^2	66.45%	-	65.96%	69.23%	70.34%	-	69.85%	70.32%

Dependent Variable: G_t (NIPA defense spending less CFC)								
	(A) Full Sample (Baseline)				(B) No Extrapolated Data and Korean War			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
BA_{t+1}			-0.008				-0.211	
			(0.067)				(0.121)	
BA_t $(1-\lambda)$	0.126	0.255	0.135	0.173	0.210	0.337	0.485	0.250
	(0.057)	(0.066)	(0.097)	(0.060)	(0.112)	(0.131)	(0.192)	(0.126)
BA_{t-1} (λ)	0.656	0.745	0.652	0.635	0.560	0.663	0.473	0.530
	(0.055)	(0.066)	(0.064)	(0.055)	(0.112)	(0.131)	(0.120)	(0.125)
NEWS_{t-1}				-0.065				-0.011
				(0.030)				(0.072)
T	72	72	72	68	64	64	64	60
Sample	1947-2019	1947-2019	1947-2019	1947-2015	1956-2019	1956-2019	1956-2019	1956-2015
Adjusted R^2	88.17%	-	88.00%	89.15%	83.10%	-	83.65%	83.87%

Notes: All variables are in real per capita values. Column (1): estimate Equation (3). Column (2): estimate Equation (3) with constrained coefficients summing to one. Column (3): estimate Equation (3) while controlling for leads of contracts/authorizations. Column (4): estimate Equation (3) while controlling for lags of defense news shocks.

Columns (1) and (2) present OLS estimates under the unconstrained and constrained specifications, respectively, where the constraint forces the coefficients to sum to one, as in the model. Data are in real per capita values, aggregated by fiscal year. In Panel (A), which uses the full sample, the coefficient on lagged contracts is larger in magnitude and more statistically significant than the coefficient on con-

[19] On the contrary, quarterly data and hump-shaped dynamics require higher order models to be well approximated.

temporaneous contracts. The implied estimate of the delay parameter λ exceeds 0.50 in the constrained specification, indicating that the moving-average component of NIPA spending is non-invertible and therefore subject to a non-fundamentalness problem.

For robustness, Panel (B) repeats the analysis using a sample that begins in fiscal year 1956, which serves two purposes: it excludes the lower-quality contract observations obtained via extrapolation and it removes the influential Korean War surge, consistent with evidence that this period is highly influential in defense-spending data (see Perotti, 2014; Ramey, 2016; Dupor and Guerrero, 2017). Results are robust.

Moreover, Column (3) augments the baseline specification with leads of contracts. We find no evidence that NIPA spending predicts future contract awards. Column (4) adds lagged defense news shocks to the regression. Even after controlling for news, lagged contracts retain strong predictive power for NIPA defense procurement spending. Lastly, the bottom panel of the Table replicates the same results using (defense) budget authority (BA) and NIPA defense spending (net of CFC). The results are identical: measurement delays appear to be severe enough to prevent the extraction of fiscal shocks from annual NIPA data.

IV.1. Granger-causality Tests on Raw Government Data

We formalize the measurement delay in NIPA using also standard Granger-causality tests. We report results for two samples and two aggregation levels. Our baseline sample uses the full available period. For robustness, we also consider a sample starting in 1956, for the same reasons discussed above.

We conduct the tests using both quarterly data and fiscal-year data, from contracts and budget authority. Quarterly data provide higher frequency variation but are also more volatile. We therefore also report results at the fiscal-year frequency, using two lags of each variable to mimic the conventional eight quarterly lags used in the test.

Defense Spending Authorizations. To perform Granger-causality tests using budget authority at quarterly frequency, we build a new quarterly series of defense spending authorizations. In the rest of the paper we will refer to this series as *spending authorizations*, or simply authorizations. The series is constructed by interpolating the original budget authority values with quarterly variation from military contracts, such that the fiscal year averages of spending authorizations return the original annual budget authority data. The time series of spending authorizations and its detailed construction is displayed in Appendix B.

Is NIPA Spending Predicted? The top panel of Table 2 reports the results for contracts, for both sample and aggregation levels. The bottom panel reports the results using spending authorizations: the fiscal year values of spending authorizations refer to budget authority, while the quarterly values refer to the interpolated series of budget authority and military contracts.

Results from the full sample are clear: military contracts Granger-cause NIPA defense procurement spending, and spending authorizations predict NIPA defense spending net of CFC, both at quarterly

TABLE 2 — CONTRACTS PREDICT NIPA *G*

(1) Do Contracts Predict NIPA?						
Predicted	*Predictor*	*Frequency*	*Sample*	*F*	*pvalue*	*Predict?*
NIPA Def. Procurement Sp.	Military Contracts	Calendar Quarter	1947-2019	10.60	0.0000	Yes
Military Contracts	NIPA Def. Procurement Sp.	Calendar Quarter	1947-2019	1.30	0.2453	No
NIPA Def. Procurement Sp.	Military Contracts	Fiscal Year	1947-2019	49.92	0.0000	Yes
Military Contracts	NIPA Def. Procurement Sp.	Fiscal Year	1947-2019	0.99	0.3776	No
NIPA Def. Procurement Sp.	Military Contracts	Calendar Quarter	1956-2019	1.98	0.0500	Yes
Military Contracts	NIPA Def. Procurement Sp.	Calendar Quarter	1956-2019	2.27	0.0233	Yes
NIPA Def. Procurement Sp.	Military Contracts	Fiscal Year	1956-2019	7.42	0.0013	Yes
Military Contracts	NIPA Def. Procurement Sp.	Fiscal Year	1956-2019	0.77	0.4671	No

(2) Do Authorizations Predict NIPA Defense Spending Less CFC?						
Predicted	*Predictor*	*Frequency*	*Sample*	*F*	*pvalue*	*Predict?*
NIPA Def. Sp. Less CFC	Spending Authorizations	Calendar Quarter	1947-2019	14.31	0.0000	Yes
Spending Authorizations	NIPA Def. Sp. Less CFC	Calendar Quarter	1947-2019	1.51	0.1534	No
NIPA Def. Sp. Less CFC	Spending Authorizations	Fiscal Year	1947-2019	61.87	0.0000	Yes
Spending Authorizations	NIPA Def. Sp. Less CFC	Fiscal Year	1947-2019	0.11	0.8987	No
NIPA Def. Sp. Less CFC	Spending Authorizations	Calendar Quarter	1956-2019	3.09	0.0024	Yes
Spending Authorizations	NIPA Def. Sp. Less CFC	Calendar Quarter	1956-2019	2.09	0.0375	Yes
NIPA Def. Sp. Less CFC	Spending Authorizations	Fiscal Year	1956-2019	7.29	0.0015	Yes
Spending Authorizations	NIPA Def. Sp. Less CFC	Fiscal Year	1956-2019	2.13	0.1275	No

Notes: We implement Granger-causality tests in Stata using the `var` command and the post-estimation command `vargranger`. In the bivariate specifications, we estimate a VAR with eight lags. We report small-sample adjusted test statistics using the `dfk` and `small` options. Operationally, the test is equivalent to (i) regressing the predicted variable on its own lags and on the lags of the predictor, and (ii) conducting a Wald test of the joint null that the predictor's lags are all zero.

and annual frequency. The reverse is not true.

The results are similar in the sample excluding the Korean War. When the data are aggregated to the fiscal-year frequency, contracts and authorizations clearly predict their NIPA counterparts, but not vice versa. This is consistent with the graphical evidence discussed in the previous section: around large swings in defense spending, NIPA visibly lags behind contracts and authorizations. Examples include the outbreak of the Vietnam War, the onset of budget sequestration after the Carter–Reagan buildup, and the increase in spending following the signing of the NDAA in November 2015. Nonetheless, the quarterly results are more mixed. In the restricted sample starting in 1956, both contracts and authorizations appear to predict their NIPA counterparts, and vice versa.

One possible explanation is that quarterly variation makes the series more volatile, which reduces the ability of Granger-causality tests to isolate the measurement delay that is clearly visible in the same sample at annual frequency. At the same time, quarterly data may capture different information sets precisely because of their higher frequency. However, papers that use NIPA to identify unanticipated government spending shocks typically report highly persistent impulse responses of government spending.[20] This pattern is more consistent with the view that the mixed quarterly Granger-causality results reflect the difficulty of isolating the measurement delay in noisier quarterly data than with the view that quarterly variation reveals a fundamentally different source of unanticipated spending shocks. Taken together, these results suggest that researchers should be cautious when using NIPA to identify unanticipated government spending shocks, especially given that at annual frequency the same NIPA series appears to contain anticipatory information. In the next section, we show that unanticipated government spending shocks extracted from NIPA G by conditioning on a broader set of lags of endogenous variables—Blanchard and Perotti (2002) shocks—are Granger-caused by shocks to spending authorizations even in the post-Korean War sample, suggesting that the additional lagged information helps isolate the measurement delay in quarterly data.

Are Contracts and Authorizations Predicted by News? Table 3 examines the relationship between military contracts, spending authorizations, and the defense news shock series of Ramey (2011), updated by Ramey and Zubairy (2018). We estimate a bivariate VAR in defense news shocks and military contracts and test for Granger-causality in both directions. The results point to bidirectional predictability. This finding indicates that the two series share some information, but are not informationally equivalent: each contains variation not captured by the other.

We interpret this pattern as consistent with the two series capturing different types of fiscal shocks. Defense news shocks are designed to measure anticipated changes in government spending. By contrast, contracts and authorizations are informative about unanticipated, or surprise, changes in spending. Their ability to forecast subsequent movements in NIPA spending does not arise from foresight about future government purchases, but from measurement delay in the construction of NIPA. In other words, contracts and authorizations lead NIPA because NIPA records procurement spending with a lag, not because these series proxy for anticipated fiscal policy. Since anticipated and unanticipated government spending shocks can have different macroeconomic effects (see Mountford and Uhlig, 2009), the two series are useful for addressing complementary questions.

NIPA, news, and contracts/authorizations. Finally, the top panel of Table 4 reports Granger-causality tests from a trivariate VAR that includes NIPA defense procurement spending, defense news shocks and military contracts. The bottom panel reports the same Granger-causality tests from a trivariate VAR that includes NIPA defense spending (net of CFC), defense news shocks and spending authorizations.

The key result is that contracts and authorizations continue to predict NIPA defense procurement spending even after controlling for lagged defense news shocks. There is only one caveat: contracts and NIPA defense procurement spending, orthogonalized by lags of defense news, do not predict each other in the post-Korean war sample. However, contracts predict NIPA defense procurement at fiscal

[20]For instance, Barnichon, Debortoli, and Matthes (2022) identify unanticipated shocks using NIPA G ordered after a measure of expected government spending designed to absorb fiscal foresight (Auerbach and Gorodnichenko, 2012), using a sample starting in 1966. The reported impulse response function of G is highly persistent, suggesting that the identified shock loads primarily on persistent, low-frequency variation in NIPA G, rather than on transitory quarterly movements.

TABLE 3 — CONTRACTS/AUTHORIZATIONS AND NEWS CAPTURE DIFFERENT INFORMATION SETS

(1) Do News Predict Contracts?						
Predicted	*Predictor*	*Frequency*	*Sample*	*F*	*pvalue*	*Predict?*
Defense News Shocks	Military Contracts	Calendar Quarter	1947-2015	2.55	0.0110	Yes
Military Contracts	Defense News Shocks	Calendar Quarter	1947-2015	8.62	0.0000	Yes
Defense News Shocks	Military Contracts	Calendar Quarter	1956-2015	2.06	0.0406	Yes
Military Contracts	Defense News Shocks	Calendar Quarter	1956-2015	1.72	0.0960	Yes

(2) Do News Predict Authorizations?						
Predicted	*Predictor*	*Frequency*	*Sample*	*F*	*pvalue*	*Predict?*
Defense News Shocks	spending authorizations	Calendar Quarter	1947-2015	4.25	0.0001	Yes
spending authorizations	Defense News Shocks	Calendar Quarter	1947-2015	4.45	0.0000	Yes
Defense News Shocks	spending authorizations	Calendar Quarter	1956-2015	2.30	0.0217	Yes
spending authorizations	Defense News Shocks	Calendar Quarter	1956-2015	2.18	0.0300	Yes

year frequency, even in this sample; moreover, spending authorizations predict NIPA defense spending less CFC, in the post-Korean war sample.

IV.2. Empirical Specification and Fiscal Shocks

Given the long and variable measurement delays in NIPA defense spending, our preferred strategy for identifying unanticipated government spending shocks is to use innovations in our new quarterly series of spending authorizations rather than innovations in measured NIPA G. Authorizations capture the fiscal shock closer to the point at which resources are committed, lead the corresponding movements in NIPA defense spending, and cover a broader share of defense spending than contracts alone. For these reasons, we use authorization-based shocks as our baseline throughout the paper and view them as the natural starting point for future work on unanticipated defense spending shocks whenever comparable authorization data are available. Results using military contracts are analogous and are reported in Appendix D.

Baseline Specification. We estimate the impulse response functions of different components of GDP in response to a shock to spending authorizations using lag-augmented local projections (Jordà, 2005; Montiel Olea and Plagborg-Møller, 2021). Our estimating equation adopts the widely used Hall-Barro-Redlick (HBR) transformation, where each variable measured in dollars X_t is first differenced and divided by lagged GDP:

$$x_t := \frac{X_t - X_{t-1}}{\text{GDP}_{t-1}}.$$

TABLE 4 — CONTRACTS/AUTHORIZATIONS PREDICT NIPA G EVEN AFTER ORTHOGONALIZING FOR LAGS OF NEWS

(1) Do Contracts Predict NIPA Defense Procurement Spending?

Predicted	*Predictor*	*Frequency*	*Sample*	*F*	*pvalue*	*Predict?*
NIPA Def. Procurement Sp.	Military Contracts	Calendar Quarter	1947-2015	4.26	0.0001	Yes
Military Contracts	NIPA Def. Procurement Sp.	Calendar Quarter	1947-2015	0.81	0.5965	No
NIPA Def. Procurement Sp.	Military Contracts	Fiscal Year	1947-2015	17.44	0.0000	Yes
Military Contracts	NIPA Def. Procurement Sp.	Fiscal Year	1947-2015	0.34	0.7119	No
NIPA Def. Procurement Sp.	Military Contracts	Calendar Quarter	1956-2015	1.62	0.1200	No
Military Contracts	NIPA Def. Procurement Sp.	Calendar Quarter	1956-2015	1.60	0.1276	No
NIPA Def. Procurement Sp.	Military Contracts	Fiscal Year	1956-2015	5.74	0.0055	Yes
Military Contracts	NIPA Def. Procurement Sp.	Fiscal Year	1956-2015	0.11	0.8997	No

(2) Do Authorizations Predict NIPA Defense Spending Less CFC?

Predicted	*Predictor*	*Frequency*	*Sample*	*F*	*pvalue*	*Predict?*
NIPA Def. Sp. Less CFC	Spending Authorizations	Calendar Quarter	1947-2015	5.25	0.0000	Yes
Spending Authorizations	NIPA Def. Sp. Less CFC	Calendar Quarter	1947-2015	1.25	0.2711	No
NIPA Def. Sp. Less CFC	Spending Authorizations	Fiscal Year	1947-2015	19.63	0.0000	Yes
Spending Authorizations	NIPA Def. Sp. Less CFC	Fiscal Year	1947-2015	0.04	0.9594	No
NIPA Def. Sp. Less CFC	Spending Authorizations	Calendar Quarter	1956-2015	2.74	0.0068	Yes
Spending Authorizations	NIPA Def. Sp. Less CFC	Calendar Quarter	1956-2015	1.06	0.3922	No
NIPA Def. Sp. Less CFC	Spending Authorizations	Fiscal Year	1956-2015	5.89	0.0049	Yes
Spending Authorizations	NIPA Def. Sp. Less CFC	Fiscal Year	1956-2015	0.72	0.4898	No

Moreover, our estimating equation is expressed in long differences (Piger and Stockwell, 2025):[21]

$$\frac{Y_{t+h} - Y_{t-1}}{\text{GDP}_{t-1}} = \beta^h \cdot \frac{\text{SA}_t - \text{SA}_{t-1}}{\text{GDP}_{t-1}} + (\text{Lagged Controls}) + \varepsilon_{t+h}, \tag{5}$$

where the OLS estimate of β^h represents the impulse response function at horizon h, to an innovation in the HBR-transformed value of spending authorizations (SA_t).

Our standard conditioning set of variables includes: GDP, NIPA government spending, the three-month Treasury bill rate (TB3, to control for monetary policy), and a measure of tax shocks to control for changes in taxes. We use the narrative series of exogenous tax shocks from Romer and Romer (2010), extended by Liu and Williams (2019). Following standard practices, we include four quarters of lagged values for our standard set of controls. When the outcome variable is not GDP, government spending or spending authorizations, we also include lagged values of the outcome. All nominal variables are

[21] Formally, the control lags are the following:
$\sum_{j=1}^{4}\left(\beta_j^h \cdot \frac{\text{SA}_{t-j}-\text{SA}_{t-1-j}}{\text{GDP}_{t-1-j}} + \gamma_j^h \cdot \frac{\text{G}_{t-j}-\text{G}_{t-1-j}}{\text{GDP}_{t-1-j}} + \phi_j^h \cdot \frac{\text{GDP}_{t-j}-\text{GDP}_{t-1-j}}{\text{GDP}_{t-1-j}} + \rho_j^h \cdot \frac{\text{Y}_{t-j}-\text{Y}_{t-1-j}}{\text{GDP}_{t-1-j}} + \pi_j^h \cdot \text{TB3}_{t-j} + \tau_j^h \cdot \frac{\text{Tax}_{t-j}^{\text{R\&R10}}}{\text{GDP}_{t-1-j}}\right)$.

deflated by the GDP price deflator.

The estimated IRFs (β_h) are asymptotically equivalent to those obtained from a recursive VAR with HBR-transformed variables, which orders spending authorizations first, treating it as an internal instrument for government spending (Ramey and Zubairy, 2018; Stock and Watson, 2018; Plagborg-Møller and Wolf, 2021).[22]

Sample. Our baseline sample spans 1947Q1–2017Q4, matching the extended tax series. For robustness, we also consider a sample that excludes the influential Korean War years, starting in 1956Q1.

Fiscal Shocks Construction. Given our specification, we essentially construct government spending shocks using forecast errors of spending authorizations:

$$\eta_t^{SA} := sa_t - \text{Proj}\left(sa_t \mid \underbrace{sa_{t-1}, y_{t-1}, g_{t-1}, TB3_{t-1}, \tau_{t-1}}_{\text{Lag 1}}, \dots, \underbrace{sa_{t-4}, y_{t-4}, g_{t-4}, TB3_{t-4}, \tau_{t-4}}_{\text{Lag 4}} \right),$$

where $Proj(x_t \mid \cdot)$ denotes the linear projection (OLS fitted value) of x_t onto the included variables (see Brockwell and Davis, 1991; Hamilton, 1994). sa_t is the HBR-transformed series of spending authorizations. g_t and y_t are the HBR-transformed values of NIPA government spending and GDP, and τ_t is the extended series of narrative tax shocks.

These shocks meet the criteria for identifying structural shocks laid out in Ramey (2016): they are (i) uncorrelated with contemporaneous and lagged values of endogenous variables, (ii) uncorrelated with other shocks, and (iii) unanticipated. Controlling for lags of endogenous variables meets condition (i) by construction. The plausible exogeneity of shocks to military spending fulfills condition (ii). The Granger-causality tests reported in Tables 2-4 corroborate the validity of assumption (iii).

In the remainder of the paper we will use the innovations in spending authorizations, η_t^{SA}, to estimate the dynamic effects of fiscal shocks as well as fiscal multipliers.

Granger-causality Tests on Identified Fiscal Shocks. The Granger-causality tests reported in Tables 2-4 are a useful and parsimonious description of the timing relationships in the raw time series data, suggesting a specific recursive ordering of the variables, useful to corroborate our identification assumption (iii). However, they do not represent a formal test of non-fundamentalness: in fact, it may well be that lags of the endogenous variables in our conditioning set may contain enough information to eliminate the predictive power of authorizations and/or contracts.

Therefore, before delving into the dynamic causal responses of spending authorizations, we perform Granger-causality tests to show that even after conditioning on our set of lagged variables, defense spending authorization shocks predict shocks identified using NIPA government spending. In

[22] We prefer local projections over VAR for two reasons. First, VARs appear to have worse small sample properties than LPs (Montiel Olea, Plagborg-Møller, et al., 2025). Second, fiscal multipliers are estimated with LP-IV methods, which are numerically equivalent to the ratio of the area underlying the estimated cumulative impulse response functions of GDP and G, estimated with LP, thus allowing for a direct mapping between impulse responses and multiplier estimates (Ramey, 2016; Stock and Watson, 2018; Ramey and Zubairy, 2018).

particular, we construct shocks to NIPA G mirroring a recursive identification à la Blanchard and Perotti (2002), in which forecast errors (i.e., innovations) to government spending are treated as structural shocks:

$$\xi_t^{BP} := g_t - \text{Proj}\left(g_t \;\middle|\; \underbrace{y_{t-1}, g_{t-1}, TB3_{t-1}, \tau_{t-1}}_{\text{Lag 1}}, \dots, \underbrace{y_{t-4}, g_{t-4}, TB3_{t-4}, \tau_{t-4}}_{\text{Lag 4}} \right),$$

Therefore, we run Granger-causality tests between the spending authorizations shocks described above, η_t^{SA}, against the BP shocks, ξ_t^{BP}. Results are reported in the top panel of Table 5 for both the baseline sample and the sample without the Korean war period in it, for robustness. This approach builds on Ramey (2011), who showed that BP shocks were Granger-caused by war dates. It is also similar in spirit to the non-fundamentalness test, implemented and discussed in Forni and Gambetti (2016).

TABLE 5 — FISCAL-SHOCK GRANGER-CAUSALITY TESTS: SPENDING-AUTHORIZATION SHOCKS

(1) Do Spending-Authorization Shocks Predict BP shocks?						
Predicted	*Predictor*	*Frequency*	*Sample*	*F*	*p-value*	*Predict?*
ξ_t^{BP} (BP Shocks)	η_t^{SA} (Spending-Authorizations Shocks)	Calendar Quarter	1947-2017	22.21	0.0000	Yes
η_t^{SA}	ξ_t^{BP}	Calendar Quarter	1947-2017	0.40	0.8099	No
ξ_t^{BP}	η_t^{SA}	Calendar Quarter	1956-2017	4.41	0.0019	Yes
η_t^{SA}	ξ_t^{BP}	Calendar Quarter	1956-2017	0.25	0.9121	No
(2) Do Spending-Authorization Shocks Predict BP shocks, Even After Controlling for News?						
ξ_t^{BP}	η_t^{SA}	Calendar Quarter	1947-2015	3.16	0.0147	Yes
η_t^{SA}	ξ_t^{BP}	Calendar Quarter	1947-2015	0.99	0.4125	No
ξ_t^{BP}	η_t^{SA}	Calendar Quarter	1956-2015	3.74	0.0058	Yes
η_t^{SA}	ξ_t^{BP}	Calendar Quarter	1956-2015	0.43	0.7838	No

Results are clear: BP shocks are predicted by the spending authorization shocks in both samples, but not vice versa. For robustness, Appendix D reports the analogue of Table 5 using shocks identified with military contracts. The results are identical: BP shocks are predicted by shocks identified with military contracts data.

The Role of Fiscal Foresight. The bottom panel of the Table reports the same Granger-causality tests, after augmenting the specification with the updated series of defense news shocks from Ramey and Zubairy (2018). This is to account for fiscal foresight, given that the raw time series of spending authorizations and defense news shocks appear to Granger-cause each others (Table 4). We find that even after controlling for lags of defense news shocks, spending authorizations shocks preserve predictive power over the BP shocks.

This finding matters for identification strategies that attempt to recover unanticipated government spending shocks using innovations in NIPA government spending, orthogonalized by measures of expected spending such as defense news shocks or forecasts from the *Survey of Professional Forecasters* (SPF). For example, Auerbach and Gorodnichenko (2012) augment a VAR with forecasts of government spending growth and interpret innovations in G orthogonalized by these forecasts as unanticipated government spending shocks.[23]

Our results complement the fiscal-foresight mechanism emphasized by Ramey (2011). Defense news are informative about anticipated changes in government spending, and orthogonalizing with respect to news is useful for purging the anticipatory component of fiscal shocks. Our point is that a separate accounting mechanism is also at work: NIPA defense procurement spending is partly a delayed reflection of earlier authorizations and contract awards. As a result, even after controlling for expectations, innovations in NIPA spending still combine current and lagged realizations of the same underlying surprise fiscal shock. This is why authorizations provide a cleaner measure of unanticipated government spending shocks.

STYLIZED FACT 3: *spending authorizations retain predictive power for NIPA spending and BP shocks even after conditioning on defense news shocks.*

IV.3. Empirical Results

Figure 10 plots the impulse responses (β^h) of nine macroeconomic variables in response to a shock to spending authorizations, using our baseline sample 1947Q1–2017Q4. For comparability across outcomes, all values are normalized by the peak response of NIPA G. Bands represent 68% and 90% confidence levels and standard errors are heteroskedasticity robust. We use robust standard errors following the recommendation in Montiel Olea and Plagborg-Møller (2021). We specifically avoid using HAC standard errors (e.g. Newey-West), given that in small sample applications they are subject to negative bias (Herbst and Johannsen, 2024).

Discussion of GDP Components. The top panel shows the impulse responses of GDP (left), NIPA government purchases (center) and spending authorizations (right). Both GDP and NIPA government purchases rise quickly—immediately, in the case of GDP—and remain elevated for four years after the initial shock. GDP responds immediately to spending authorizations—significantly faster than it responds to NIPA G. NIPA G peaks 7 to 11 quarters after the shock to spending authorizations.

The middle panel shows the impulse response of consumption, as broken down into durables (left) and non-durables plus services (center). Durables consumption rises significantly in the immediate aftermath of the shock before falling to zero.[24] In contrast, non-durables and services consumption

[23]This approach has been widely adopted; see Forni and Gambetti (2016), D'Alessandro, Fella, and Melosi (2019), Barnichon, Debortoli, and Matthes (2022), and Jørgensen and Ravn (2022), among others. Related strategies use SPF forecast revisions or forecast errors (e.g., Ricco, Callegari, and Cimadomo, 2016; Auerbach and Gorodnichenko, 2012).

[24]The negative response of durables from quarters 6 to 9 may reflect the outbreak of the Korean War: households anticipated shortages of goods, as they had during WWII, and bought in advance of anticipated shortages (see Hickman, 1955; Ramey, 2016; Binder and Brunet, 2021). In samples beginning after 1950, the dip in durables consumption in quarters 6 to 9 is not statistically significant. Notably, however, the initial increase in durables consumption also lacks statistical significance when

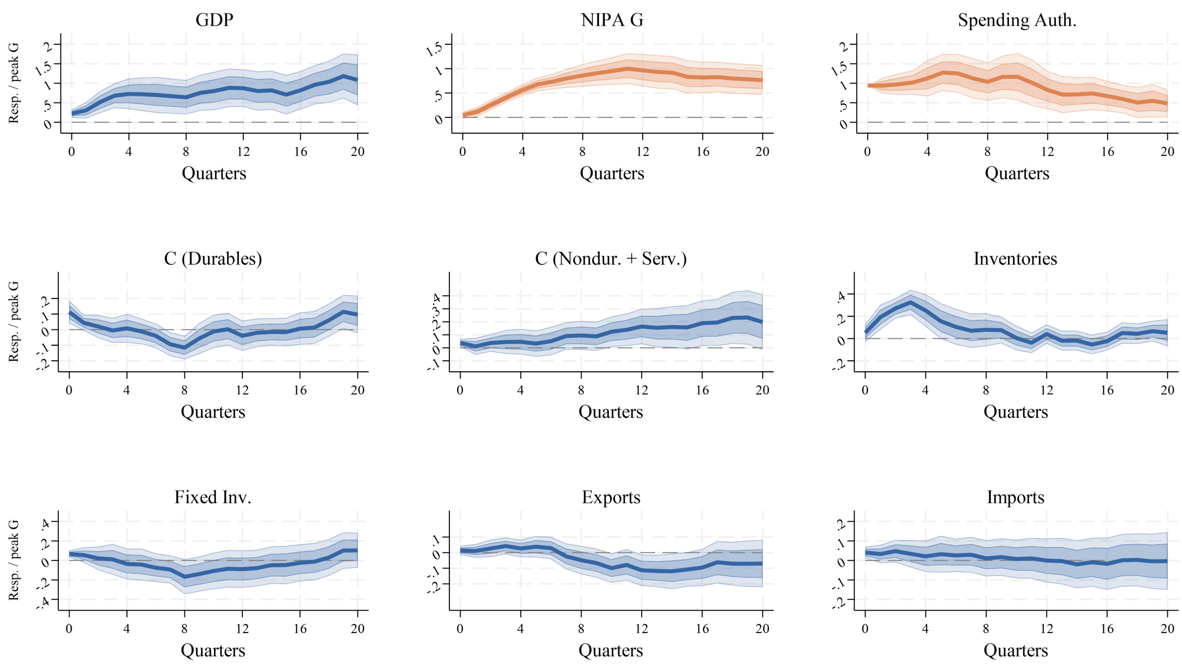

FIGURE 10 — IRFS TO A DEFENSE SPENDING AUTHORIZATION SHOCK

Notes: Baseline specification of Equation (5). Sample is 1947Q1–2017Q4 sample. Bands are 68% and 90% confidence bands, constructed using heteroskedasticity robust standard errors.

initially remains unchanged, before slowly rising more than a year after the initial shock—though statistically the increase in non-durables and services consumption is only marginally significant.

Fixed investment initially responds positively to the spending shock, but then turns negative. Samples excluding the Korean War yield similar results, although the positive response of investment is more pronounced and the negative response is never significant.[25] The bottom middle and right panels display the responses of exports and imports. Exports temporarily decline seven quarters after the shock, while imports increase slightly on impact and then show no further response.

The Response of Inventories. A full accounting of the dynamic responses of all GDP components lies outside the scope of the paper. What matters for our purposes is that a complete decomposition makes clear that the faster response of GDP relative to NIPA G is driven almost entirely by inventories (middle-right panel of Figure 10). Inventories rise immediately on impact and peak roughly three quarters after the shock at about 40% of the eventual peak response of government spending, which occurs many quarters later. Inventories subsequently fall below zero after about three years, although this decline is not statistically different from zero. Overall, the magnitude and timing of the inventory response are

the outbreak of the Korean War is excluded (see Appendix E).

[25]The weaker response of fixed investment when including the outbreak of the Korean War may reflect regulations limiting investment during the Korean War (see Perotti, 2014).

sufficiently large to account for the early GDP response, and inventories are the only GDP component that reacts so strongly at short horizons.

These dynamics are robust to excluding the Korean War from the sample (Appendix E). In this shorter sample, the rise in inventories is still precisely estimated, while the subsequent decline becomes significantly negative, consistent with the rise-and-fall accounting mechanism described earlier in the paper.

STYLIZED FACT 4: *Following shocks to spending authorizations, inventories rise immediately and account for most of the early response of GDP, while NIPA G responds only with delay.*

Inventory Responses Mostly Reflect Defense-Sector Production. As discussed in Section III and in standard macroeconomic textbooks (e.g., Jones), national income accounting uses inventories to reconcile production and expenditure measures of GDP. A rise in inventories therefore can reflect increased production of goods not yet sold. Thus, the increase in inventories in response to spending-authorization shocks can be interpreted as an increase in total production. While, in principle, a sufficiently large shock to spending authorizations could stimulate broad-based production across the economy by stimulating aggregate demand, we believe the observed response is primarily driven by higher production in the defense sector.

First, as shown in Section III, inventories explain the timing mismatch between GDP and NIPA G, according to national accounting standards: newly initiated military production is recorded immediately in GDP as inventory investment, while NIPA G only reflects transfer of ownership to the government made later in the production process. Second, the scale of newly awarded military contracts at the onset of conflicts is consistent with the observed increase in inventories. For example, during the Vietnam War buildup, inventories rose by approximately 0.75% of GDP, while military contracts increased by roughly 1.7% of GDP.[26] Third, Appendix H.1 documents that military production is heavily concentrated in specific manufacturing subsectors (Ramey and Shapiro, 1998; Perotti, 2007; Fisher and Peters, 2010; Nekarda and Ramey, 2011; Cox et al., 2024). A shock to spending authorizations therefore primarily affects a narrow set of defense-related industries rather than generating a broad-based stimulus, which helps explain the rapid rise in inventories. Consistent with this mechanism, we show in Appendix H.2 that the inventory response to military build-ups is concentrated in manufacturing subsectors with the highest exposure to military contracts—such as "other transportation equipment" (NAICS 336) and "computer and electronic products" (NAICS 334). Using a monthly panel of inventories by manufacturing subsector, we find that during military buildups (a) inventories rise significantly, and (b) this rise is entirely driven by sectors receiving military orders.

Inventories and Defense News. In principle, inventories in the defense sector can increase for two reasons. First, firms and defense contractors may anticipate higher future orders and therefore ramp

[26] Much of this surge reflects government purchases of McDonnell–Douglas F-4 Phantom II aircraft. Sales beginning in 1966:3 amounted to 1.2 billion dollars—about 0.15% of GDP—accounting for nearly 9% of the rise in total contract value. This figure covers only the airframe; additional systems (navigation, communication, weapons, engines) were procured separately from other contractors such as General Electric.

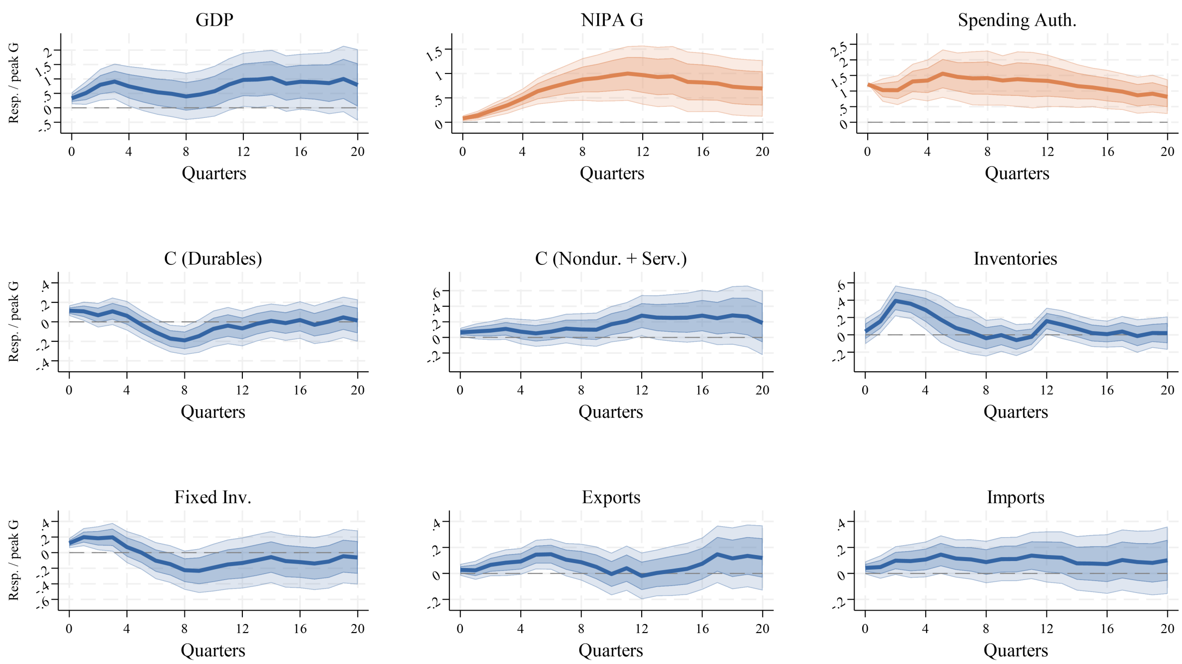

FIGURE 11 — LP RESPONSES OF GDP COMPONENTS TO SPENDING AUTHORIZATION SHOCKS ORTHOGONAL TO DEFENSE NEWS SHOCKS

Notes: Sample: 1947Q1 to 2015Q4, the end date of the defense news shock series constructed by Ramey and Zubairy (2018). All else as in Figure 10.

up production today in response to news. Second, defense contractors and their suppliers may increase production in response to unanticipated newly awarded contracts. These two mechanisms are not substitutes but are complementary. For example, in Section II, we show that a defense news shock triggers a new wave of military contracts and newly authorized defense funds on impact (Figure 2), consistent with the fact that defense news and spending authorizations Granger-cause each other (bottom panel of Table 3).

However, if the response of inventories to shocks to spending authorizations were mostly driven by anticipation effects unrelated to contract awards, we would expect the response of inventories to shocks to spending authorizations orthogonal to news to be significantly smaller. We test this hypothesis empirically by identifying an unanticipated government spending shock using innovations to spending authorizations that are orthogonal to defense news shocks. In principle, this is asymptotically equivalent to estimating a VAR that orders spending authorizations after defense news shocks. Figure 11 shows the results.

From the middle right panel, inventories respond faster than NIPA G, peaking two to three quarters after the shock at a value equivalent to about 40% of the peak response of NIPA G. Fixed investment and durables also display significant early positive responses, but none is as large or as precisely estimated as the response of inventories. The other components of GDP have insignificant responses. Overall, the faster response of GDP to an unanticipated government spending shock identified with spending authorizations, relative to NIPA G, is still driven by inventories, even when the shock is orthogonal to defense news shocks.

V. Multiplier: Underestimation and Accounting-Correction

We estimate standard cumulative fiscal multipliers (Mountford and Uhlig, 2009) using local-projection instrumental variables (LP-IV), following Ramey and Zubairy (2018).[27] Specifically, we estimate the following equation:

$$\sum_{h=0}^{H} \frac{\text{GDP}_{t+h} - \text{GDP}_{t-1}}{\text{GDP}_{t-1}} = \mathcal{M}_H \cdot \underbrace{\sum_{h=0}^{H} \frac{\text{G}_{t+h} - \text{G}_{t-1}}{\text{GDP}_{t-1}}}_{\text{Instrument with } Z_t} + \text{lags} + \nu_t, \tag{6}$$

where the lags include four lags of the variables in Equation (5). The cumulative change in G is instrumented with Z_t. The 2SLS estimate of $\mathcal{M}_H$ provides a direct estimate of the cumulative fiscal multiplier at horizon H.

Multipliers via Spending Authorizations. By setting the instrument equal to:

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

we recover the fiscal multiplier associated with spending authorizations. In particular, the identified structural shocks in this case are innovations to the HBR-transformed spending-authorization series (η_t^{SA} in the previous section) and the point estimates of $\mathcal{M}_H$ are numerically equivalent to the ratio of the cumulative impulse response of GDP (top-left panel of Figure 10) to the cumulative impulse response of G (top-middle panel of Figure 10).

The left panel of Figure 12 displays the estimated cumulative fiscal multipliers obtained by instrumenting NIPA G with spending authorizations, together with the 68% and 90% confidence bands.

Ignoring for the moment the very large impact multipliers—discussed further in the next subsection—we estimate one-year multipliers well above unity, around 1.7. The multiplier then declines toward 1 and remains near that level at longer horizons. Exact point estimates and the associated Montiel Olea and Pflueger (2013) effective F-statistics are reported in Appendix G.1 (Table G1).

Multipliers via BP shocks. The right panel shows the cumulative multipliers estimated using NIPA G in the same baseline specification, but omitting spending authorizations:

$$Z_t := \frac{G_t - G_{t-1}}{\text{GDP}_{t-1}},$$

[27] For more technical references, see Ramey (2016), Stock and Watson (2018), and Plagborg-Møller and Wolf (2021).

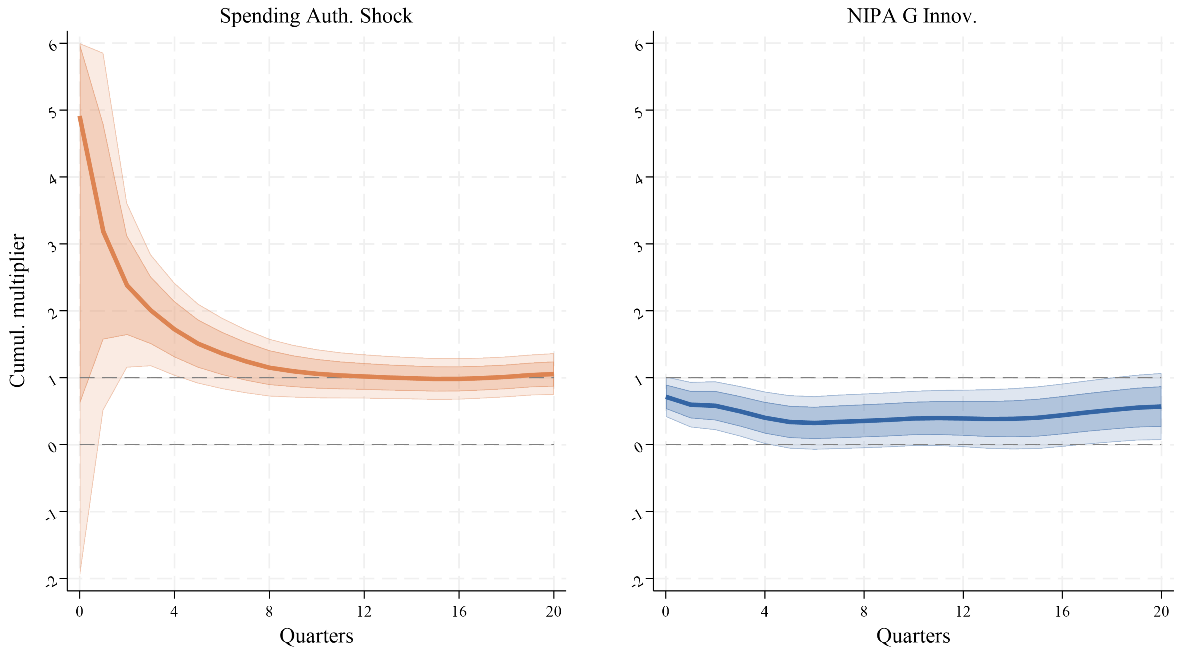

FIGURE 12 — FISCAL MULTIPLIERS
SPENDING AUTHORIZATIONS SHOCK VERSUS BP SHOCK (NIPA G INNOVATIONS)

Notes: Left panel: LP-IV estimates that instrument cumulative NIPA G with spending authorizations. Right panel: LP-IV estimates that instrument cumulative NIPA G with the innovation to NIPA G from the same baseline specification after omitting spending authorizations. Baseline HBR specification, no defense-news controls, 1947Q1–2017Q4 sample. Confidence bands are 68% and 90%, calculated with heteroskedasticity robust standard errors following (Montiel Olea and Plagborg-Møller, 2021). Bands are capped at five for illustrative reasons.

an approach that mimics a recursive identification a la Blanchard and Perotti (2002) in which NIPA G is ordered first to recover the structural shocks (ξ_t^{BP} in the previous section). Following Ramey (2016)-notation, we refer to these shocks as BP shocks. The so calculated fiscal multipliers range from roughly 0.5 to 0.8, consistent with earlier findings in the aggregate multiplier literature (e.g., Hall, 2009).

Baseline Results: Underestimation of Fiscal Multipliers. Multipliers estimated using BP shocks (ξ_t^{BP}), reported in the right panel of Figure 12 are noticeably lower than those obtained using spending authorizations, left panel of the same figure. We argue that measurement delays are the source of the lower estimates.

First, Figure 13 shows the contributions of three components of spending—defense procurement spending, defense wages, and the residual component—to the IRF of NIPA G in response to BP shocks.

On average, about 75% of the IRF of NIPA G to a BP shock is accounted for by the response of defense spending less CFC, and about 60% is accounted for by defense procurement. This strong exposure of BP shocks to defense spending is consistent with the findings of Cox et al. (2024). Consequently, BP

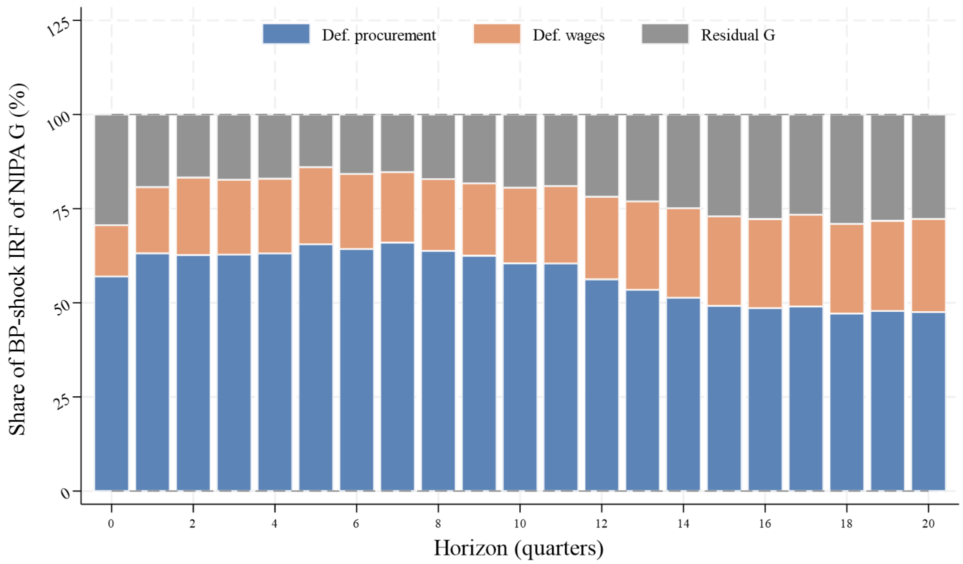

FIGURE 13 — BP SHOCK IRF OF NIPA G: COMPONENT SHARES

Notes: Baseline specification without spending authorizations. Sample is 1947Q1–2017Q4 sample. The BP shock is the innovation to NIPA G from the recursive local-projection specification that omits spending authorizations. Each bar reports, horizon by horizon, the share of the NIPA G IRF accounted for by defense procurement, defense wages, and residual G. Shares sum to 100 by construction.

shocks load heavily on variation in NIPA G that originates in the component of government spending most affected by the measurement-delay problem.

Second, NIPA defense spending and defense procurement are predicted by authorizations and military contracts (Table 2). Moreover, BP shocks are predicted by spending-authorization shocks (Table 5). Therefore, when measurement delays are sufficiently long, BP shocks pick up delayed variation in defense spending, which leads to an understated response of inventories. To corroborate this point, Figure 14 plots the IRF of inventories in response to a BP shock alongside the response of inventories to a spending-authorization shock, with each normalized by the corresponding peak response of NIPA G, to allow for a fair comparison.

The inventory response captured by BP shocks is smaller than that captured by spending-authorization shocks. Therefore, we argue that the lower estimates arise from the timing mismatch between NIPA G and spending authorizations: shocks identified from innovations to NIPA G—BP shocks—understate the output response through inventories, leading to lower estimates of the standard cumulative multiplier.

Robustness: Excluding the Korean War. The literature has long noted that fiscal multipliers become imprecisely estimated when the Korean War is excluded, primarily because military spending exhibits much less variation once both World War II and the Korean War are removed from the sample (see Dupor and Guerrero, 2017). Although our estimates are also less precise in this restricted sample, we obtain reasonably precise multiplier estimates at the one-year horizon.

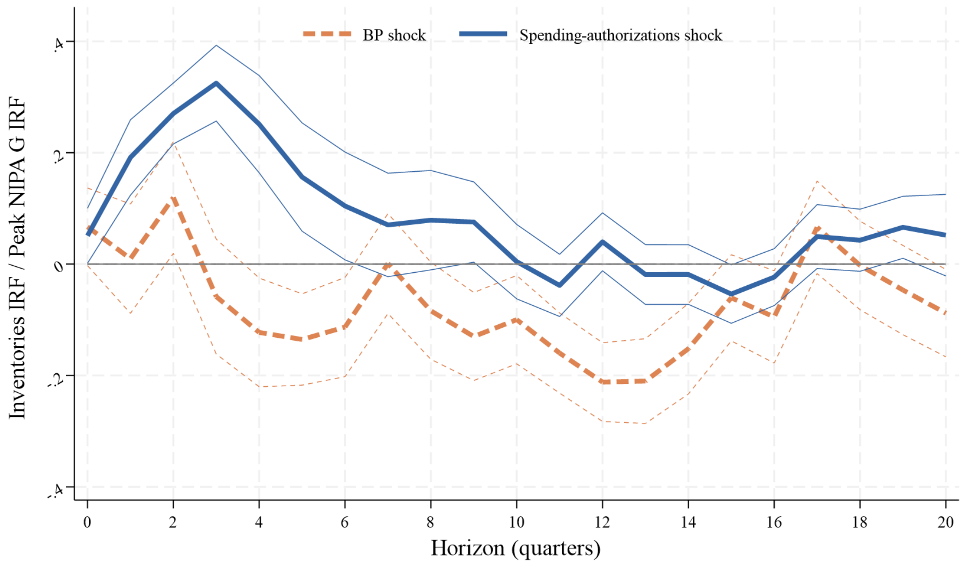

FIGURE 14 — INVENTORIES RESPOND MUCH MORE TO SPENDING-AUTHORIZATION SHOCKS THAN TO BP SHOCKS

Notes: Baseline HBR specification, no defense-news controls, 1947Q1–2017Q4 sample. The orange dashed line plots the IRF of inventories to a BP shock, normalized by the peak response of NIPA G to that BP shock. The blue solid line plots the IRF of inventories to a spending-authorization shock, normalized by the peak response of NIPA G to that spending-authorization shock. Dashed lines of the same color denote 68% confidence bands.

In this sample, multipliers rise rapidly and peak three quarters after the shock at 1.73, while the one-year multiplier is 1.66. In the short run, they are consistently larger than multipliers estimated using NIPA. Thereafter, they decline and reach 1.09 after two years. This pattern is not surprising: it aligns with the rise and subsequent decline of inventories, whose impulse response is estimated precisely even in the shorter sample (see Appendix E). Consistent with this mechanism, multipliers identified using NIPA G fall below one after one year, whereas those identified using spending authorizations remain significantly above one at the same horizon, reflecting the peak inventory response.

Robustness: News Orthogonalization. Appendix G.1 report analogous results after orthogonalizing fiscal shocks with respect to defense news (with-news columns of Table G1). Specifically, we use the defense news series of Ramey and Zubairy (2018) to purge both shock measures of their anticipatory component. The qualitative result is unchanged: even after controlling for news, multipliers identified using shocks to NIPA G remain smaller than those identified using shocks to spending authorizations.

This finding has a practical implication for identification. Orthogonalizing with respect to defense news is valuable for addressing fiscal foresight, but it is not a substitute for observing the fiscal shock earlier in the spending process (NIPA G). Our preferred approach to identify unanticipated shocks is therefore using spending authorizations and, when desired, to orthogonalize those shocks with respect to defense news. When shocks are instead identified from NIPA G, the measurement-delay problem remains: part of the initial output response still appears through inventories before it is recorded as government spending. For that reason, news-orthogonalized NIPA shocks continue to yield smaller multipliers than authorization-based shocks.

V.1. A New Accounting-Corrected Measure of Fiscal Multiplier.

The very large multiplier estimates at short horizons reflect a denominator problem created by measurement delays: shocks to spending authorizations generate an immediate response in output, but they take several quarters to appear in NIPA G. As a result, when the multiplier is constructed using the cumulative response of GDP divided by the cumulative response of NIPA G, the denominator remains small at short horizons. The estimated multiplier is therefore both very large and highly imprecise (left panel of Figure 12). A related issue appears in Ramey (2016), where the impact multiplier of a defense news shock is reported as -7.5: output rises on impact, while measured government spending initially declines slightly (see Figure 1) so the ratio becomes very large in absolute value.

Whether this is a problem depends on the object policy-makers want to measure. If the fiscal multiplier is defined as the effect on output per dollar of *government spending* as recorded in NIPA, then these short-run estimates are conceptually appropriate, even if they are estimated imprecisely. If instead the multiplier is intended to measure the effect on output per dollar *authorized*, then the same estimates are too large at short horizons.

We therefore introduce an accounting-corrected fiscal multiplier that addresses the measurement delay problem. In particular, the accounting-correction is designed to remove the mechanical inflation generated by delayed measurement in NIPA G, while preserving the standard interpretation of the multiplier at longer horizons, when authorizations and measured NIPA government spending converge.

Illustrative Framework. To introduce the accounting-corrected multiplier we first return to the stylized model to illustrate the basic intuition behind it. In particular, consider the following simple data-generating process for output:

$$Y_t = \phi \cdot Y_{t-1} + \gamma \cdot A_t + \epsilon_t, \tag{7}$$

where output exhibits persistence ϕ and responds contemporaneously and positively to authorizations, A_t with $\gamma > 0$. In this setup, given Equation (2) from the previous section—replace MPC_t with authorizations A_t—and Equation (7), the impulse response of output at time $t + h$ to a fiscal shock is given by:

$$\frac{\partial Y_{t+h}}{\partial \eta_t} = \gamma \cdot \frac{\rho^{h+1} - \phi^{h+1}}{\rho - \phi}.$$

Similarly, using Equation (4), the response of NIPA G at time $t + h$ to a fiscal shock at time t can be expressed as:

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

The fiscal multiplier is defined as the ratio of the cumulative output response to the cumulative response of government spending, and can be interpreted as *the effect of one dollar of government spending on GDP*. Formally, the multiplier at horizon H is:

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

where $f(x) := \frac{x(1-x^{H+1})}{1-x}$.

Undeterminedness of Impact Multipliers. Expression (8) highlights a key issue in estimating fiscal multipliers when government spending is measured with delay ($\lambda > 0$). When delays in NIPA G relative to authorizations are substantial (i.e., λ close to one), the impact multiplier $\mathcal{M}_0$ is mechanically inflated and diverges to infinity:

$$\lim_{\lambda \to 1} \mathcal{M}_0 = \infty.$$

This result has a clear policy implication: because NIPA G is measured with delay, it is not possible to meaningfully quantify the instantaneous (period 0) fiscal multiplier, since GDP reacts before NIPA G.

The undeterminedness of impact multipliers arises because newly authorized defense spending appears in NIPA only gradually. At short horizons, a standard multiplier that uses contemporaneous NIPA spending in the denominator therefore compares current GDP movements with a spending measure that has not yet caught up to the underlying authorization shock.

Convergence of Long-run Multipliers. Another key message from Expression (8) is that the population long-run multiplier is not affected by measurement delays:

$$\lim_{H \to \infty} \mathcal{M}_H = \frac{\gamma}{1-\phi},$$

which is independent of λ. This result formalizes the fact that since every \$1 of authorized fund is ultimately spent, in the limit we have that authorizations and spending are identical, and the fiscal multiplier is not affected by measurement delays.

We can leverage this accounting point to construct an *accounting corrected multiplier*, whose magnitude does not depend on measurement delays.

Breakdown the Fiscal Multiplier. First, we need to isolate the component of NIPA G which is the source of the measurement problem. Therefore, we break down NIPA G into two parts:

$$G_t = G_t^a + G_t^r,$$

where G_t^a is the component of NIPA G that ultimately accrues to spending authorizations, i.e. NIPA defense spending less CFC, and G_t^r is the residual component of NIPA G, which is not the object measured with delay relative to spending authorizations. On the contrary, \$1 of spending authorizations will eventually flow into G_t^a.

Let the cumulative IRFs to an authorization shock, η_t, be expressed using the following notation:

$$\Gamma_H := \sum_{h=0}^{H} \frac{\partial \text{SA}_{t+h}}{\partial \eta_t}, \qquad \Theta_H := \sum_{h=0}^{H} \frac{\partial \text{PRSP}_{t+h}}{\partial \eta_t}, \qquad \Phi_H := \sum_{h=0}^{H} \frac{\partial G^r_{t+h}}{\partial \eta_t}, \qquad \Psi_H := \sum_{h=0}^{H} \frac{\partial G^a_{t+h}}{\partial \eta_t},$$

where SA_t is spending authorizations and $\text{PRSP}_t := Y_t - G_t$ is private spending in the sense of Ramey (2013).[28] Because GDP can be written as private spending plus the authorization-linked government component plus the residual government component, we have:

$$\text{GDP}_t = \underbrace{\text{GDP}_t - G_t}_{\text{PRSP}_t} + G_t = \text{PRSP}_t + G_t^a + G_t^r,$$

and the standard cumulative fiscal multiplier can thus be written as:

$$\mathcal{M}_H := \frac{\Theta_H + \Psi_H + \Phi_H}{\Psi_H + \Phi_H} = 1 + \frac{\Theta_H}{\Psi_H + \Phi_H}.$$

The standard multiplier $\mathcal{M}_H$ is above one as long as the fiscal shock η_t has a positive effect on private spending ($\Theta_H > 0$) and is below one if the fiscal shock crowds out private spending ($\Theta_H < 0$).

In practice, at short horizons ($H \to 0$), GDP responds through inventories, reflecting the work in progress of military contractors responding to newly authorized funds ($\Theta_H > 0$), but money is not spent immediately; hence, Ψ_H is very small, if not zero, while the effect of spending authorizations on other government spending components, Φ_H, is also very likely small. As a consequence, this shrinks the denominator of the multiplier, which then diverges to meaningless large values at short horizons, as the illustrative model makes clear, and consistent with empirical evidence (left panel of Figure 12).

Second, divide and multiply the multiplier $\mathcal{M}_H$ by the cumulative IRF of spending authorizations, Γ_H:

$$\mathcal{M}_H = 1 + \frac{\Theta_H/\Gamma_H}{\Psi_H/\Gamma_H + \Phi_H/\Gamma_H} = 1 + \frac{\mathcal{P}(H)}{\mathcal{G}^a(H) + \mathcal{G}^r(H)},$$

where we define the ratios of cumulative IRFs as follows:

$$\mathcal{P}(H) := \frac{\Theta_H}{\Gamma_H}, \qquad \mathcal{G}^a(H) := \frac{\Psi_H}{\Gamma_H}, \qquad \mathcal{G}^r(H) := \frac{\Phi_H}{\Gamma_H},$$

here $\mathcal{P}(H)$ is the private-spending multiplier per dollar authorized, $\mathcal{G}^r(H)$ is the residual NIPA government-spending multiplier per dollar authorized, and $\mathcal{G}^a(H)$ is the multiplier on the authorization-linked component of NIPA spending. The three components allows us to break down the standard cumulative multiplier $\mathcal{M}_H$. Figure 15 displays the empirical estimates of $\mathcal{P}(H)$, $\mathcal{G}^r(H)$, and $\mathcal{G}^a(H)$ obtained via

[28] Private spending is distinct from private GDP. Private GDP equals private spending plus government purchases of privately produced goods and services, such as domestically produced defense procurement.

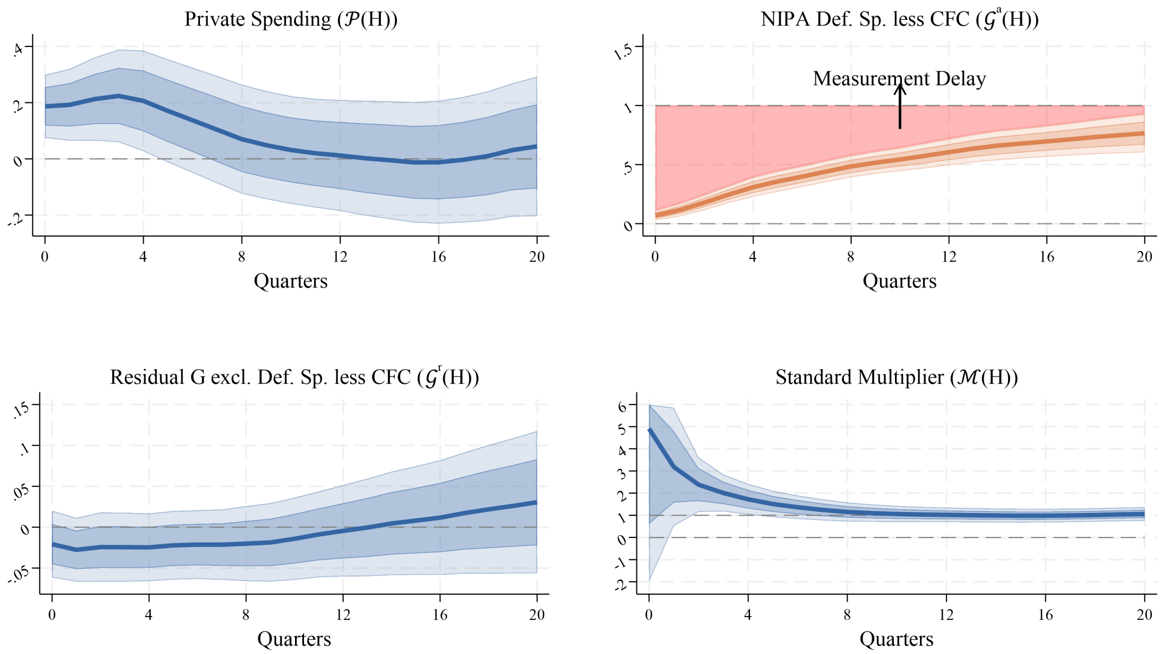

FIGURE 15 — MULTIPLIER BREAKDOWN FOR A SPENDING-AUTHORIZATION SHOCK

Notes: Baseline HBR specification, spending-authorization shock, no defense-news controls, 1947Q1–2017Q4 sample. Top-left panel: private spending multiplier to a spending-authorization shock. Top-right panel: multiplier on NIPA defense spending less CFC. Bottom-left panel: residual NIPA G multiplier, excluding defense spending less CFC. Bottom-right panel: standard GDP multiplier. Robust standard errors with 68% and 90% confidence bands.

LP-IV using the baseline specification.

First, the top-left panel plots the cumulative private spending response to a dollar of authorized funds, $\mathcal{P}(H)$, which rises on impact and remains elevated for roughly the first year, peaking after about three to four quarters before gradually returning toward zero. The positive estimates of $\mathcal{P}(H)$ indicate that new spending authorizations crowd-in private resources across the economy, a result which may reflect an endogenous rise in labor productivity, induced by new military production, typically characterized by learning-by-doing (see McGrattan and Ohanian, 2010; D'Alessandro, Fella, and Melosi, 2019; Ilzetzki, 2023; Briganti, 2023).[29]

Second, the bottom-left panel plots $\mathcal{G}^r(H)$, the residual NIPA G multiplier, defined as total NIPA G net of defense spending less CFC. $\mathcal{G}^r(H)$ is slightly negative at short horizons and turns positive only gradually at medium horizons, but never significant, suggesting limited evidence that defense spending crowds out the residual non-defense spending component. This finding supports the idea that "guns

[29] Other evidence which supports productivity enhancements induced by military spending are provided in Gross and Sampat (2023) (WWII), Antolin-Diaz and Surico (2025) (long-run effects). On the contrary, highly specialized public R&D programs may have weaker effects on short-term productivity: Kantor and Whalley (2025) (moonshot program) and Fieldhouse and Mertens (2026) (public R&D shocks).

did not come at the expense of butter", a result broadly consistent with the evidence brought forward in Ilzetzki (2025).

Third, the top-right panel plots $\mathcal{G}^a(H)$, the multiplier on NIPA defense spending less CFC, that is, the portion of NIPA spending that ultimately absorbs the authorization shock. The path of $\mathcal{G}^a(H)$ is crucial for understanding the large observed values of the multiplier $\mathcal{M}_H$ at short horizons: it is close to zero on impact and slowly converges to 1, confirming empirically the validity of our accounting principle that every \$1 of spending authorizations ultimately flows into NIPA defense spending less CFC. Nonetheless, even after five years, its point estimate remains below one. Extending the horizon H, we find that it takes about eight years to spend all authorized funds. The shaded gap relative to the value of one highlights the time-to-spend wedge created by measurement delays: it takes time for the national accounts to record the work in progress of military contractors as NIPA government spending.

Finally, the bottom-right panel plots the standard GDP multiplier $\mathcal{M}_H$, which can be constructed by combining the three other components. Since $\mathcal{G}^a(H)$ reacts with delay and the residual component of NIPA G captured by $\mathcal{G}^r(H)$ also does not react significantly to a spending authorization shock, this implies that NIPA G does not react on impact, leading to high values of the fiscal multiplier at short horizon, as confirmed by the dynamics of $\mathcal{M}_H$ in the figure.

The Accounting-Corrected Multiplier. We can now construct an accounting-corrected multiplier by exploiting the simple accounting property of spending authorizations: \$1 of authorization translates, over time, into \$1 of NIPA spending; formally:

$$\lim_{H\to\infty} \mathcal{G}^a(H) = \lim_{H\to\infty} \frac{\Psi_H}{\Gamma_H} = 1$$

that is, the cumulative amount of authorization-linked NIPA spending, Ψ_H, will eventually coincide with the cumulative amount of authorized dollars, Γ_H.

Therefore, we define the *accounting-corrected multiplier* as the standard cumulative multiplier under the assumption that every authorized dollar is ultimately spent:

$$\mathcal{M}^c(H) := 1 + \frac{\mathcal{P}(H)}{1 + \mathcal{G}^r(H)}. \tag{9}$$

Notice that, even if the residual part of NIPA G does not respond strongly at short horizons ($\lim_{H\to 0} \Phi_H = 0$), the denominator of the accounting-corrected multiplier does not approach zero, since we are now counting authorized dollars rather than dollars spent. Hence the accounting-corrected multiplier is well-defined even at short horizons and on impact.

Another way to interpret the accounting correction is to note that this is equivalent to replacing the delayed cumulative response of NIPA defense spending (net of CFC), Ψ_H, with the cumulative response of authorizations themselves, Γ_H, while leaving the residual component of NIPA G, Φ_H, in the denominator:

$$\mathcal{M}_H := \frac{\Theta_H + \Psi_H + \Phi_H}{\Psi_H + \Phi_H} \Longleftrightarrow \mathcal{M}^c(H) := \frac{\Theta_H + \Gamma_H + \Phi_H}{\Gamma_H + \Phi_H}$$

In the limit, as authorized money is spent and the authorization-linked component catches up, $\Psi_H \to \Gamma_H$, the accounting-corrected and standard multipliers coincide:

$$\lim_{H\to\infty} \mathcal{M}^c(H) = \lim_{H\to\infty} \mathcal{M}_H.$$

Lastly, Figure 16 plots the resulting accounting-corrected multiplier $\mathcal{M}^c(H)$. The point estimate is about 1.2 on impact, peaks at roughly 1.25 after three to four quarters, and then gradually drifts back toward one. Relative to the standard GDP multiplier in the bottom-right panel of Figure 15, the accounting-corrected series removes the short-run distortion created by the slow pass-through from authorizations into NIPA defense spending less CFC. The main message is therefore straightforward: a dollar of spending authorization raises GDP by about a dollar, and slightly more in the first year, once the accounting delay is corrected for.

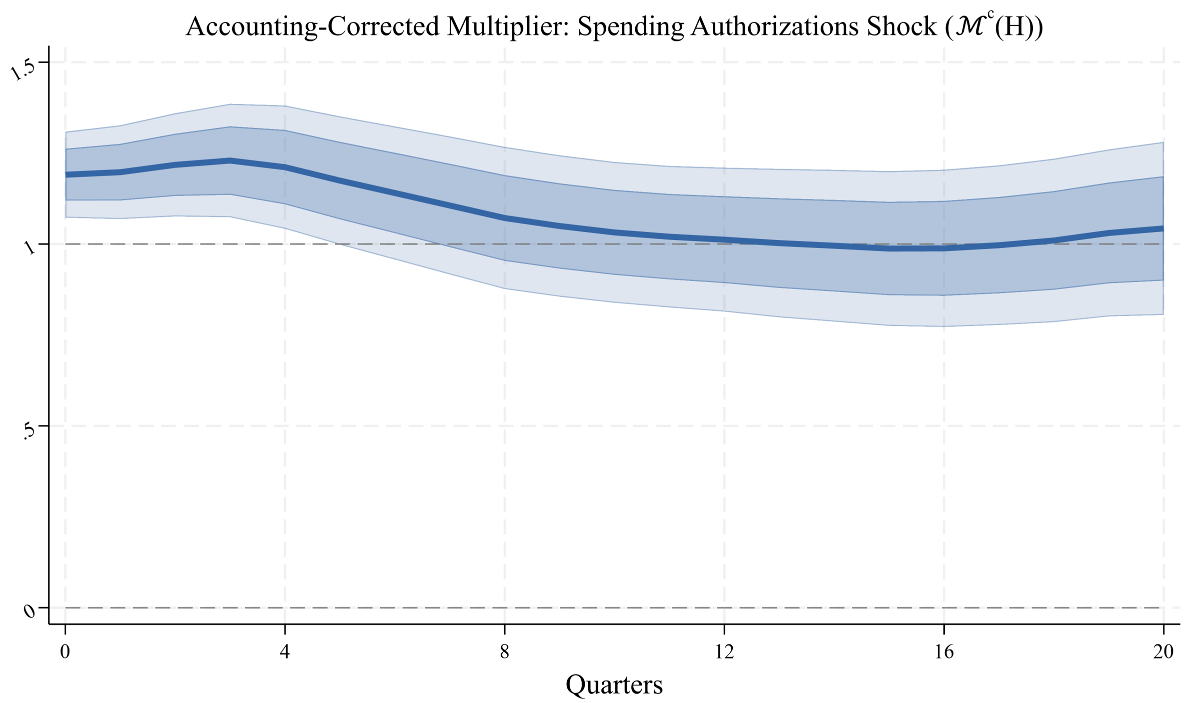

FIGURE 16 — ACCOUNTING-CORRECTED MULTIPLIER FOR A SPENDING-AUTHORIZATION SHOCK

Notes: Baseline HBR specification, spending-authorization shock, no defense-news controls, 1947Q1–2017Q4 sample. The figure plots $\mathcal{M}^c(H) = 1 + \mathcal{P}(H)/(1 + \mathcal{G}^r(H))$. Robust standard errors with 68% and 90% confidence bands.

Robustness. Supporting figures are reported in Appendix G.3. Two robustness exercises are especially informative. First, starting the sample in 1956 leaves the basic decomposition intact but yields wider confidence bands. The point estimate rises more sharply in the first year and then falls back toward one. Using military prime contracts rather than spending authorizations over the full sample delivers a similar message, although the profile is flatter and somewhat lower at medium horizons. In both cases, the central lesson is unchanged: once the delayed authorization-linked component of NIPA spending is corrected for, the short-run accounting-corrected multiplier remains close or above one and remains above short-run multipliers obtained from NIPA G.

Policy Implications. We believe that the relevant policy question is not whether output rises per dollar of NIPA spending recorded today, but whether output rises per dollar of spending that has been authorized and will ultimately be disbursed. Our accounting-corrected measure answers that question while respecting the national-income accounting decomposition: it replaces the delayed authorization-linked component of NIPA spending with authorizations themselves. This produces multipliers near or slightly above one rather than short-run estimates dominated by the timing conventions of NIPA recording. More broadly, our results suggest little evidence of substantial crowding out.

VI. Conclusion

This paper shows that the timing of government spending measurement has first-order implications for both shock identification and multiplier estimation. In defense procurement, GDP can rise before measured NIPA government spending not only because agents receive news about future fiscal expansions, as emphasized by Ramey (2011), but also because national-income accounting records ongoing military production as private inventory investment until delivery. The timing mismatch between output and measured government spending therefore reflects two distinct mechanisms: fiscal foresight and delayed expenditure recording in the national accounts.

We document the second mechanism using newly constructed quarterly series of military contracts and spending authorizations. Both series lead their NIPA counterparts by roughly three to four quarters and continue to predict measured defense spending even after annual aggregation and after controlling for defense news shocks. This implies that innovations in recorded NIPA G are not, by themselves, a clean measure of unanticipated government spending shocks when spending is recorded late in the production process. Defense news remain essential for isolating anticipated variation, but orthogonalizing with respect to news does not eliminate the separate measurement problem embedded in NIPA spending.

Following authorization shocks, GDP rises immediately while measured NIPA G responds only gradually, and the early output response is accounted for largely by inventories. This is consistent with work in progress being recorded as private inventories before final delivery to the government. These timing differences matter quantitatively for fiscal multipliers. Multipliers identified from innovations in measured NIPA G are biased downward at short horizons because they miss part of the initial output response. When shocks are identified using spending authorizations instead, the estimated multiplier exceeds one within the first year and converges toward one at longer horizons. Our accounting-corrected multiplier formalizes this point by replacing the delayed component of recorded spending with authorizations. It is well defined even on impact and implies a simple quantitative message: a dollar of defense spending authorization raises GDP by about a dollar, and slightly more in the first year. In our setting, these estimates provide little evidence of short-run crowding out of private spending following a defense expansion.

The practical implication for applied work is straightforward. Researchers interested in identifying unanticipated government spending shocks should not rely mechanically on innovations in recorded NIPA spending when the underlying expenditure is subject to long and variable accounting delays, as is typical in defense procurement. In such settings, earlier administrative measures should be used to identify the shock at the time of commitment. When fiscal foresight is also a concern, the appropriate strategy is to combine the two insights: use defense news to account for anticipated variation, and use authorizations to recover the surprise component that NIPA records only with delay.

More broadly, our results show that measurement conventions in the national accounts are not a secondary data issue. They can materially affect identification, estimated multipliers, and the interpretation of fiscal transmission. For policymakers interested in the real-time fiscal stance, authorizations may provide a more informative measure of the timing and magnitude of fiscal actions than ex post spending flows. Because the treatment of work in progress follows the broader System of National Accounts, the same logic is likely to apply beyond U.S. defense procurement. More generally, our findings suggest that, in fiscal analysis, measurement is part of the identification problem rather than a secondary bookkeeping detail.

Appendix

A Military Contracts Data

A.1. 1947Q1 to 1950Q4 - 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 predict real military prime contracts per capita from 1947Q1 to 1950Q4 using information from (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:

$$\text{MPC}_t = \kappa + \beta \cdot (\text{Avg.Hours.Aircraft})_t + \sum_{h=0}^{4} \psi_h \cdot \text{NIPA}_{t+h} + \varepsilon_t \tag{10}$$

We estimate Equation (10) via OLS, using a sample spanning from 1951Q1 to 1980Q4. The regression delivers an in-sample R^2 of 64.8%.

Why Using 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 Ramey and Shapiro, 1998; Fisher and Peters, 2010; McGrattan and Ohanian, 2010; Nekarda and Ramey, 2011; Ilzetzki, 2023). Average weekly hours of production workers is an excellent proxy for capital utilization and tracks ongoing production (see Bils and Cho, 1994; Fernald, 2012) . The bottom panel of Figure A1 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).

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.

Why Leads of NIPA Defense Procurement Spending? 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.

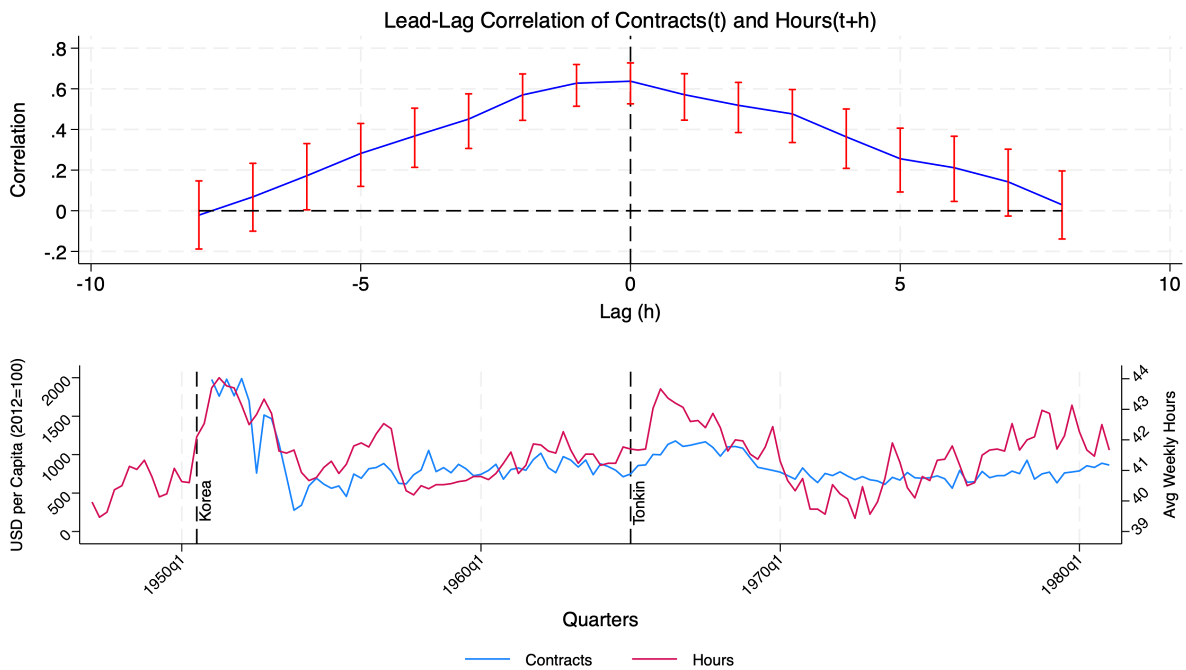

FIGURE A1 — COMOVEMENT IN WEEKLY HOURS OF AIRCRAFT PRODUCTION WORKERS AND MILITARY CONTRACTS

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 original series.

Evidence from the Survey of Current Business. 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 Chart 2 on page 3 of SCB August 1950 as well as page 4 (section "Expansion centered in durables"):

"*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 [...].* "

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:

"*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.*"

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.2. 1983Q1 to 2003Q4 — Federal Procurement Summary Report

The Official *Federal Procurement Summary Report* (FPSR), produced annually by the Directorate for Information Operations and Reports (DIOR), provides (i) the annual value of military prime contract awards and (ii) bar charts reporting quarterly values of *total* federal procurement contracts (military plus non-military). Since quarterly military contract data are unavailable for this period, we construct a quarterly series using the information embedded in these quarterly federal totals.

We proceed in two steps.

1. We seasonally adjust the quarterly (annualized) real series of total federal contracts using the X-13 ARIMA-SEATS program from the Census Bureau.

2. We interpolate the annual military contract series by scaling the quarterly seasonally adjusted federal series so that, within each fiscal year, its average matches the corresponding annual value of military prime contract awards reported in the FPSR. This adjustment aligns the quarterly series with the FPSR fiscal-year totals and removes any gap between the fiscal-year average of quarterly contracts and the reported military amounts.

Figure A2 summarizes this procedure. The top-left panel compares the original quarterly non–seasonally-adjusted federal series with its X-13 seasonally adjusted counterpart (step 1). The top-right panel shows the quarterly series before (red) and after (blue) the fiscal-year adjustment: because military procurement represents roughly 80% of federal procurement over this period, the fiscal-year adjustment produces only a small and uniform downward shift (step 2). The bottom-left panel displays the resulting quarterly series of military contracts (blue), whose fiscal-year averages match exactly the annual military values reported in the FPSR.

Federal Contracts Driven by Military Contracts. A natural concern is that quarterly variation in federal procurement might reflect fluctuations in non-military awards. However, military procurement dominates federal procurement during these years. The left panel of Figure A3 compares annual

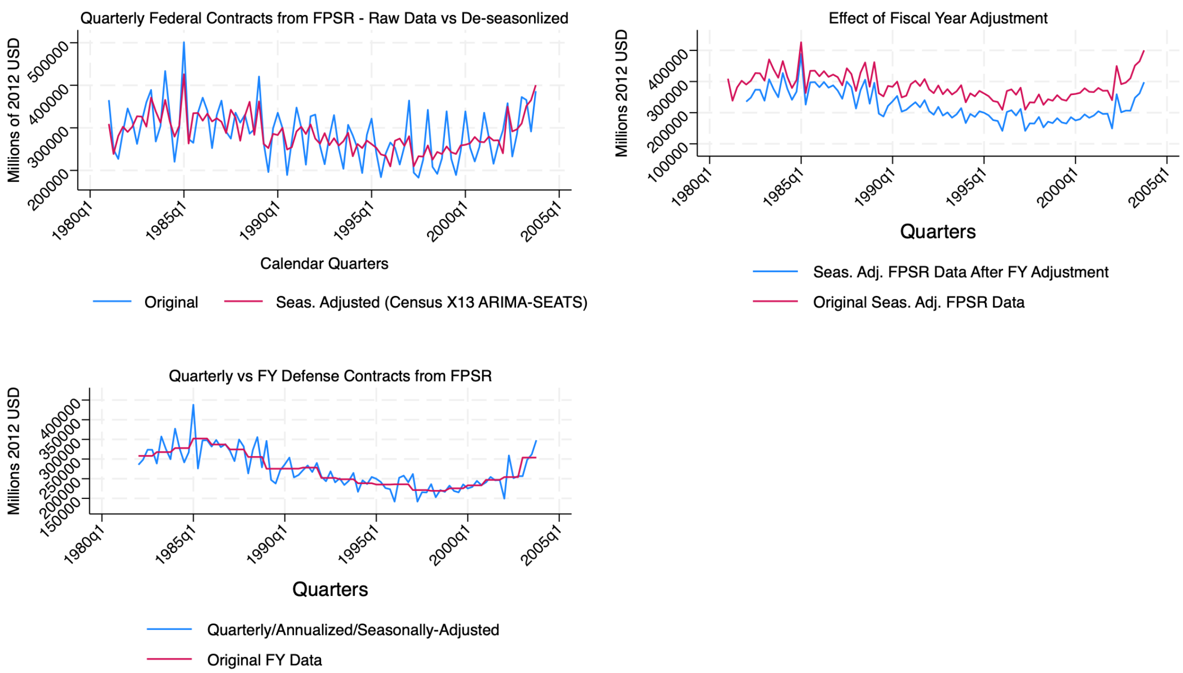

FIGURE A2 — FEDERAL PROCUREMENT SUMMARY REPORT:
SEASONAL AND FISCAL YEAR ADJUSTMENTS

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

military contracts with quarterly federal procurement aggregated to the fiscal-year level; the two series move closely together. The right panel documents that military procurement accounts for approximately 80% of total federal procurement throughout this period. Hence, quarterly variation in federal procurement primarily reflects the timing of military contracts, supporting our use of quarterly federal totals to interpolate quarterly military procurement.

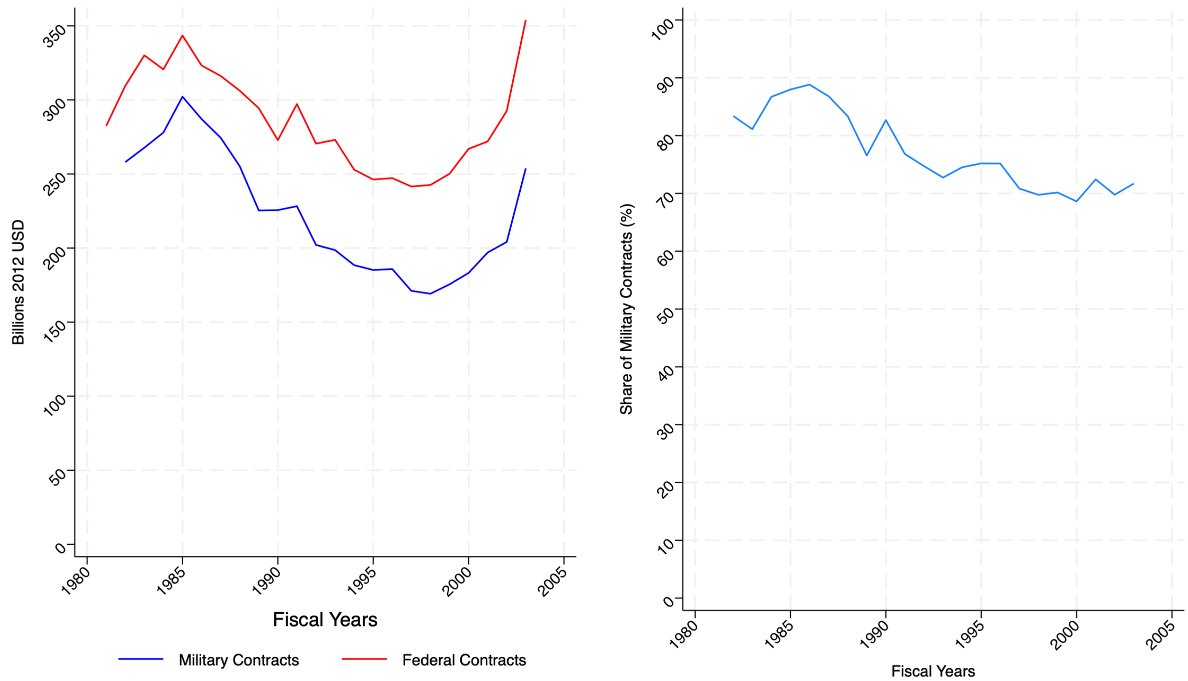

FIGURE A3 — FEDERAL PROCUREMENT SUMMARY REPORT ANNUAL DATA
FEDERAL VS. MILITARY CONTRACTS

Notes: Left panel: annual values of federal and military contracts. Quarterly federal procurement (military plus non-military) is aggregated to the fiscal-year frequency. Right panel: military procurement as a share of total federal procurement by fiscal year.

A.3. 2004Q1 Onward — Federal Procurement Data System Next Generation

For 2004Q1 onward, we identify military contracts using the universe of federal procurement transactions reported in the Federal Procurement Data System–Next Generation (FPDS-NG), restricting attention to contracts awarded by the Department of Defense.[30]

We aggregate DoD contracts to the fiscal-year level and interpolate a quarterly series using information from quarterly variation in newly awarded military contracts. This procedure is preferred to aggregating all obligations at the quarterly frequency because large de-obligations often occur in the quarters following the initial obligation. These negative adjustments cancel previously recorded obligations and introduce substantial noise into the quarterly data. Aggregation at the fiscal-year level balances out de-obligations, which have no real economic effect, while preserving annual totals.

The interpolation method parallels that used for the FPSR data:

1. We seasonally adjust the quarterly series of newly awarded military contracts using the X-13 ARIMA-SEATS program from the Census Bureau. The top-left panel of Figure A4 compares the seasonally adjusted series (red) with the raw series (blue).

[30] Military 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 in this period.

2. We then scale the seasonally adjusted quarterly series so that fiscal-year averages align with the original fiscal-year values of military contracts. This fiscal-year adjustment is shown in the top-right panel of Figure A4. Because fiscal-year totals incorporate both new awards and contract modifications, the adjusted series (blue) lies above the seasonally adjusted new-awards series.

The bottom-left panel of Figure A4 compares the raw fiscal-year series of military contracts with the resulting quarterly series. By construction, fiscal-year averages of the quarterly data match the original FPSR fiscal-year values, while quarterly fluctuations capture variation in newly awarded, seasonally adjusted military contracts.

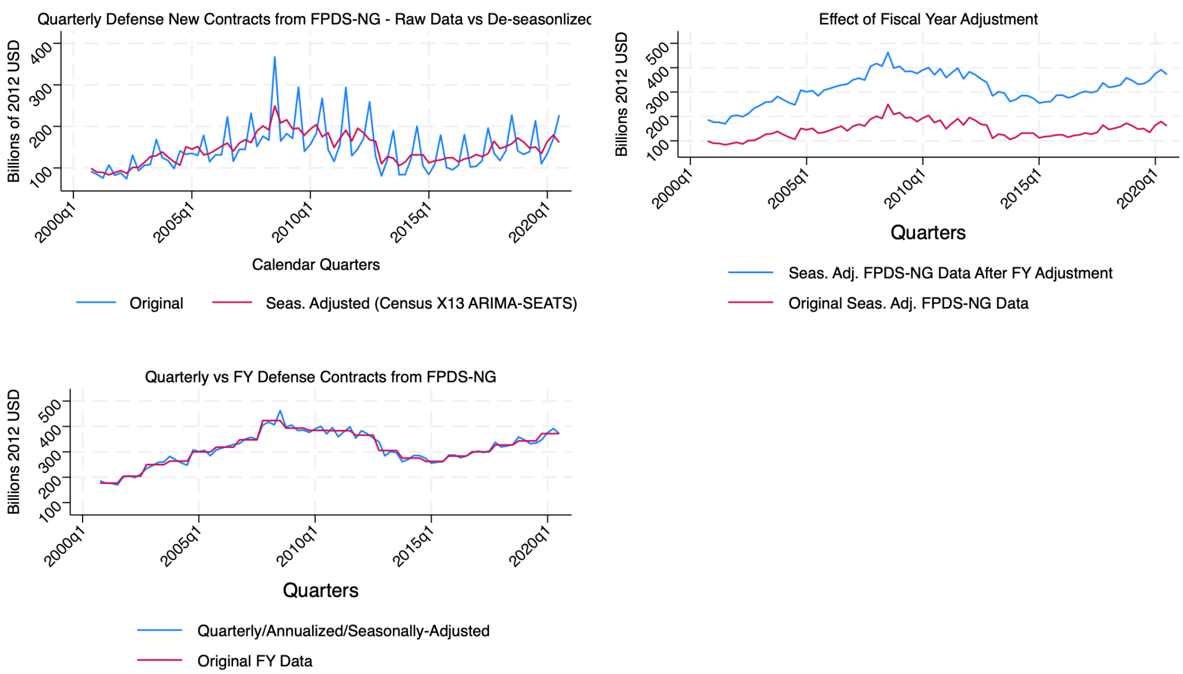

Figure A4 — Construction of Quarterly Military Contracts from the Federal Procurement Data System

Notes: Price deflator is the GDP price deflator.

A.4. Comparison of Contract Series with Dupor&Guerrero 2017 (DG17)

As a sanity check, we compare our newly constructed series of defense contracts, aggregated by fiscal year, with the annual series of defense contracts from Dupor and Guerrero (2017) (DG17), a benchmark in the literature. 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. Figure A5 shows the two 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.

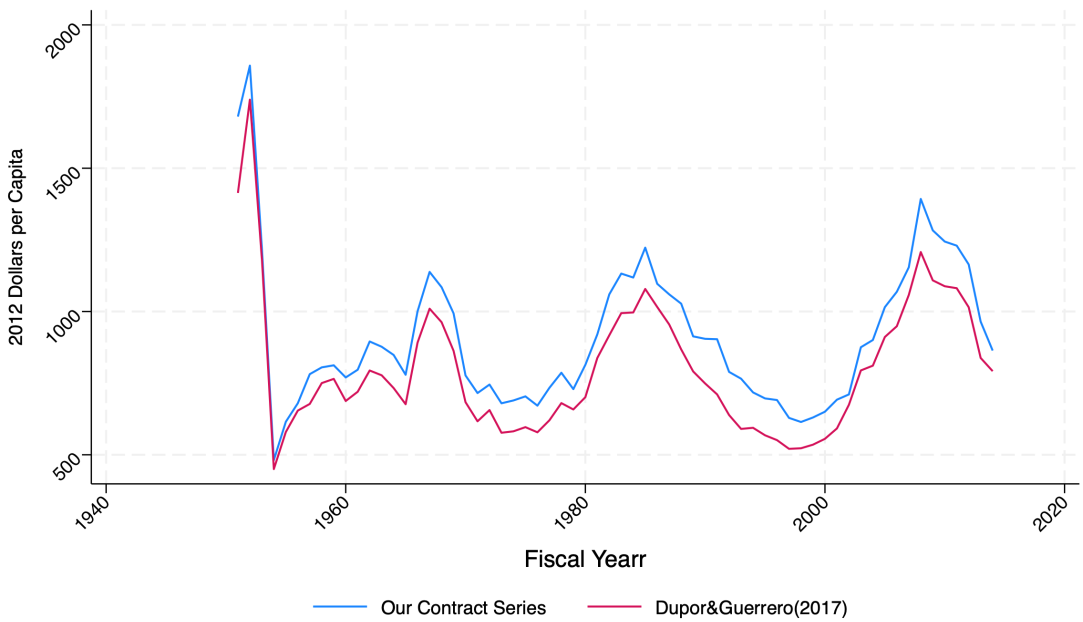

FIGURE A5 — CONTRACT SERIES IS CONSISTENT WITH DUPOR AND GUERRERO (2017) BUT IS SLIGHTLY HIGHER DUE TO INCLUSION OF SMALL (<$10K) CONTRACTS

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

B spending authorizations

This appendix describes the construction of the quarterly spending authorizations series used in the baseline quarterly analysis. The goal is to combine the broad coverage of annual defense budget authority with the higher-frequency timing information contained in quarterly military prime contracts. The resulting quarterly series preserves the annual level of budget authority within each federal fiscal year while allowing the timing of newly authorized spending to vary within the year.

B.1. Annual Defense Budget Authority

Our starting point is an annual fiscal-year series of defense budget authority. Budget authority is the legal authority of the federal government to incur obligations. For defense, it is the earliest broad measure of the government's commitment to spend and therefore the natural annual counterpart to the quarterly authorization series used in Section IV.

Annual totals by budget function are available from the Office of Management and Budget beginning in fiscal year 1976. To extend the series backward, we use the historical *Budgets of the United States* and record the previous year's actual totals rather than contemporaneous estimates or requested amounts. We begin the series in 1938 because, before the New Deal, the federal budget was understood more as an upper bound on spending than as a spending target, making earlier figures less comparable over time (Schick, 1990).

For the pre-1976 period, the terminology and reporting conventions change somewhat. The modern term "budget authority" only becomes standard in the late 1960s; earlier budgets often report "obligation authority," and before the 1952 budget appropriations and other authorizations are reported separately. For defense spending, however, these distinctions can be reconciled with relatively small adjustments. In particular, when the underlying budget documents report "new contract authority" and "liquidating appropriations," we convert obligation authority into a budget-authority concept by adding the former and subtracting the latter. This adjustment can be made from the 1945 budget onward; for 1938–1942 we use obligation authority directly because liquidating appropriations are not reported separately.

The reconstructed annual defense series also adjusts the historical classification of several spending categories to maintain a consistent concept of defense spending over time. In particular, the Coast Guard and the Selective Service System are treated as defense spending in all years, and atomic energy spending before 1974 is apportioned between defense and non-defense using the first year in which that split is reported separately. The resulting annual series is the fiscal-year defense budget authority series used throughout the paper.

B.2. From Annual Budget Authority to Quarterly spending authorizations

Budget authority is observed only at the fiscal-year level, whereas our main quarterly analysis requires a within-year timing measure. We therefore construct quarterly spending authorizations by combining annual budget authority with the quarterly path of military prime contracts. The guiding idea is simple:

budget authority pins down how much defense spending is authorized over a fiscal year, while contracts provide information about when those authorizations are translated into concrete commitments within that year.

We first map each calendar quarter into the corresponding federal fiscal year using the official fiscal-year definition. Before FY1977, the federal fiscal year ran from July 1 to June 30, so calendar quarters 3 and 4 belong to the following fiscal year. From FY1977 onward, the fiscal year runs from October 1 to September 30, so calendar quarter 4 belongs to the following fiscal year. The 1976 transition quarter (July–September 1976) is handled separately using the published transition-quarter budget-authority value, annualized so that it is comparable to the level interpretation of the fiscal-year series. When we later collapse the quarterly data to fiscal years, that transition quarter is omitted, matching the official fiscal-year accounting.

Within each fiscal year f, let BA_f denote annual defense budget authority and let MPC_t denote quarterly military prime contracts. We compute the fiscal-year average of quarterly contracts, $\overline{\text{MPC}}_f$, and define quarterly spending authorizations as

$$\text{SA}_t = \text{MPC}_t + \left(\text{BA}_{f(t)} - \overline{\text{MPC}}_{f(t)}\right),$$

where $f(t)$ is the fiscal year containing quarter t. This is exactly the interpolation rule implemented in the quarterly build. It shifts the level of the quarterly contract series by a fiscal-year-specific constant, preserving the within-year timing pattern of contracts while forcing the fiscal-year average of the quarterly series to equal annual budget authority.

Two properties of the resulting series are immediate. First, at quarterly frequency, spending authorizations inherit their within-fiscal-year timing from military contracts, which is what allows the series to anticipate NIPA spending in the quarterly analysis. Second, at fiscal-year frequency, the interpolated series collapses back to annual defense budget authority by construction. In other words, the fiscal-year values of spending authorizations are identical to budget authority, while the quarterly series provides the higher-frequency measure used in the Granger-causality tests and local-projection analysis in the main text.

C Defense Procurement and Defense Spending in the National Accounts

This appendix clarifies how defense procurement enters NIPA government spending and how the two NIPA-based defense series used in the paper are constructed. The key distinction is between a narrow procurement series, used as the NIPA counterpart to military contracts, and a broader defense-spending-less-CFC series, used as the NIPA counterpart to budget authority and spending authorizations.

C.1. Where Defense Procurement Appears in NIPA G

In the NIPAs, government purchases are recorded within government consumption expenditures and gross investment. For defense procurement, this means purchases of privately produced goods and services can appear in two places. Services and nondurable inputs purchased by the Department of Defense enter government consumption through federal defense intermediate goods and services. Purchases of long-lived assets, such as aircraft, missiles, structures, and software, enter government gross investment through federal defense fixed assets. Accordingly, the paper defines NIPA defense procurement spending as the sum of these two components, following Cox et al. (2024):

$$\text{Defense Procurement}_t = \text{Defense Intermediate Goods and Services}_t + \text{Defense Gross Investment}_t.$$

This is the series used throughout the paper as the NIPA counterpart to military contracts.

The accounting logic is central to the paper's timing argument. When a contractor begins producing a long-lived defense good, the associated output is recorded in the national accounts as work in progress and therefore appears in private inventories while production is ongoing. Only when the good is delivered and ownership transfers to the government does NIPA reclassify that output from inventories into government investment. By contrast, procurement-related services that are delivered contemporaneously can enter government consumption without such a delay. This distinction is why GDP can rise before measured NIPA G: production begins at the contract stage, but part of the associated government purchase is recorded only later, at delivery.

A useful way to think about the accounting is through a stylized missile purchase. The missile itself is recorded as federal defense equipment investment, while installation, maintenance, testing, and related services are recorded as intermediate goods and services purchased by the government. During production, the unfinished missile shows up in private inventories; at delivery, inventories fall and government investment rises by the same amount. Aggregate GDP does not jump at delivery because the production was already counted earlier; what changes at delivery is the composition of expenditure, from inventories to government spending.

C.2. The Two NIPA Defense Series Used in the Paper

The paper uses two related but distinct NIPA-based defense series. The first is the defense procurement series defined above. This is the natural counterpart to military prime contracts because both series refer to procurement activity and exclude categories such as military pay that are not driven directly by private prime-contract awards.

The second is NIPA defense spending less consumption of fixed capital (CFC). This broader series is the appropriate NIPA counterpart to defense budget authority and to our quarterly spending authorizations, because budget authority covers a wider set of defense expenditures than procurement alone. We therefore construct

$$\text{Defense Spending Less CFC}_t = \text{Federal Defense Spending}_t - \text{Defense CFC}_t.$$

Removing CFC aligns the NIPA concept more closely with the budget concept used for authorizations, since budget authority measures newly authorized spending rather than depreciation of previously accumulated government capital.

This distinction between the two NIPA series matches the empirical design in the main text. When we study the timing relationship between contracts and recorded spending, we compare military contracts to NIPA defense procurement spending. When we study budget authority and spending authorizations, we compare them to NIPA defense spending less CFC. In both cases, the relevant NIPA series is chosen to match the coverage of the administrative measure as closely as possible while preserving the accounting logic emphasized in the paper.

D Results using Military Contracts

This appendix reports the quarterly results obtained when military prime contracts, rather than spending authorizations, are used as the internal instrument. The goal is to show that the main identification results do not depend on the interpolation from annual budget authority to quarterly authorizations. Because contracts are narrower than spending authorizations and line up most closely with defense procurement spending, we continue to treat spending authorizations as the preferred baseline instrument in the main text. Nonetheless, the contract-based results tell the same qualitative story.

Identification with Military-Contract Shocks. Under the HBR transformation, let

$$mpc_t := \frac{\text{MPC}_t - \text{MPC}_{t-1}}{\text{GDP}_{t-1}}.$$

We define the military-contract shock as the innovation to mpc_t after projecting it on the same lagged conditioning set used in the baseline authorization specification:

$$\eta_t^{\text{MPC}} := mpc_t - \text{Proj}\left(mpc_t \;\middle|\; \underbrace{mpc_{t-1}, y_{t-1}, g_{t-1}, TB3_{t-1}, \tau_{t-1}}_{\text{Lag 1}}, \ldots, \underbrace{mpc_{t-4}, y_{t-4}, g_{t-4}, TB3_{t-4}, \tau_{t-4}}_{\text{Lag 4}} \right).$$

This contract-based shock is narrower than η_t^{SA}, but it is measured at the award stage of the spending process and therefore also captures the fiscal shock before it appears in measured NIPA G.

Military-Contract Shocks and BP shocks. Table D1 reports the analogue of Table 5 when the internal instrument is military prime contracts rather than spending authorizations. The results closely mirror the authorization-shock evidence. In both the full sample and the post-Korean-War sample, BP shocks are predicted by military-contract shocks, while the reverse is not true. After augmenting the specification with lags of defense news shocks, the same directional pattern remains. Thus replacing spending authorizations with contracts does not overturn the main identification point: innovations in NIPA G continue to embed delayed realizations of earlier military spending commitments.

Macroeconomic Responses to Military-Contract Shocks. Figure D1 reports the baseline HBR responses to a military-contract shock in the full postwar sample without news orthogonalization. The responses are qualitatively similar to those obtained with spending authorizations. GDP rises on impact, NIPA G reacts more slowly, and inventories account for most of the early output response. These patterns are consistent with military contracts capturing the same underlying timing mechanism as spending authorizations, even though contracts cover only the procurement-related portion of defense spending.

Figure D2 repeats the exercise after orthogonalizing military-contract shocks with respect to defense news shocks. The main pattern remains intact. The early response of GDP still arrives ahead of

TABLE D1 — FISCAL-SHOCK GRANGER-CAUSALITY TESTS: MILITARY-CONTRACT SHOCKS

Predicted	*Predictor*	*Sample*	*F*	*p-value*	*Predict?*
(1) Do Military-Contract Shocks Predict BP shocks?					
ξ_t^{BP} (BP shocks)	η_t^{MPC} (Military Contracts Shocks)	1947-2017	21.28	0.0000	Yes
η_t^{MPC}	ξ_t^{BP}	1947-2017	0.24	0.9166	No
ξ_t^{BP}	η_t^{MPC}	1956-2017	4.84	0.0009	Yes
η_t^{MPC}	ξ_t^{BP}	1956-2017	0.44	0.7817	No
(2) Do Military-Contract Shocks Predict BP shocks, Even After Controlling for News?					
ξ_t^{BP}	η_t^{MPC}	1947-2015	4.64	0.0013	Yes
η_t^{MPC}	ξ_t^{BP}	1947-2015	1.69	0.1518	No
ξ_t^{BP}	η_t^{MPC}	1956-2015	4.44	0.0018	Yes
η_t^{MPC}	ξ_t^{BP}	1956-2015	0.70	0.5951	No

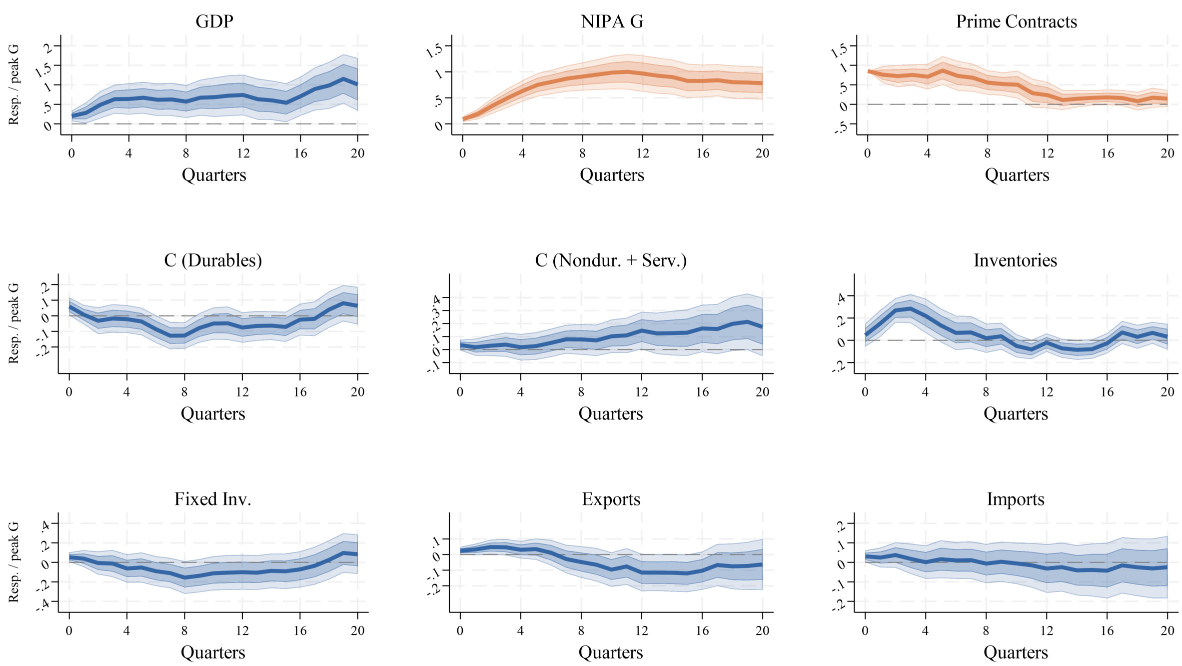

FIGURE D1 — IRFS TO A MILITARY-CONTRACT SHOCK

Notes: Baseline HBR specification, military-prime-contract shock, no defense-news controls, 1947Q1–2017Q4 sample. For comparability across outcomes, all values are normalized by the peak response of NIPA G. Bands are 68% and 90% confidence bands, constructed using heteroskedasticity-robust standard errors.

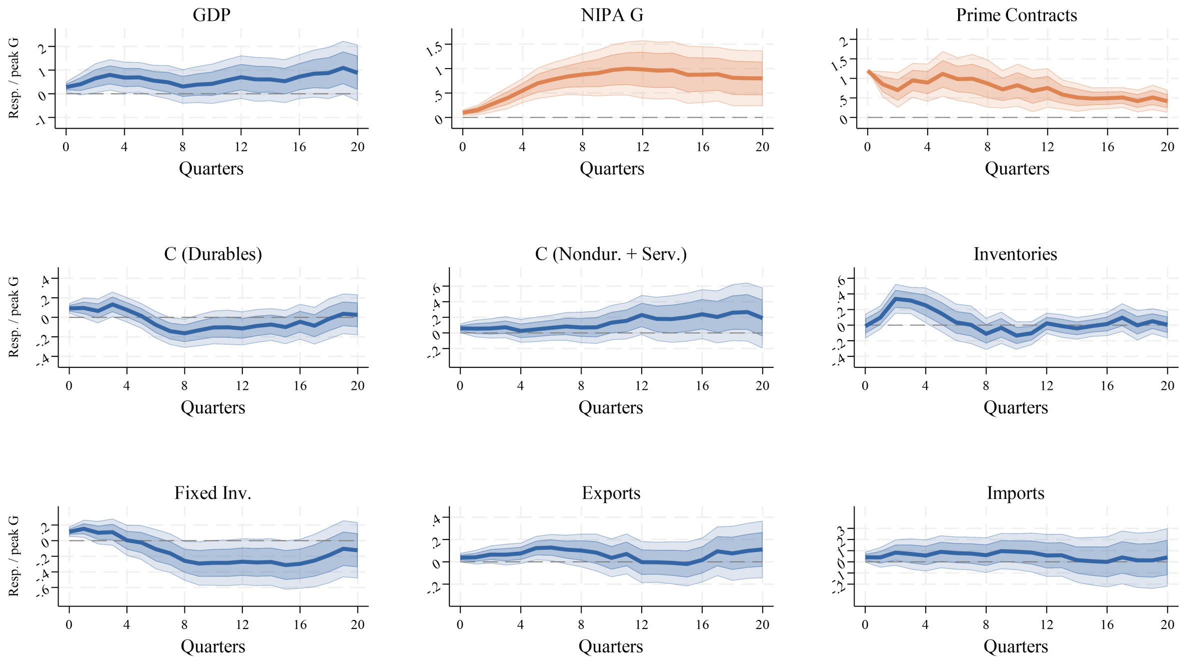

FIGURE D2 — LP RESPONSES OF GDP COMPONENTS TO MILITARY-CONTRACT SHOCKS ORTHOGONAL TO DEFENSE NEWS SHOCKS

Notes: Baseline HBR specification, military-prime-contract shock, with defense-news controls, 1947Q1–2015Q4 sample. For comparability across outcomes, all values are normalized by the peak response of NIPA G. Bands are 68% and 90% confidence bands, constructed using heteroskedasticity-robust standard errors.

measured NIPA G, and inventories remain the clearest margin through which output responds before government spending is recorded in the national accounts.

Taken together, the military-contract results support the same interpretation as the baseline authorization results. Contracts are narrower and therefore not our preferred baseline instrument, but they support the same identification logic and the same interpretation of the inventory channel. Multiplier results based on military-contract shocks are reported in Appendix G.3; as discussed in the main text, they deliver the same qualitative message, although the profile is flatter and somewhat lower at medium horizons.

E Robustness of Macroeconomic Responses

This appendix collects the two robustness exercises for Section IV.3 that are most relevant for the paper's main mechanism. The first excludes the Korean War by starting the sample in 1956Q1. The second applies the same sample restriction while also orthogonalizing spending-authorization shocks with respect to defense news. Together, these figures show that the central qualitative result is not driven by the Korean War and does not disappear once fiscal foresight is controlled for.

Figure E1 reports the baseline HBR responses to a spending-authorization shock in the 1956Q1–2017Q4 sample without defense-news controls. Relative to the full-sample baseline, confidence bands widen, as expected, because removing the Korean War substantially reduces the amount of defense-spending variation available for identification. Even so, the same qualitative pattern remains. GDP responds before measured NIPA G, inventories rise immediately and continue to account for most of the early output response, and the subsequent reversal in inventories becomes more visible in the shorter sample.

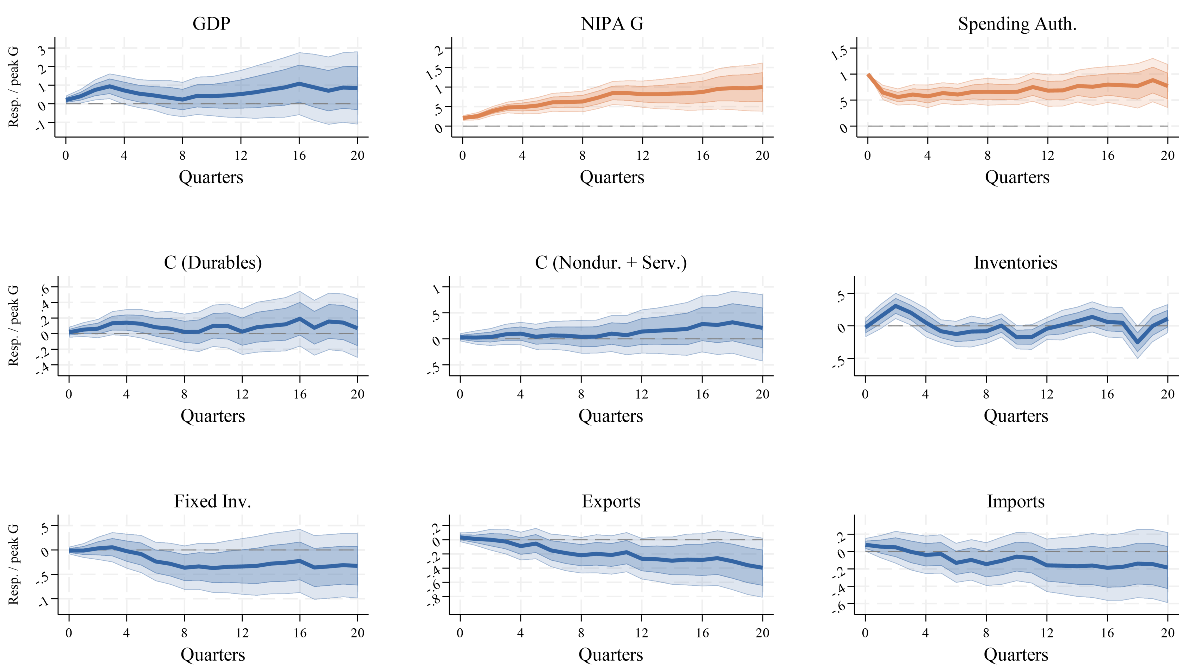

FIGURE E1 — IRFS TO A DEFENSE SPENDING AUTHORIZATION SHOCK: 1956 SAMPLE

Notes: Baseline HBR specification, spending-authorization shock, no defense-news controls, 1956Q1–2017Q4 sample. For comparability across outcomes, all values are normalized by the peak response of NIPA G. Bands are 68% and 90% confidence bands, constructed using heteroskedasticity-robust standard errors.

Figure E2 repeats the exercise after orthogonalizing spending-authorization shocks with respect to defense news in the same 1956Q1–2015Q4 sample (reflecting the availability of defense news shocks). The inventory channel again survives. GDP still rises earlier than NIPA G, and inventories remain

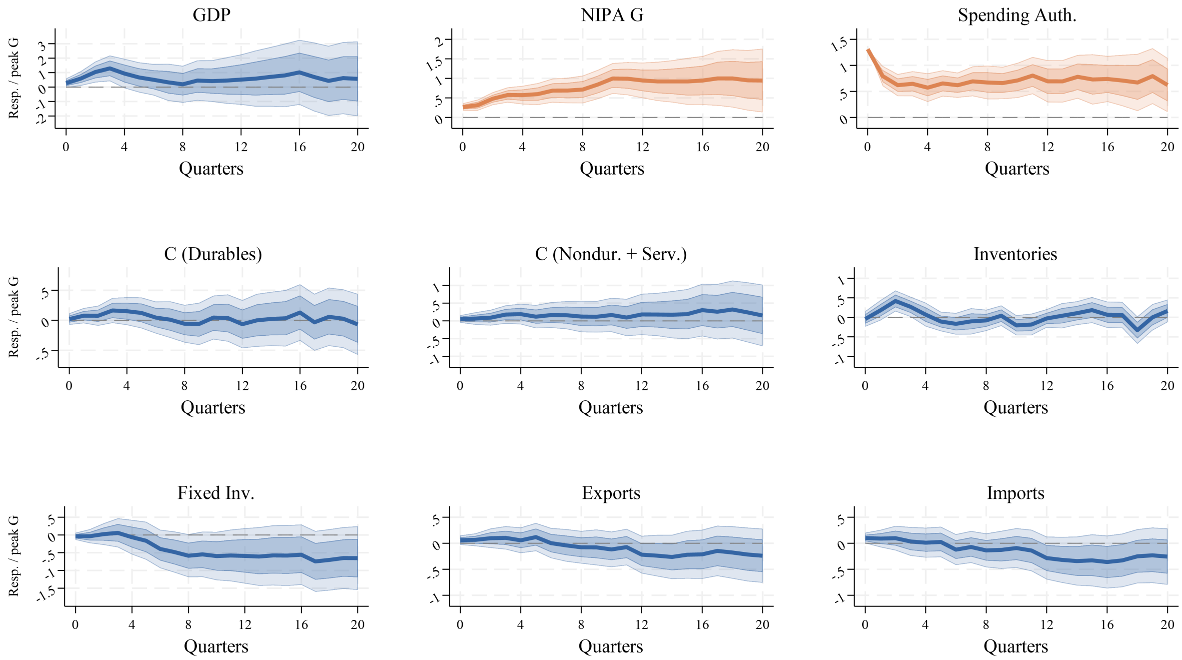

Figure E2 — LP Responses of GDP Components to Spending-Authorization Shocks Orthogonal to Defense News: 1956 Sample

Notes: Baseline HBR specification, spending-authorization shock, with defense-news controls, 1956Q1–2015Q4 sample. For comparability across outcomes, all values are normalized by the peak response of NIPA G. Bands are 68% and 90% confidence bands, constructed using heteroskedasticity-robust standard errors.

the clearest short-run margin through which output reacts before spending is recorded in the national accounts. Thus the timing mismatch documented in the paper is robust both to excluding the Korean War and to controlling for defense-news shocks.

Taken together, these robustness exercises reinforce the interpretation in the main text. Precision declines when the Korean War is excluded, but the core mechanism is unchanged: administrative measures that capture defense spending earlier in the process continue to produce an immediate inventory response and a faster GDP response than measured NIPA G.

F Accounting-Corrected Multiplier: Estimation and Standard Errors

This appendix documents the reparameterization used to estimate the accounting-corrected multiplier and to construct its pointwise confidence bands.

Reparameterization. The accounting-corrected multiplier at horizon H is

$$\mathcal{M}^c(H) := 1 + \frac{\mathcal{P}(H)}{1 + \mathcal{G}^r(H)},$$

where $\mathcal{P}(H)$ is the private-spending multiplier to authorizations and $\mathcal{G}^r(H)$ is the residual NIPA government-spending multiplier to authorizations. Define

$$\Theta_H := \sum_{h=0}^{H} \frac{\partial \text{PRSP}_{t+h}}{\partial \eta_t}, \qquad \Gamma_H := \sum_{h=0}^{H} \frac{\partial \text{SA}_{t+h}}{\partial \eta_t}, \qquad \Phi_H := \sum_{h=0}^{H} \frac{\partial \text{NIPA}^r_{t+h}}{\partial \eta_t},$$

where η_t denotes the authorization shock. Then

$$\mathcal{P}(H) = \frac{\Theta_H}{\Gamma_H}, \qquad \mathcal{G}^r(H) = \frac{\Phi_H}{\Gamma_H},$$

and therefore

$$\mathcal{M}^c(H) := 1 + \frac{\Theta_H/\Gamma_H}{1 + \Phi_H/\Gamma_H} := 1 + \frac{\Theta_H}{\Gamma_H + \Phi_H}.$$

Common-Sample Estimation. For the HBR specification, let

$$\eta_t \equiv sa_t := \frac{\text{SA}_t - \text{SA}_{t-1}}{\text{GDP}_{t-1}}.$$

For each horizon H, define the cumulative transformed variables

$$C_t^{prsp}(H) := \sum_{h=0}^{H} \frac{\text{PRSP}_{t+h} - \text{PRSP}_{t-1}}{\text{GDP}_{t-1}},$$
$$C_t^{auth}(H) := \sum_{h=0}^{H} \frac{\text{SA}_{t+h} - \text{SA}_{t-1}}{\text{GDP}_{t-1}},$$
$$C_t^{nipar}(H) := \sum_{h=0}^{H} \frac{\text{NIPA}^r_{t+h} - \text{NIPA}^r_{t-1}}{\text{GDP}_{t-1}}.$$

We then estimate the three reduced-form regressions

$$C_t^{prsp}(H) = \theta_H \eta_t + \mathbf{W}_t' \delta_H^{prsp} + \varepsilon_{t,H}^{prsp},$$
$$C_t^{auth}(H) = \gamma_H \eta_t + \mathbf{W}_t' \delta_H^{auth} + \varepsilon_{t,H}^{auth},$$
$$C_t^{nipar}(H) = \phi_H \eta_t + \mathbf{W}_t' \delta_H^{nipar} + \varepsilon_{t,H}^{nipar},$$

where $\mathbf{W}_t$ collects the same lagged conditioning variables used in the baseline specification. The OLS coefficients θ_H, γ_H, and ϕ_H are the empirical counterparts of Θ_H, Γ_H, and Φ_H, respectively. Hence, the accounting-corrected multiplier can be estimated as:

$$\widehat{\mathcal{M}}^c(H) := 1 + \frac{\hat{\theta}_H}{\hat{\gamma}_H + \widehat{\phi}_H}.$$

The key advantage of this reparameterization is that the three equations use the same regressor of interest and the same conditioning set. As a result, the joint covariance matrix of

$$\widehat{\beta}_H := \begin{pmatrix} \hat{\theta}_H \\ \hat{\gamma}_H \\ \widehat{\phi}_H \end{pmatrix}$$

can be estimated directly.

To see this, partial out the controls $\mathbf{W}_t$ from all three dependent variables and from η_t. Let $\tilde{\eta}_t$ denote the residual from regressing η_t on $\mathbf{W}_t$, and let

$$\widetilde{\mathbf{C}}_t(H) := \begin{pmatrix} \widetilde{C}_t^{prsp}(H) \\ \widetilde{C}_t^{auth}(H) \\ \widetilde{C}_t^{nipar}(H) \end{pmatrix}, \qquad \widetilde{\mathbf{u}}_{t,H} := \begin{pmatrix} \tilde{\varepsilon}_{t,H}^{prsp} \\ \tilde{\varepsilon}_{t,H}^{auth} \\ \tilde{\varepsilon}_{t,H}^{nipar} \end{pmatrix}.$$

By the Frisch–Waugh–Lovell theorem, the system can be written as

$$\widetilde{\mathbf{C}}_t(H) = \beta_H \tilde{\eta}_t + \widetilde{\mathbf{u}}_{t,H}, \qquad \beta_H := \begin{pmatrix} \theta_H \\ \gamma_H \\ \phi_H \end{pmatrix}.$$

The OLS estimator is therefore

$$\widehat{\beta}_H := \left(\sum_{t=1}^{T} \tilde{\eta}_t^2 \right)^{-1} \sum_{t=1}^{T} \tilde{\eta}_t \widetilde{\mathbf{C}}_t(H).$$

A heteroskedasticity-robust estimator of its covariance matrix is

$$\widehat{\mathbf{V}}_{\beta,H} := \left(\sum_{t=1}^{T} \tilde{\eta}_t^2 \right)^{-1} \left(\sum_{t=1}^{T} \tilde{\eta}_t^2 \, \widehat{\widetilde{\mathbf{u}}}_{t,H} \widehat{\widetilde{\mathbf{u}}}_{t,H}' \right) \left(\sum_{t=1}^{T} \tilde{\eta}_t^2 \right)^{-1}.$$

The off-diagonal elements of $\widehat{\mathbf{V}}_{\beta,H}$ deliver $\widehat{\text{Cov}}(\hat{\theta}_H, \hat{\gamma}_H)$, $\widehat{\text{Cov}}(\hat{\theta}_H, \widehat{\phi}_H)$, and $\widehat{\text{Cov}}(\hat{\gamma}_H, \widehat{\phi}_H)$.

Standard Errors and Confidence Bands. We can now apply the Δ-method. Define

$$m(\theta, \gamma, \phi) := 1 + \frac{\theta}{\gamma + \phi}, \qquad \widehat{\mathcal{M}}^c(H) = m(\hat{\theta}_H, \hat{\gamma}_H, \widehat{\phi}_H).$$

Its gradient is

$$\nabla m(\theta,\gamma,\phi) := \begin{pmatrix} \dfrac{1}{\gamma+\phi} \\ -\dfrac{\theta}{(\gamma+\phi)^2} \\ -\dfrac{\theta}{(\gamma+\phi)^2} \end{pmatrix}.$$

If

$$\sqrt{T}\left(\widehat{\beta}_H - \beta_H\right) \xrightarrow{d} \mathcal{N}\big(0, \Omega_{\beta,H}\big),$$

then

$$\sqrt{T}\left(\widehat{\mathcal{M}}^c(H) - \mathcal{M}^c(H)\right) \xrightarrow{d} \mathcal{N}\big(0, \nabla m(\beta_H)'\,\Omega_{\beta,H}\,\nabla m(\beta_H)\big).$$

Accordingly, the Δ-method variance estimator is

$$\widehat{\mathrm{Var}}\Big(\widehat{\mathcal{M}}^c(H)\Big) := \nabla m(\widehat{\beta}_H)'\,\widehat{\mathbf{V}}_{\beta,H}\,\nabla m(\widehat{\beta}_H).$$

Writing this expression out explicitly yields

$$\begin{aligned}\widehat{\mathrm{Var}}\Big(\widehat{\mathcal{M}}^c(H)\Big) = \frac{\widehat{\mathrm{Var}}(\hat{\theta}_H)}{(\hat{\gamma}_H+\widehat{\phi}_H)^2} + \frac{\hat{\theta}_H^2}{(\hat{\gamma}_H+\widehat{\phi}_H)^4}\Big[\widehat{\mathrm{Var}}(\hat{\gamma}_H) + \widehat{\mathrm{Var}}(\widehat{\phi}_H) + 2\widehat{\mathrm{Cov}}(\hat{\gamma}_H,\widehat{\phi}_H)\Big] \\ - \frac{2\hat{\theta}_H}{(\hat{\gamma}_H+\widehat{\phi}_H)^3}\Big[\widehat{\mathrm{Cov}}(\hat{\theta}_H,\hat{\gamma}_H) + \widehat{\mathrm{Cov}}(\hat{\theta}_H,\widehat{\phi}_H)\Big].\end{aligned}$$

This expression makes clear why one cannot recover valid confidence bands by combining the marginal bands of $\widehat{\mathcal{P}}(H)$ and $\widehat{\mathcal{G}}^r(H)$ alone: the covariance terms are an essential part of the sampling uncertainty of the accounting-corrected multiplier.

Finally, a pointwise $(1-\alpha)$ confidence interval for $\mathcal{M}^c(H)$ is given by

$$\widehat{\mathcal{M}}^c(H) \pm c_{1-\alpha/2}\sqrt{\widehat{\mathrm{Var}}\Big(\widehat{\mathcal{M}}^c(H)\Big)},$$

where $c_{1-\alpha/2}$ denotes the appropriate critical value from either the standard normal distribution or a finite-sample t approximation.

Implementation. In the empirical exercises, we estimate the three reduced-form equations on the exact common sample at each horizon, using the same shock and the same conditioning set in all three regressions. This makes the covariance terms in the variance expression above directly estimable and prevents the confidence bands from being driven by sample mismatches across equations. The appendix figures report pointwise 68% and 90% confidence bands based on the Δ-method standard errors derived here.

G Additional Multiplier Results and Robustness

This appendix collects the exact estimates and supporting robustness figures for Section V.

G.1. Exact Estimates and First-Stage Strength

TABLE G1 — CUMULATIVE FISCAL MULTIPLIERS: EXACT ESTIMATES AND EFFECTIVE F-STATISTICS

	(A) Full Sample						(B) Post-Korean-War Sample					
	No News			With News			No News			With News		
Horizon	Auth.	BP	Eff. F	Auth.	BP	Eff. F	Auth.	BP	Eff. F	Auth.	BP	Eff. F
0	4.912	0.716	1.081	4.194	0.787	4.951	0.968	1.062	24.079	1.018	1.040	20.151
	(4.326)	(0.182)		(1.948)	(0.257)		(0.600)	(0.256)		(0.631)	(0.254)	
1	3.184	0.598	2.828	3.927	0.560	6.112	1.318	0.962	21.577	1.479	0.973	17.937
	(1.618)	(0.206)		(1.479)	(0.328)		(0.691)	(0.338)		(0.725)	(0.340)	
2	2.381	0.582	7.699	3.546	0.442	7.916	1.595	1.016	25.423	1.790	1.032	21.693
	(0.745)	(0.221)		(1.168)	(0.396)		(0.721)	(0.401)		(0.759)	(0.408)	
3	2.008	0.498	17.523	3.136	0.242	10.195	1.731	0.964	28.362	1.951	0.981	24.297
	(0.505)	(0.228)		(0.952)	(0.442)		(0.752)	(0.427)		(0.803)	(0.439)	
4	1.724	0.402	34.796	2.538	0.066	13.454	1.660	0.921	26.324	1.870	0.936	22.410
	(0.420)	(0.237)		(0.793)	(0.490)		(0.786)	(0.440)		(0.857)	(0.460)	
5	1.509	0.340	57.050	2.032	-0.049	16.741	1.522	0.896	24.077	1.709	0.911	20.148
	(0.361)	(0.241)		(0.654)	(0.543)		(0.785)	(0.445)		(0.871)	(0.470)	
6	1.364	0.325	75.789	1.680	-0.084	18.541	1.361	0.917	23.053	1.516	0.937	19.027
	(0.322)	(0.242)		(0.562)	(0.592)		(0.756)	(0.441)		(0.846)	(0.470)	
7	1.247	0.341	85.979	1.428	-0.080	19.277	1.221	0.923	21.953	1.345	0.946	17.998
	(0.289)	(0.245)		(0.497)	(0.643)		(0.745)	(0.444)		(0.841)	(0.476)	
8	1.151	0.355	84.400	1.231	-0.075	18.609	1.092	0.885	20.895	1.187	0.908	17.035
	(0.261)	(0.247)		(0.455)	(0.682)		(0.733)	(0.443)		(0.830)	(0.476)	
9	1.098	0.371	76.823	1.106	-0.028	17.240	1.018	0.868	20.337	1.092	0.888	16.980
	(0.238)	(0.248)		(0.428)	(0.703)		(0.726)	(0.447)		(0.815)	(0.480)	
10	1.060	0.392	68.670	1.028	0.058	15.948	0.938	0.883	20.891	0.991	0.904	17.599
	(0.221)	(0.250)		(0.406)	(0.709)		(0.697)	(0.446)		(0.780)	(0.482)	
11	1.035	0.398	61.381	0.998	0.112	14.859	0.887	0.895	21.195	0.922	0.916	17.884
	(0.208)	(0.254)		(0.389)	(0.721)		(0.682)	(0.447)		(0.763)	(0.485)	
12	1.020	0.392	54.790	0.998	0.135	13.619	0.860	0.903	20.955	0.881	0.925	17.653
	(0.199)	(0.260)		(0.376)	(0.735)		(0.683)	(0.447)		(0.767)	(0.487)	
13	1.004	0.383	49.481	1.004	0.139	12.596	0.852	0.911	20.394	0.858	0.931	17.073
	(0.195)	(0.268)		(0.369)	(0.757)		(0.694)	(0.451)		(0.782)	(0.493)	
14	0.993	0.386	45.592	1.012	0.159	11.854	0.857	0.921	19.856	0.849	0.935	16.395
	(0.191)	(0.276)		(0.362)	(0.772)		(0.703)	(0.456)		(0.800)	(0.499)	
15	0.982	0.403	41.927	1.012	0.206	11.031	0.873	0.927	19.202	0.851	0.938	15.610
	(0.189)	(0.283)		(0.359)	(0.772)		(0.712)	(0.461)		(0.816)	(0.506)	
16	0.983	0.440	39.498	1.019	0.277	10.431	0.903	0.937	18.570	0.869	0.941	14.846
	(0.187)	(0.286)		(0.358)	(0.758)		(0.717)	(0.467)		(0.829)	(0.514)	
17	0.996	0.482	37.699	1.029	0.344	10.000	0.914	0.946	18.035	0.865	0.944	14.136
	(0.185)	(0.289)		(0.357)	(0.743)		(0.720)	(0.471)		(0.839)	(0.521)	
18	1.013	0.520	36.087	1.039	0.389	9.537	0.900	0.946	17.313	0.835	0.938	13.287
	(0.184)	(0.293)		(0.359)	(0.734)		(0.724)	(0.477)		(0.851)	(0.530)	
19	1.040	0.553	34.715	1.060	0.421	9.133	0.903	0.955	16.547	0.824	0.943	12.415
	(0.185)	(0.297)		(0.363)	(0.731)		(0.728)	(0.480)		(0.865)	(0.537)	
20	1.056	0.571	33.447	1.059	0.422	8.769	0.884	0.961	15.669	0.789	0.946	11.461
	(0.188)	(0.303)		(0.373)	(0.739)		(0.734)	(0.482)		(0.884)	(0.542)	

Notes: Each row reports the cumulative multiplier at horizon H. Authorization columns instrument cumulative NIPA G with spending authorizations. BP columns use the BP-style innovation to NIPA G, obtained from the same recursive specification after omitting spending authorizations. Eff. F reports the Montiel Olea–Pflueger effective F-statistic for the authorization specification. Heteroskedasticity-robust standard errors are reported in parentheses beneath point estimates. The HBR no-news columns use samples ending in 2017Q4; the with-news columns end in 2015Q4 because of defense-news data availability.

Table G1 reports the full set of HBR standard multipliers. Authorization-based multipliers are generally larger than BP-based multipliers, especially at short and medium horizons, while the effective first stage strengthens quickly outside the impact horizon.

G.2. Robustness of Standard Multipliers

The next three figures report the standard multiplier robustness exercises that are most informative for the paper's main interpretation. Restricting the sample to start in 1956 narrows the role of the Korean War, while orthogonalizing with respect to defense news shows that the gap between authorization-based and BP multipliers is not driven by anticipatory shocks alone.

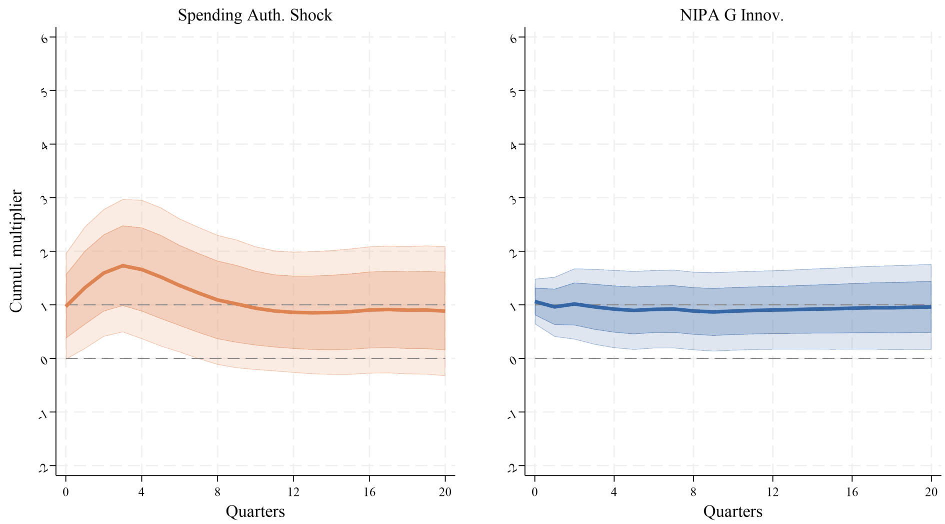

FIGURE G1 — STANDARD CUMULATIVE MULTIPLIERS: 1956 SAMPLE

Notes: Baseline HBR specification, spending-authorization shock on the left and BP shock on the right, no defense-news controls, 1956Q1–2017Q4 sample. Confidence bands are 68% and 90%.

Taken together, these figures show that the basic multiplier comparison is stable across the main identification choices. Authorization-based multipliers remain around one or above one at medium horizons, whereas BP multipliers remain substantially smaller and can be close to zero at short horizons.

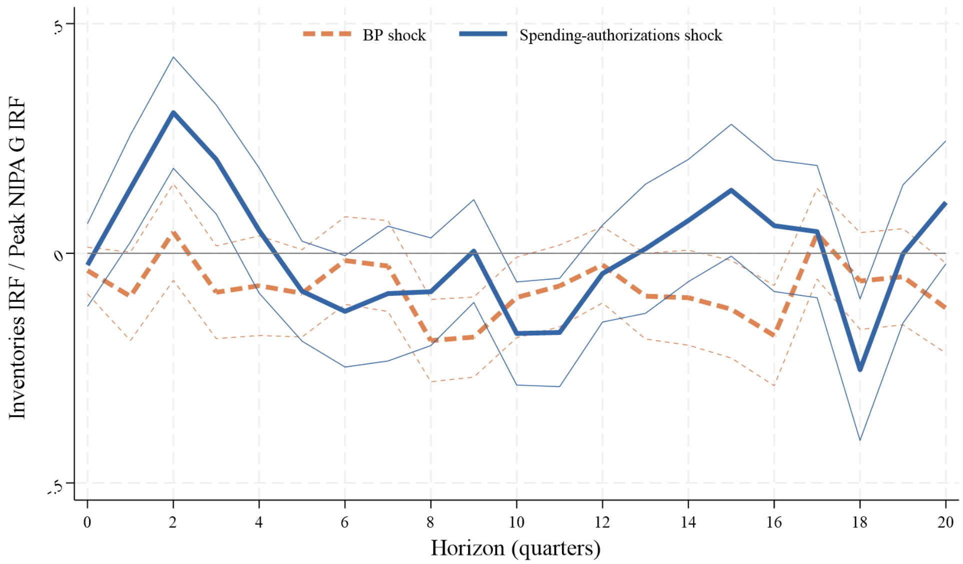

Figure G2 — Inventories, BP Shocks, and Spending-Authorization Shocks: 1956 Sample

Notes: Baseline HBR specification, no defense-news controls, 1956Q1–2017Q4 sample. The orange dashed line plots the IRF of inventories to a BP shock, normalized by the peak response of NIPA G to that BP shock. The blue solid line plots the IRF of inventories to a spending-authorization shock, normalized by the peak response of NIPA G to that spending-authorization shock. Dashed lines of the same color denote 68% confidence bands.

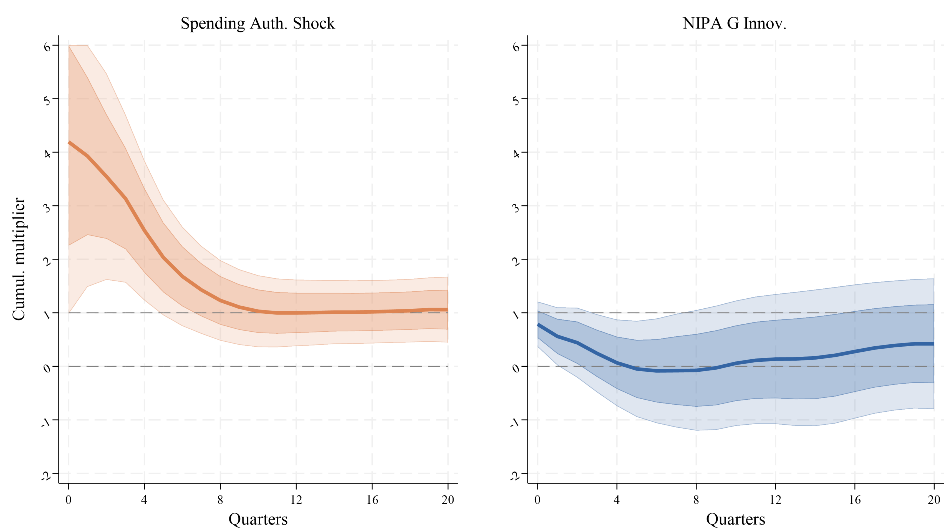

Figure G3 — Standard Cumulative Multipliers with News Orthogonalization

Notes: Baseline HBR specification, spending-authorization shock on the left and BP shock on the right, with defense-news controls, 1947Q1–2015Q4 sample. Confidence bands are 68% and 90%.

G.3. Robustness of the Accounting-Corrected Multiplier

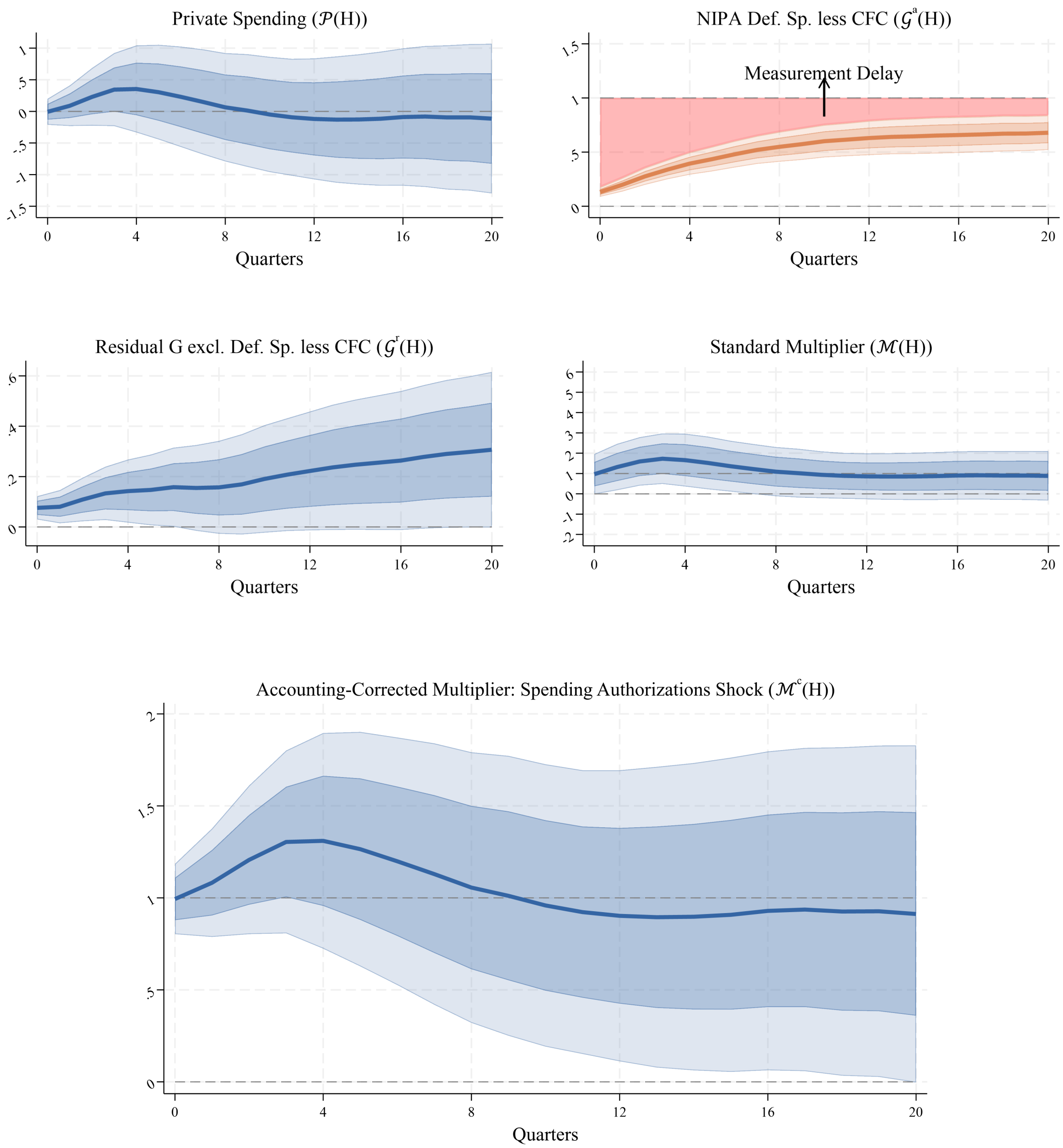

FIGURE G4 — ACCOUNTING-CORRECTED MULTIPLIER ROBUSTNESS: 1956 SAMPLE

Notes: Left panel: multiplier breakdown. Right panel: accounting-corrected multiplier. Baseline HBR specification, spending-authorization shock, no defense-news controls, 1956Q1–2017Q4 sample.

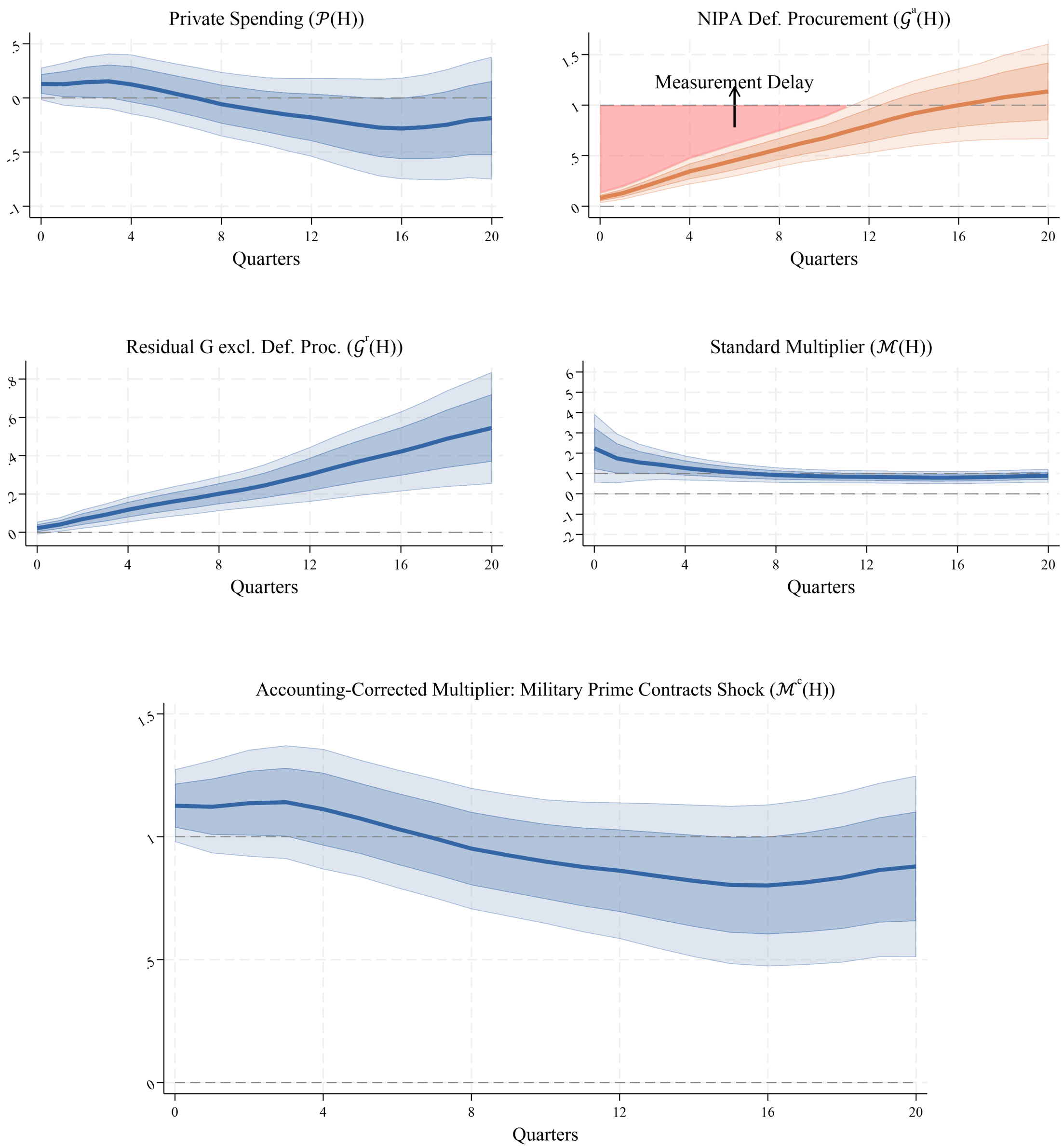

FIGURE G5 — ACCOUNTING-CORRECTED MULTIPLIER ROBUSTNESS: MILITARY PRIME CONTRACTS

Notes: Left panel: multiplier breakdown. Right panel: accounting-corrected multiplier. Baseline HBR specification, military prime-contract shock, no defense-news controls, 1947Q1–2017Q4 sample.

The accounting-corrected multiplier is robust to the two main appendix exercises discussed in the main text. Excluding the Korean War yields wider confidence bands but similar dynamics. Replacing spending authorizations with military prime contracts produces a somewhat flatter profile, while preserving the central result that the corrected multiplier remains near or above one at short horizons.

G.4. Selected Exact Estimates of the Accounting-Corrected Multiplier

TABLE G2 — ACCOUNTING-CORRECTED MULTIPLIER: SELECTED EXACT ESTIMATES

	Baseline	Post-Korean-War	Prime-Contract Shock
Horizon	Auth. (1947)	Auth. (1956)	MPC (1947)
0 (Impact)	1.191	0.994	1.127
	(0.072)	(0.116)	(0.089)
4 (one year)	1.211	1.310	1.112
	(0.103)	(0.355)	(0.148)
8 (two years)	1.072	1.056	0.952
	(0.118)	(0.445)	(0.149)
12 (three years)	1.012	0.903	0.862
	(0.120)	(0.479)	(0.168)
20 (five years)	1.043	0.913	0.879
	(0.144)	(0.555)	(0.223)

Notes: Entries report selected values of the accounting-corrected multiplier $\mathcal{M}^c(H)$ for the HBR specification without defense-news controls. Heteroskedasticity-robust Δ-method standard errors are reported in parentheses beneath point estimates. The baseline and post-Korean-War columns use spending-authorization shocks, while the final column uses military prime contracts. Reported horizons are $H = 0, 4, 8, 12, 20$.

Table G2 reports selected exact values for the baseline accounting-corrected multiplier together with the two robustness cases shown above. Reporting a few benchmark horizons makes transparent that the corrected multiplier is well defined on impact and remains close to one at longer horizons.

H Industry Analysis: Inventory Responses Are Concentrated in Defense-Exposed Manufacturing

This appendix provides supporting evidence for the mechanism emphasized in the main text. If the aggregate inventory response reflects work in progress by defense contractors, it should be concentrated in manufacturing industries that are most exposed to federal purchases rather than spread uniformly across sectors.

H.1. Measuring Manufacturing Sectors' Reliance on Federal Purchases

We combine information from the BEA Make and Use tables for more than 60 non-government sectors over 1963–1996. Following Horowitz and Planting (2009), we construct exposure measures that capture both direct sales to the federal government and indirect exposure through downstream input-output linkages.

Direct purchases. Let A_t denote the direct-requirement matrix and let $\text{SALES}_{i\to G,t}$ denote sales from industry i to the federal government. We begin with the vector of direct government sales 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, and G denotes the federal general government. The subscript 0 indicates that this measure only includes direct sales to the government. $\text{SALES}_{i\to G,t}$ includes both government gross investment, which appears as final use in the Use tables, and direct requirements. We report the time-average values of $\boldsymbol{\gamma}_{0,t}$ in Table H1.

Government indirect purchases. Following Nekarda and Ramey (2011), we also incorporate downstream input-output linkages. Let the (i, j) element of the yearly $n \times n$ input-output matrix A_t be

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

We then construct the vector of direct and first-order indirect sales shares as

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

Its ith element is

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

TABLE H1 — MANUFACTURING RELIANCE ON FEDERAL PURCHASES IS HIGHLY CONCENTRATED

Sector (i)	*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

Notes: The last column divides $\gamma_{2,i}$ by its maximum value, which is attained by other transportation equipment. The weights θ_i used in the empirical analysis therefore include second-order connections and are normalized to lie in $[0, 1]$.

We report the time-average of $\boldsymbol{\gamma}_{1,t}$ in the fourth column of Table H1. Similarly, we construct direct, first-order, and second-order indirect sales to the government as shares of total output,

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

We report the time-average values of $\boldsymbol{\gamma}_{2,t}$ in the fifth column of Table H1. Because second-order downstream connections add little beyond first-order connections, that is, $\gamma_1 \approx \gamma_2$, we use second-order exposure as our preferred measure and do not go further down the input-output network.

Industry weights. We define industry exposure weights as

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

This normalization makes the most defense-exposed industry equal to one. Table H1 shows that federal-purchase exposure is highly concentrated. Other transportation equipment is by far the most exposed sector, followed by computer and electronic products, whereas sectors such as wood products or food and beverage manufacturing have very limited exposure. This concentration is exactly what one would expect if military buildups primarily operate through a narrow set of defense-producing industries.

H.2. Heterogeneous Inventory Responses Across Manufacturing Industries

We next ask whether the aggregate inventory response documented in the main text is concentrated in the industries that are most exposed to federal purchases.

Data. We use monthly BEA data to construct a panel of real inventories for 18 manufacturing industries from January 1959 to December 1997. Because the series are reported in chained dollars, the industry-level results are best interpreted qualitatively rather than as exact quantitative decompositions of the aggregate inventory response.

Empirical specification. We estimate a panel local-projection with lags and expressed in long-differences:

$$\frac{\text{Invt}_{i,t+h} - \text{Invt}_{i,t-1}}{\text{Invt}_{i,t-1}} = \lambda_{ih} + \underbrace{\alpha_h \cdot \text{War}_t + \beta_h \cdot \text{War}_t \cdot \theta_i}_{=(\alpha_h+\beta_h \cdot \theta_i)\cdot \text{War}_t} + \sum_{p=1}^{12} \varphi_{ph} \cdot \frac{\text{Invt}_{i,t-p} - \text{Invt}_{i,t-p-1}}{\text{Invt}_{i,t-p-1}} + \varepsilon_{i,t+h}. \quad (11)$$

Here $h = 0, 1, \ldots, 84$ (7 years), $\text{Invt}_{i,t}$ is real inventories in industry i at month t, λ_{ih} is an industry fixed effect which absorbs time invariant sectoral characteristics, such as reliance θ_i. Finally, War_t is a monthly military-buildup indicator. Consistent with Ramey and Shapiro (1998) and Eichenbaum and Fisher (2005), we use a weighted event measure that takes value 1 in March 1965 and 0.3 in January 1984. Results using unweighted war events are qualitatively very similar, and they are reported below.

In this setup, we prefer war dates rather than defense news shocks because of two reasons. First, they are well suited for an event case studies aimed at comparing the different dynamics of sectoral inventories in sectors differentially exposed to federal purchases, during a military build-up. Second, war dates can be aggregated naturally to monthly frequency.

The coefficient α_h captures the response of a sector with no government exposure, that is, $\theta_i = 0$. The coefficient β_h captures how that response changes with exposure to federal purchases. A positive estimate of β_h therefore implies that inventories rise more strongly in sectors that are more closely tied to the government.

Results. Figure H1 reports the baseline estimates. The left panel plots α_h, the response of a sector with $\theta_i = 0$, while the right panel plots $\alpha_h + \beta_h$, the response of the most exposed sector, normalized to $\theta_i = 1$.

The estimates imply a sharp difference across sectors. Inventories in industries with little government exposure respond very little. By contrast, inventories in the most exposed industries rise substantially during military buildups. This pattern is difficult to reconcile with a broad-based demand expansion that would lift inventories uniformly across manufacturing. Instead, it points to a defense-production mechanism in which inventories capture work in progress at firms receiving military orders.

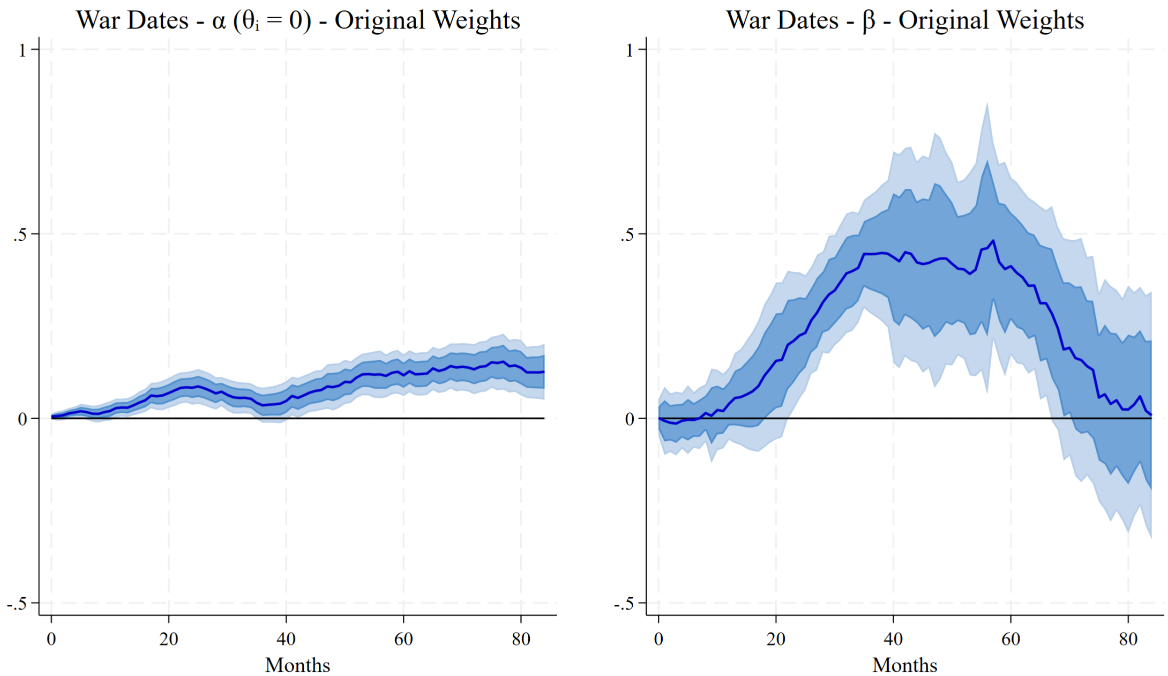

Figure H1 — Response of Sectoral Inventories to War Events

Notes: The left panel reports estimates of α_h, the response when $\theta_i = 0$. The right panel reports estimates of $\alpha_h + \beta_h$, the response when $\theta_i = 1$. The sample runs from January 1959 to December 1997 and covers 18 manufacturing industries. Confidence bands are 68% and 90%. Standard errors are obtained by bootstrap using Stata's `vce(boot)` routine for `xtreg`. Weights are normalized by the maximum exposure weight, corresponding to other transportation equipment manufacturing. The unit of real inventories is millions of chained 2005 dollars.

Robustness. Figure H2 summarizes two robustness checks. The first column reproduces the baseline estimates (i.e., those of Figure H1). The second column randomly reallocates the exposure weights θ_i across industries. The third column replaces war events with a monetary policy shock, constructed from the narrative series of Romer and Romer (2004) updated by Wieland and Yang (2020). In both exercises, the heterogeneous response disappears: once weights are randomized or the shock is unrelated to military buildups, the estimated differential response across industries becomes statistically indistinguishable from zero.

These results suggest that the baseline heterogeneity is not driven by generic business-cycle sensitivity or by spurious correlation between inventories and industry weights. Instead, it is specific to military buildups and strongest in the industries most exposed to federal purchases.[31]

Robustness: Unweighted war events. Because the sample contains only two major military buildups, we also repeat the baseline exercise without weighting the two events. Figure H3 shows that the qualitative pattern is unchanged.

[31] We thank Juan Herreno for suggesting these robustness tests.

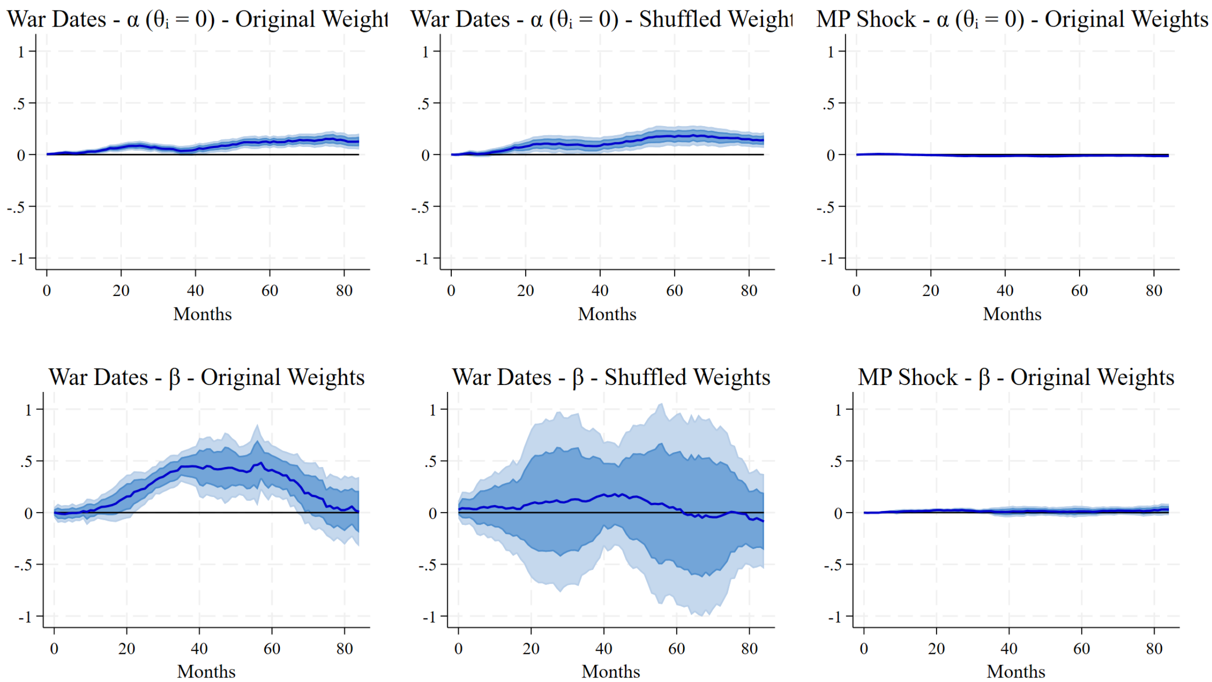

FIGURE H2 — RESPONSE OF SECTORAL INVENTORIES TO WAR EVENTS ROBUSTNESS

Notes: The first column reproduces the baseline estimates from Figure H1. The second column randomly reallocates industry exposure weights. The third column replaces war events with monetary policy shocks.

Summary. The industry evidence supports the paper's main mechanism. During military buildups, inventories rise disproportionately in the manufacturing industries most exposed to federal purchases. This is exactly what one would expect if inventories are recording unfinished defense production that has not yet appeared in measured NIPA government spending. If instead the aggregate inventory response reflected a broad macroeconomic expansion unrelated to defense production, the response would be much more uniform across industries, which is not what we find.

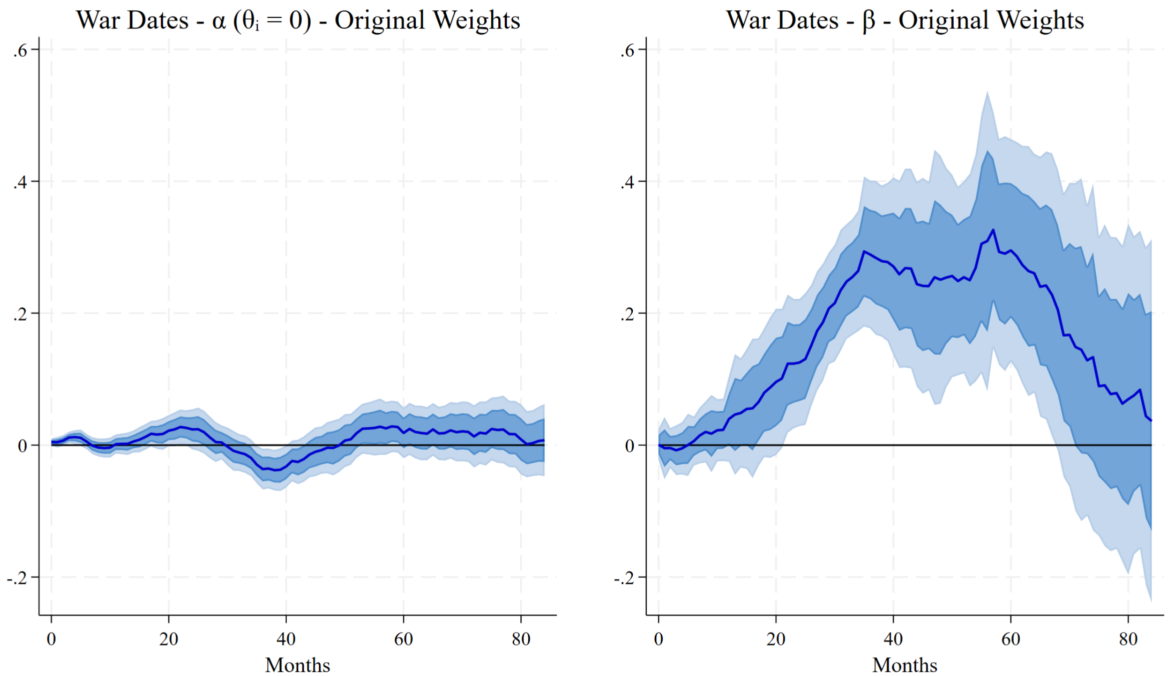

FIGURE H3 — RESPONSE OF SECTORAL INVENTORIES TO WAR EVENTS
UNWEIGHTED WAR EVENTS

Notes: Same as in Figure H1, except that the war-event indicator gives equal weight to the two military buildups.

References

Alloza, **Mario**, **Domınguez-Dıaz**, **Ruben**, **Dura**, **Pau**, and **Varotto**, **Iacopo** (Mar. 2026). The Macroeconomic Effects of Defence Expenditure: Evidence from Spain. *Working Paper*. 2026.

Antolin-Diaz, **Juan** and **Surico**, **Paolo** (July 1, 2025). The Long-Run Effects of Government Spending. *American Economic Review*. 2025. 115, pp. 2376–2413.

Ascari, **Guido**, **Beck-Friis**, **Peder**, **Florio**, **Anna**, and **Gobbi**, **Alessandro** (Mar. 2023). Fiscal foresight and the effects of government spending: It's all in the monetary-fiscal mix. *Journal of Monetary Economics*. 2023. 134, pp. 1–15.

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.

Barnichon, **Regis**, **Debortoli**, **Davide**, and **Matthes**, **Christian** (Jan. 10, 2022). Understanding the Size of the Government Spending Multiplier: It's in the Sign. *The Review of Economic Studies*. 2022. 89, pp. 87–117.

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

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

Binder, **Carola Conces** and **Brunet**, **Gillian** (Nov. 12, 2021). Inflation expectations and consumption: Evidence from 1951. *Economic Inquiry*. 2021, ecin.13051.

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.

Briganti, **Edoardo** (Oct. 2023). On the Effects of Government Purchases and Their Transmission Mechanism. *UCSD Manuscript*. 2023.

Brockwell, **Peter J.** and **Davis**, **Richard A.** (1991). *Time Series: Theory and Methods*. Springer Series in Statistics. New York, NY: Springer New York.

Cox, **Lydia**, **Müller**, **Gernot J.**, **Pastén**, **Ernesto**, **Schoenle**, **Raphael**, and **Weber**, **Michael** (Oct. 1, 2024). Big *G*. *Journal of Political Economy*. 2024. 132, pp. 3260–3297.

D'Alessandro, **Antonello**, **Fella**, **Giulio**, and **Melosi**, **Leonardo** (Aug. 2019). Fiscal Stimulus with Learning-by-Doing. *International Economic Review*. 2019. 60, pp. 1413–1432.

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. 1, 2012). A Quarterly, Utilization-Adjusted Series on Total Factor Productivity. *Federal Reserve Bank of San Francisco, Working Paper Series*. 2012, pp. 01–28.

Fieldhouse, **Andrew J** and **Mertens**, **Karel** (Feb. 2026). The Returns to Government R&D: Evidence from U.S. Appropriations Shocks. *Conditionally Accepted, American Economic Review*. 2026.

Fisher, **Jonas** and **Peters**, **Ryan** (May 1, 2010). Using Stock Returns to Identify Government Spending Shocks. *The Economic Journal*. 2010. 120, pp. 414–436.

Forni, **Mario** and **Gambetti**, **Luca** (Mar. 2016). Government spending shocks in open economy VARs. *Journal of International Economics*. 2016. 99, pp. 68–84.

Gross, **Daniel** and **Sampat**, **Bhaven** (2023). America, Jump-started: World War II R&D and the Takeoff of the U.S. Innovation System. *American Economic Review*. 2023. 113, pp. 3323–3356.

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

Hall, **Robert E.** (2009). By How Much Does GDP Rise If the Government Buys More Output? *Brookings Papers on Economic Activity*. 2009, pp. 183–221.

Hamilton, **James D.** (1994). *Time series analysis*. Princeton, N.J: Princeton University Press. 799 pp. ISBN: 978-0-691-04289-3.

Herbst, **Edward P.** and **Johannsen**, **Benjamin K.** (Mar. 2024). Bias in local projections. *Journal of Econometrics*. 2024. 240, p. 105655.

Hickman, **Bert G.** (Jan. 1955). *The Korean War and United States Economic Activity, 1950-1952.* hick55-1. National Bureau of Economic Research.

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

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

— (Feb. 2025). *Guns and Growth: The Economic Consequences of Defense Buildups*. Kiel Report 2. Kiel Institute for the World Economy.

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

Jørgensen, **Peter L.** and **Ravn**, **Søren H.** (Jan. 2022). The inflation response to government spending shocks: A fiscal price puzzle? *European Economic Review*. 2022. 141, p. 103982.

Kantor, **Shawn** and **Whalley**, **Alexander** (Sept. 1, 2025). Moonshot: Public R&D and Growth. *American Economic Review*. 2025. 115, pp. 2891–2925.

Leeper, **Eric**, **Walker**, **Todd**, and **Yang**, **Shu-Chun Susan** (2013). Fiscal Foresight and Information Flows. *Econometrica*. 2013. 81, pp. 1115–1145.

Leeper, **Eric M.**, **Walker**, **Todd B.**, and **Yang**, **Shu-Chun S.** (Nov. 2010). Government investment and fiscal stimulus. *Journal of Monetary Economics*. 2010. 57, pp. 1000–1012.

Liu, **Chang** and **Williams**, **Noah** (Aug. 2019). State-level implications of federal tax policies. *Journal of Monetary Economics*. 2019. 105, pp. 74–90.

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.

Mertens, **Karel** and **Ravn**, **Morten O.** (May 1, 2010). Measuring the Impact of Fiscal Policy in the Face of Anticipation: A Structural VAR Approach. *The Economic Journal*. 2010. 120, pp. 393–413.

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

Montiel Olea, **José Luis** and **Plagborg-Møller**, **Mikkel** (2021). Local Projection Inference Is Simpler and More Robust Than You Think. *Econometrica*. 2021. 89, pp. 1789–1823.

Montiel Olea, **José Luis**, **Plagborg-Møller**, **Mikkel**, **Qian**, **Eric**, and **Wolf**, **Christian** (2025). Local Projections or VARs? A Primer for Macroeconomists. *NBER Macroeconomics Annual*. 2025.

Mountford, **Andrew** and **Uhlig**, **Harald** (Sept. 2009). What are the effects of fiscal policy shocks? *Journal of Applied Econometrics*. 2009. 24, pp. 960–992.

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.

Piger, **Jeremy** and **Stockwell**, **Thomas** (Aug. 2025). Differences From Differencing: Should Local Projections With Observed Shocks Be Estimated in Levels or Differences? *Journal of Applied Econometrics*. 2025. 40, pp. 759–787.

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

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.

— (2013). Government Spending and Private Activity. *Fiscal Policy after the Financial Crisis, edited by Alberto Aleina and Francesco Giavazzi*. University of Chicago Press, pp. 19–62. URL: https://doi.org/10.7208/9780226018584-003.

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

— (2021). The Macroeconomic Consequences of Infrastructure Investment. *Economic Analysis and Infrastructure Investment. Chicago: University of Chicago Press*. 2021. 28215.

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.

Ricco, **Giovanni**, **Callegari**, **Giovanni**, and **Cimadomo**, **Jacopo** (Sept. 2016). Signals from the government: Policy disagreement and the transmission of fiscal shocks. *Journal of Monetary Economics*. 2016. 82, pp. 107–118.

Romer, **Christina D** and **Romer**, **David H** (June 1, 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.

Schick, **Allen** (1990). *The capacity to budget*. Washington, D.C. : Lanham, MD: Urban Institute Press ; Distributed by University Press of America. 225 pp. ISBN: 9780877664383 9780877664390.

Stock, **James H.** and **Watson**, **Mark W.** (May 1, 2018). Identification and Estimation of Dynamic Causal Effects in Macroeconomics Using External Instruments. *The Economic Journal*. 2018. 128, pp. 917–948.

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