Why Does GDP Move Before G? It's all in the Measurement*

Edoardo Briganti
University of California, San Diego

Victor Sellemi
University of California, San Diego

December 9, 2022

Abstract

The early response of GDP to narratively identified defense news shocks is not captured when using government spending shocks identified recursively from a VAR. We show that this early differential response is driven primarily by a robust and positive increase in business inventories. This also explains why fiscal multipliers are higher when estimated with defense news shocks compared to recursive shocks. More specifically, inventories increase to account for federal contractors ramping up early-stage production associated with newly awarded defense procurement contracts. However, the value of these contracts does not show up in government spending (G) until the associated payment-on-delivery, which occurs an average of 2-3 quarters later because of time-to-build. Using novel data on defense procurement obligations, we show that shocks to contract awards Granger-cause recursively identified VAR shocks but not defense news shocks. Consequently, the early response of GDP relative to G following a defense news shock is largely explained by the delayed accounting of defense procurement contracts. We also show that 40% of the early inventory effect of news shocks is attributed to production smoothing of defense industries which is unrelated to new contract awards. Finally, we use procurement obligations as an instrument for G and estimate a 4-year spending multiplier of 1.17.

Keywords: Macroeconomics, Government Spending, Public Procurement (JEL: E60, E62).

*We are particularly thankful to Valerie Ramey for her constant feedback and help. We also acknowledge helpful comments from Jim Hamilton, Johannes Wieland, Marc Muendler, Juan Herreño, Fabian Trottner, Carlos Goes, Francesco Amodeo, Ali Uppal and other UCSD macro seminar's participants. Without implicating, we also thank Elizabeth Roer, Bryan Goudie, and other industry experts for valuable comments and discussion.

1 Introduction

Forward-looking economic agents react to both changes in current government spending (G) and expectations about future government spending. Accounting for anticipation effects is crucial to identify government spending shocks. Narrative methods accomplish this goal by using exogenous instruments that directly capture changes in expectations about future government spending. Effects that are driven directly by the instrument are deemed anticipation effects while the rest is a response to changes in current spending. Recent empirical research also documents payment delays in government spending, meaning that government funds are often legally obligated to contractors long before they are paid (and recorded as spending). As a result, payment delays can mechanically generate anticipation effects of government spending.

In this work, we propose a novel inventory-based channel for anticipation effects of government spending. We provide evidence that the early response of GDP relative to G following a Ramey (2011) defense news shock is driven almost exclusively by an increase in the inventories of federal defense contractors. The early increase in inventories can be attributed primarily to an increase in newly awarded defense procurement contracts following a defense news shock. However, war-related contract awards and associated early-stage production occur several quarters before federal funds are paid, and are thus recorded in government spending with a delay. Until payments-on-delivery occur, the associated increase shows up in aggregate inventories.

We begin with an empirical decomposition of the increase in GDP following a defense news shock using variables obtained from the National Income and Product Accounts (NIPA). At the aggregate level, we find that the majority of this increase is driven by the robust increase in investment and specifically the business inventories component of investment. We show further in the panel of manufacturing industries that the increase in inventories following war events is driven by higher real inventories by defense contractors.

To document this, we construct novel quarterly time-series of defense procurement spending and defense procurement obligations. In particular, procurement spending is the component of G capturing outlays or payment to defense contractors, while obligations sums the total value of prime contract awards. We find that obligations lead payments (and thus G) by an average of 2-3 quarters. However, we are not the first to claim that the most widely used measure of G, based on NIPA data, measures government spending too late in the process. Leduc and Wilson (2013) find that a dollar of federal funds allocated to highway construction can take up to six years to be paid after it is obligated at the state level. Similarly, Brunet (2020) finds that the annual budget leads NIPA's measure of G by one year. Brunet (2020) also proposes the idea that production-in-progress associated with unpaid federal funds is recorded in inventories until it is moved to G.

However, these studies do not directly document the response of defense industries to obligated defense funds. By contrast, we use defense procurement obligations to decompose the early response of inventories to news shocks into two components. We call the first and main component *time-to-build* production, since it captures the work-in process of defense contractors in response to new defense procurement contract awards. Since defense contracts have a long dollar-weighted duration (median of more than 4 years), the early response of inventories accounts for this unpaid production-in-progress which does not yet show up in G. Secondly, defense news shocks also have an effect on inventories which cannot be explained by newly awarded contracts. Basically, contractors increase production because they foresee future contracts being awarded,namely *production smoothing*. This effect is only relevant in the quarters after the news shock, but accounts for about 40% of the early response of inventories.

Our findings have important implications regarding the identification of government spending shocks. We show that recursively identified shocks to government spending with a VAR lead to underestimates of the response of inventories relative to narratively identified defense news shocks. The differential response of inventories virtually accounts for the entire difference in estimates of the aggregate fiscal multiplier using these two identification strategies. This difference can also be explained in terms of payment delays in G. In particular, we verify that shocks to federal procurement obligations Granger-cause VAR-identified shocks to government spending. As a result, we argue that researchers should account for time delays in obligation and payment of federal funds when studying the effects of government spending.

Related Literature First of all, our work relates to the fiscal policy literature which uses war dates and news to study the effects of government spending. Among the most notable works is Ramey and Shapiro (1998), Edelberg, Eichenbaum, and J. Fisher (1999), Burnside, Eichenbaum, and J. D. Fisher (2004), Eichenbaum and J. Fisher (2005), Ramey (2011), Barro and Redlick (2011), Ben Zeev and Pappa (2017) and Ramey and Zubairy (2018). One key empirical finding in these papers is the early response of GDP relative to G following news about future defense spending. The literature considers this early response to be evidence of anticipation effects of government spending, since output reacts directly to news before any corresponding changes in G. We contribute to this literature by showing that this early response is driven by an increase in inventories, and more specifically defense industries reacting via time-to-build production and production smoothing.

Recent work also documents long implementation lags in government spending (i.e., delays in spending government funds which have already been obligated).

Most notably, Leduc and Wilson (2013) discuss payment delays in the context of highway construction, while Brunet (2020) highlights the delay between congressional budgeting and outlays of federal funds. We add to this strand of literature by providing direct measurement of a time delay between defense procurement obligations and spending of 2-3 quarters.

Finally, our work reinforces the results of Ramey (2011), who shows that government spending is anticipated, and that anticipation complicates identification of government spending shocks. In particular, we show that shocks identified with the recursive method of Blanchard and Perotti (2002) miss the initial production of defense industries at the onset of a military build-up. The underestimated effect, captured in inventories, leads to lower estimates of the multiplier.

The paper is organized as follows. Section 2 establishes the positive response of contractor inventories following a defense news shock. Section 3 studies the underlying economic and accounting mechanisms driving the response of inventories using novel procurement data. Section 4 shows that our findings pose a threat for the identification of government spending shocks. Section 5 compares our estimates of the multiplier using obligations instead of spending with other estimates from the literature. Section 6 concludes.

2 Response of Inventories to Fiscal Shocks

In this section, we decompose changes in the components of real output that are driven by news about future government spending rather than actual government spending. We find that the early response of GDP to defense news shocks is driven primarily by a positive and robust response in business inventories.[1] Our starting point is Ramey (2011), who finds that aggregate output reacts immediately to news about future war-related defense spending (defense news shocks), while government spending itself has a delayed response.[2] We replicate this result in Figure 1.

[1] Note that inventories is one component (along with residential and non-residential investment) of I in the decomposition $GDP = C + I + G + NX$.

[2] Related work also documents an early move to GDP relative to G (Ramey and Shapiro (1998), Edelberg, Eichenbaum, and J. Fisher (1999), Eichenbaum and J. Fisher (2005), Ben Zeev and Pappa (2017)).

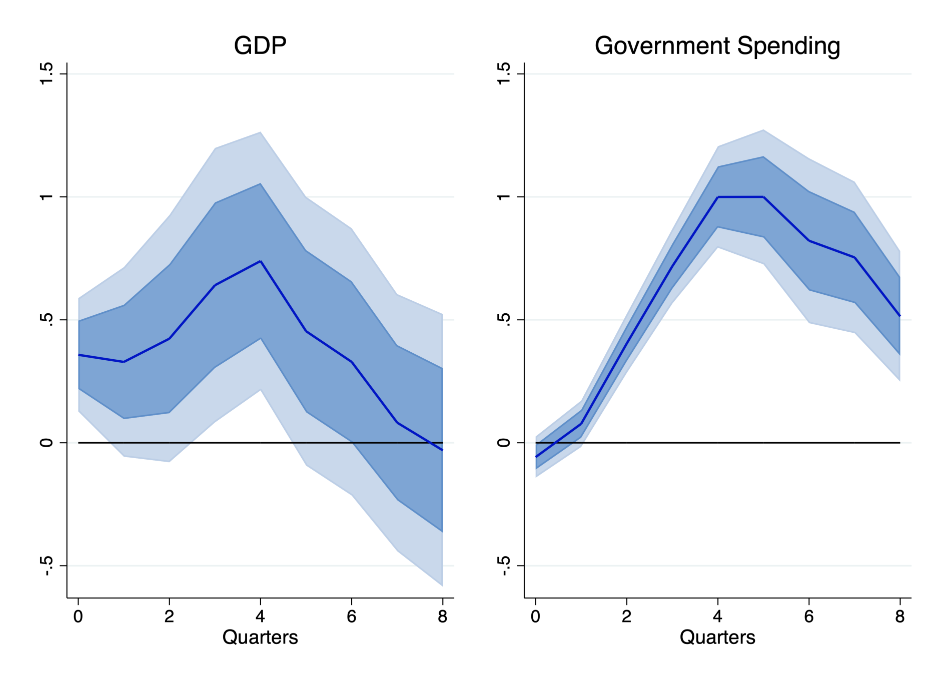


Figure 1: **Response of GDP and G to a Defense News Shock**: *IRFs of G and GDP to a defense news shock are obtained via lag-augmented local projections while controlling for lags of GDP, G, consumption, investment, hours worked and 3-month T-Bill rate. Band represent the 68% and 90% heteroskedasticity robust standard errors. Defense news shocks are obtained from the updated series in Ramey and Zubairy (2018). Sample goes from 1947Q1 to 2015Q4.*

To further investigate the underlying mechanism here, we decompose GDP and estimate the aggregate response of consumption, investment, government spending, and net-exports to defense news shocks.[3] In particular, we estimate the quarterly impulse response function (IRF) of each component using lag-augmented local projections:[4]

$$y_{t+h} = \theta_h \cdot \text{Shock}_t + \beta \cdot \mathbf{X}_t + \varepsilon_{t+h} \tag{1}$$

where y_{t+h} is the change in outcome of interest (e.g., inventories), Shock_t is the updated series of narratively identified defense news shocks from Ramey and Zubairy (2018), and $\mathbf{X}_t$ is a vector of controls for lagged shocks, government spending, consumption, inventories, hours worked by the private sector, and the three-month Treasury Bill rate. Following Ramey and Zubairy (2018), we divide all nominal variables by real potential output and the GDP price deflator.

[3]See Appendix D for detailed results.

[4]See Jordà (2005) for local projections and Montiel Olea and Plagborg-Møller (2020) for econometric details on lag-augmented local projections, which simply require heteroskedasticity robust standard errors.

We show that the early increase in GDP relative to G following a defense news shock initially shows up as an increase in inventories.[5] We report the impulse response of aggregate inventories to a defense news shock in Figure 2. Our results are robust to the exclusion of the Korean War (the largest military build-up after World War II) from the sample, indicating that the response of inventories is not driven by periods in which defense shocks dominate.[6]

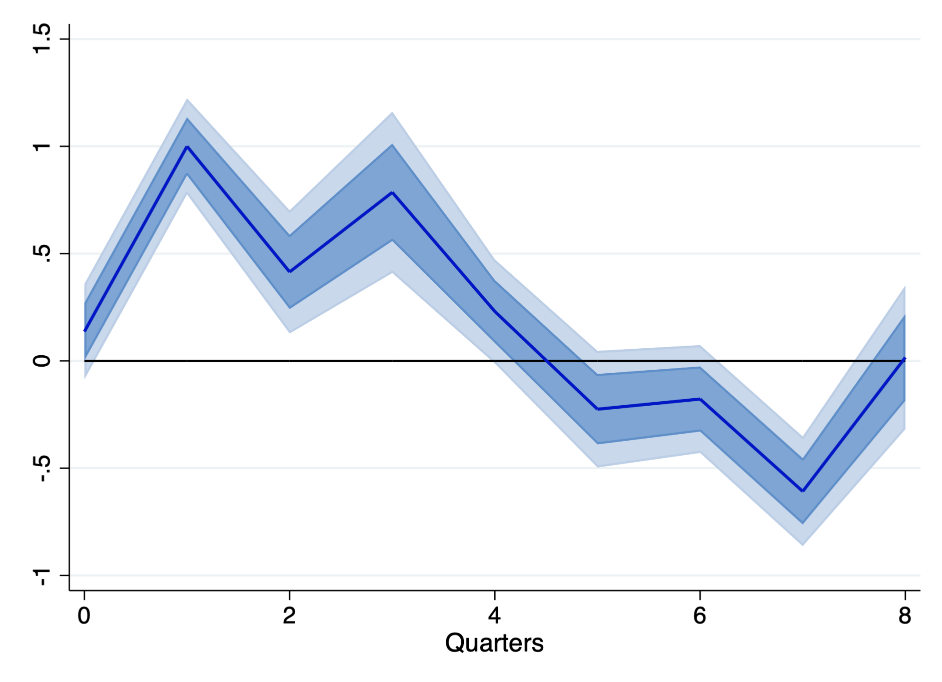


Figure 2: **Response of Inventories to a Defense News Shock**: *IRFs of inventories to a defense news shock are obtained via lag-augmented local projections while controlling for lags of GDP, G, consumption, investment, hours worked and 3-month T-Bill rate. Values are normalized by their respective peak responses. Band represent the 68% and 90% heteroskedasticity robust standard errors. Defense news shocks are obtained from the updated series in Ramey and Zubairy (2018). Sample goes from 1947Q1 to 2015Q4.*

Notice that when we use defense news shocks (left-panel), the response of inventories is positive, significant and spikes at horizon 1, that is, even before any

[5]We do not observe any significant early movement in net-export. However, in samples which include the Korean War, we also observe a positive response in durable consumption at horizon 0. Unlike the response of inventories, the response of durables is not robust to the exclusion of the war, and is likely driven by a fear of long-term rationing and price control which was uniquely prevalent at the onset of the Korean War. See e.g., Ramey (2016) and Binder and Brunet (2021).

[6]We believe that it is important to include the largest war events in the sample, since they mimic natural experiments involving government spending. However, we are aware of potential confounding factors pointed out by the literature (e.g., Perotti (2007), J. D. Fisher and Peters (2010), Perotti (2014) and Ramey (2016)). We thus also report results without the Korean war in Appendix E.

movement in G (recall Figure 1). [7] The response of inventories is robust to the use of recursive shocks.

Next, we show in the panel of manufacturing industries that the aggregate response of inventories is driven by an increase in industries which heavily contract to the federal government.

2.1 Industry Analysis

Given the positive and robust aggregate response of inventories, we study heterogeneity in this response across industries in response to war events. First of all, we use monthly data from the Bureau of Economic Analysis (BEA) to construct a panel of real inventories for 18 manufacturing industries between January 1959 and December 1997.[8]

Secondly, we construct a weight θ_i for each industry which captures the long-run average share of industry sales coming from government purchases. We follow Nekarda and Ramey (2011) and also include downstream linkages to control for indirect sales to the government via input-output connections. For instance, Other Transportation Equipment Manufacturing has direct sales to the government which accounts for 34% of its total sales. However, after including first order input-output connections, the dependence on government purchases jumps up to 42% and 44% when including first and second order connections respectively.[9]

Moreover, manufacturing industries heavily specialize in defense production (see e.g., Ramey and Shapiro (1998), Nekarda and Ramey (2011) and Cox et al. (2021)). For instance, the strong dependence of Other Transportation Equipment Manufacturing on government purchases is driven by its defense sub-sectors, such as Aircraft, Ship Building, Guided Missiles and Space Vehicles. Therefore, we refer to these industries as "defense industries".

Finally, we estimate the following equation:

$$\mathrm{Invt}_{i,t+h} = \lambda_{ih} + \alpha_h \cdot \mathrm{War}_t + \beta_h \cdot \mathrm{War}_t \cdot \theta_i + \sum_{p=1}^{P} \varphi_{ph} \cdot \mathrm{Invt}_{i,t-p} + \varepsilon_{i,t+h} \qquad (2)$$

[7] We highlight that recursively identified shocks also capture a positive but delayed response of inventories. We report this result in Appendix F. Fatas and Mihov (2001) estimate the effect of recursive shocks on a multitude of variables and also find a positive response of inventories, even if they do not discuss this result in their paper. To the best of our knowledge, they are the only one who looked at inventories.

[8] Many thanks to Valerie Ramey for providing this data. Our data stops in 1997, however, most of the variation in defense spending comes from periods before the Nineties (Vietnam War and Soviet invasion of Afghanistan).

[9] See Appendix B.1 for detailed derivation of industry weights.

where $h = 0, 1, ..., H$, Invt_{it} is total real inventories of industry i in month t, λ_{ih} is an industry fixed-effect, and War_t is war dates.[10] Consistent with Ramey and Shapiro (1998) and Eichenbaum and J. Fisher (2005), our war event variable is a weighted dummy with value 1 on March 1965 and 0.3 on January 1980 to indicate the start of the Vietnam War and Soviet invasion of Afghanistan, respectively.

We are interested in the estimands α_h and $\alpha_h + \beta_h$. The former represents is the response of inventories for those industries not connected to the government, i.e. $\theta_i = 0$. The latter is the response of industries highly connected to the government through government purchases i.e. $\theta_i = 1$.

If war dates has a differential positive effect on sectoral inventories, proportional to the degree of connection to the government, we expect $\beta_h > 0$.

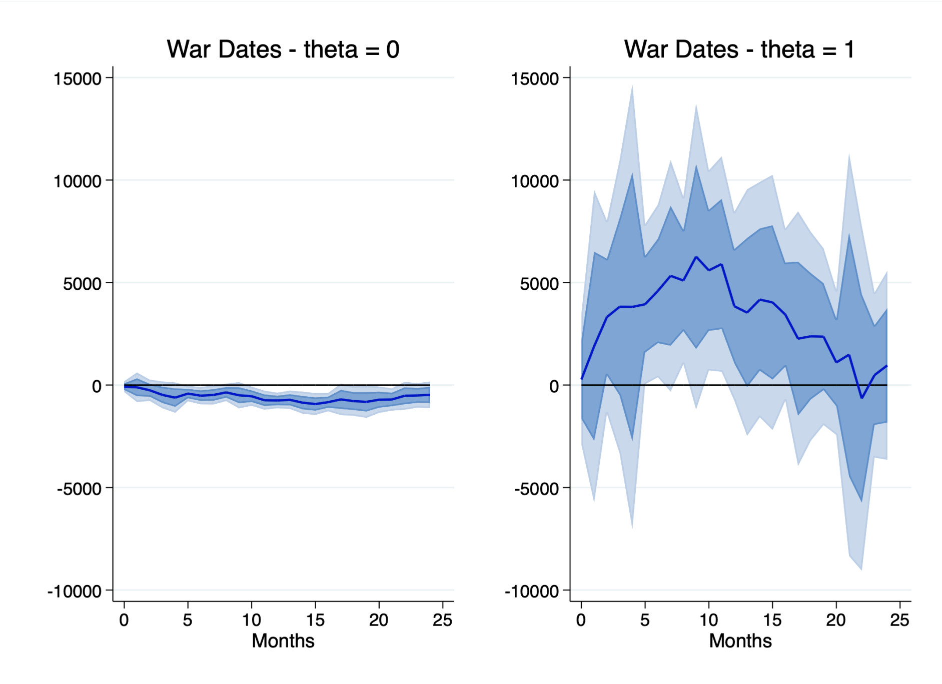


Figure 3: **Response of Sectoral Inventories to War Events.** *Left panel shows estimates of α_h (response when $\theta_i = 0$), right panel reports estimates of $\alpha_h + \beta_h$ (response when $\theta_i = 1$). Weights are normalized by maximum weight (i.e. the one of Other Transportation Equipment Manufacturing). Since Real Inventories are trending, data is filtered using Hamilton (2018)'s filter (we set $h = 24$ and $p = 12$, that is two years lag plus one more year of lags). The unit of real inventories is millions of 2005 chained dollars. Sample goes from 1959-Jan to 1997-Dec and uses 18 sectors breakdown of Manufacturing. Confidence bands are 68% and 90%. Standard errors are obtained via Bootstrap.*

Figure 3 shows a significant positive and long-term differential response ($\alpha_h + \beta_h$) of federal contractor inventories to war dates. On the other hand, the change in

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

inventories for those industries who do not supply the government (α_h) is negative and close to zero. Therefore, all of the effect of war dates on inventories is explained by the degree of connection of each sector to the government.

Robustness We verify that this differential response of contractor inventories is not driven by different sensitivities of defense industries to the business-cycle. In particular, we replace War_t with monetary policy shocks constructed narratively by C. D. Romer and D. H. Romer (2004) and updated by Wieland and Yang (2020) and estimate the differential response (β_h) to be statistically indistinguishable from zero. This confirms that the reaction of federal contractors to defense news shocks is driven by war-related forces and not the associated business-cycle fluctuations.

Furthermore, we make sure that the differential response of defense industries during a military build-up is not driven by spurious correlation. In particular, we re-estimate Equation (2) using randomly re-shuffled weights as commonly done in the production network literature (e.g. see Ozdagli and Weber (2020)). In this case we no longer detect a significant response of defense industries to war events. We report the results of these robustness checks in Appendix B.2.

Finally, one might argue that inventories are not a direct measure of production. Therefore, we proxy defense production using average weekly hours of production and non-supervisory workers in the Aircraft industries. We find a positive and strong effect on this variable in response to both defense news and defense prime contract awards. This result is a direct evidence of defense industries quickly ramping-up production in response to military build-ups.[11]

3 What Drives the Increase in Inventories?

In this section, we show that the increase in federal contractor inventories following a defense news shock is primarily driven by an increase in early-stage production associated with defense procurement contracts. However, the value of this production does not show up in government spending until 2-3 quarters later, when the final goods are delivered and paid for by government agencies. This mechanically generates the early response of inventories relative to G following a defense news shock. To our knowledge, our work is the first to directly document this time delay between defense obligations and payments, and decompose its role in generating an early response of inventories.

In principle, contractor inventories might increase in response to changes in ex-

[11] We explain why we use average hours in the Aircraft industry to measure defense production in Appendix C.2. In the same part of the appendix we also show the estimated IRFs of hours worked in the Aircraft industry in response to news and contracts.

pectations about future contract awards and not just actual contract awards.[12] We refer to changes in inventories coming from new contract awards as *time-to-build* production, since firms are building up inventories as part of a long production process. On the other hand, we refer to inventory changes that cannot be explained by contract awards as production smoothing.[13] Using novel data on procurement contracts, we measure the relative effects of these two channels in the positive response of inventories following a defense news shock. We find that almost all of the impact effect is explained by early stage time-to-build production, although production smoothing accounts for about 40% of the increase in later horizons.[14]

3.1 Data

We construct a novel database of defense procurement spending and obligations. Spending measures payments (or outlays) from federal agencies to contractors, while obligations measure the total value of federal funds as soon as they are contractually obligated to firms. To construct the spending series, we use the accounting identity discussed in Cox et al. (2021):

$$\begin{aligned}(\text{Procurement Spending})_t \approx\ & (\text{Intermediate Goods \& Services Purchased})_t + \\ & + (\text{Change in Government Fixed Assets})_t + \\ & - (\text{Investment R\&D})_t \\ \approx\ & (\text{Payment to Contractors})_t,\end{aligned}$$

where all variables are obtained from the National Income and Product Accounts (NIPA). Figure 4 plots this measure of defense procurement spending along with the annual measure of procurement spending of Dupor and Guerrero (2017), aggregated over states. The two measures are virtually identical before 1984, but afterwards the Dupor and Guerrero (2017) series omits contract actions with value less than $25,000 and thus systematically underestimates our NIPA-based series. From 2000 onwards, we also aggregate federal agency payments from the universe of procurement contracts, available in the Federal Procurement Data System (FPDS), and find that our measure is consistent.

[12]See e.g., Chapter 17 of Jones (2008) for details on the economic mechanisms underlying changes in inventories.

[13]Firms might start production early in anticipation of future contract awards either to smooth convex adjustment costs or lower delivery times. However, we do not take a stance on the underlying behavioral motivation here.

[14]We assume that any changes in inventories reflect actual production. In Chapter 7 of NIPA's 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.*"

To construct the obligations series, we aggregate the value of procurement contracts awarded by the DoD from the universe of procurement contracts recorded in the Federal Procurement Data System (FPDS). Since this data is only available from 2000 onwards, we also collect historical information from the periodical *Business Conditions Digest* (henceforth BCD) which is available from January 1951 to November 1988.[15] We use information from the contract and spending data to impute missing quarters.[16]

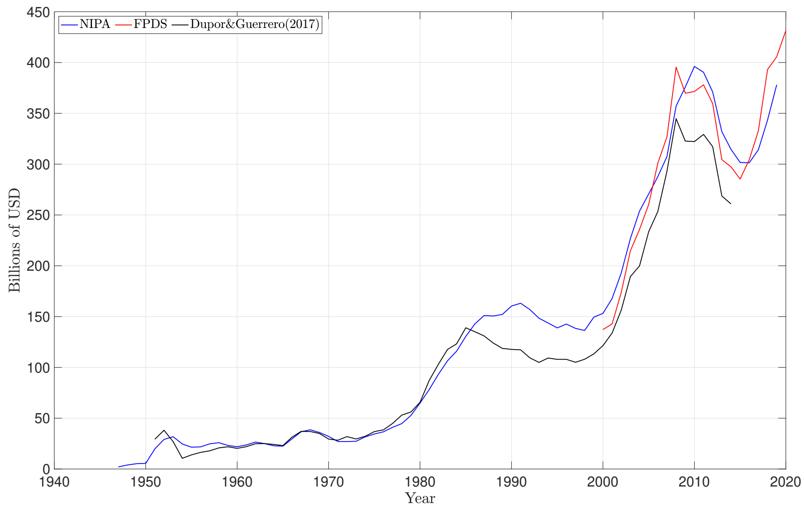


Figure 4: **Annual Defense Procurement Spending.** *Data from FPDS and our newly constructed defense procurement spending variable from NIPA data are aggregated at yearly frequency to match Dupor and Guerrero (2017)'s variable.*

3.2 Granular Evidence of Time-to-Build

Obligations and spending capture different moments in the defense procurement process. To better understand the timing, we summarize the defense procurement process in Figure 5. The timeline begins when a contract is awarded (i.e., when the federal government becomes contractually *obligated* to pay for a pre-specified set of goods/services). Although contractors become aware of contract opportunities before the award date in the form of pre-award solicitations, we find that these solicitations largely occur in the same quarter as the award date.[17]

[15]Many thanks to Valerie Ramey for providing this data. See Appendix A.2 for details on the sources of contract level data.

[16]See Appendix A.4 for details on the construction of the series.

[17]See Appendix A.5 for more details on the timing of pre-award notices for a detailed description of these notices.

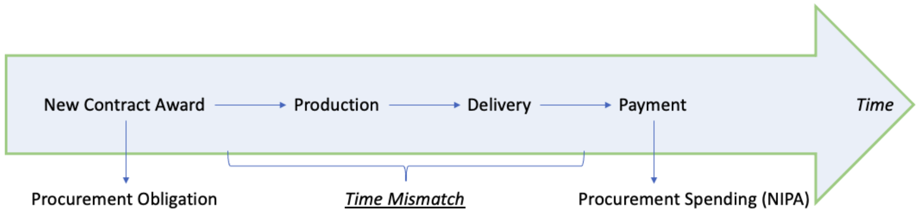


Figure 5: **Timeline of Procurement Contracts.** *The procurement timeline follows information from the Federal Acquisition Regulation (FAR) and the BEA's Concepts and Methods of NIPA.*

After contracts are awarded, contractors launch a potentially long production process. In particular, contract-level data indicates that the mean and median duration of a $1 defense procurement contract are 4.2 and 5.4 years, respectively. We measure duration as the period of performance, or the number of days between award date and contract end (full delivery) date. We find that total defense procurement spending is dominated by few very large contracts with very long duration. Using the same data, Cox et al. (2021) report a very short average contract duration. However, their estimated duration is not weighted by contract size and is thus consistent with out findings.[18]

Given that production takes a long time, when do associated payments actually occur? According to the Federal Acquisition Regulation (FAR), the canonical rule (FAR 32.904) for payments to federal contractors from government agencies is *payments-after-delivery*. Certain contracts are also subject to partial-delivery-payments. However, given the multiple year average duration of contracts, we still observe several quarter-long delays in partial deliveries.[19]

The left panel of Figure 6 plots spending and obligations from Jan 1951-Nov 1988, and the right panel reports the lead-lag correlation.[20] We find that obligations lead spending by 2-3 quarters, with the average lead-lag correlation significantly peaking in the North-East quadrant of the map. This suggests that changes in obligations are more highly correlated with delayed changes in spending rather than current changes in spending. The results replicate for more recent obligations data obtained from FPDS. We report the full set of results and robustness checks in Appendix A.2.

[18]See Table 4 in Appendix A.3 for more details.

[19]See Appendix A.3 for more details on partial delivery payments.

[20]Lead-lag correlations are useful for studying relationships in time between variables. One recent example is Smets, Tielens, and Van Hove (2019), who study the propagation of inflation from upstream to downstream sectors.

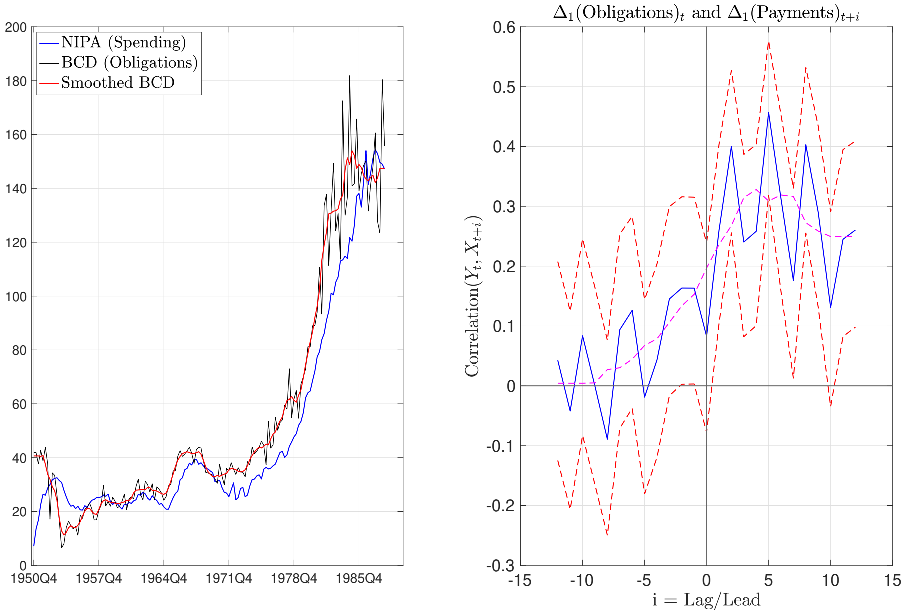


Figure 6: **Federal Defense Procurement Obligations and Spending**. *Source: BCD for Obligations and NIPA for Spending. Sample: January 1951 to November 1988. The lead-lag correlation-map plots the following: Corr* $(\Delta_1(Obligations)_t, \Delta_1(Payments)_{t+i})$*, where* Δ_1 *is the first difference operator and payments are proxied with our measure of defense procurement spending from the NIPA data.*

The payment (government outlay) thus occurs several quarters after the contract award. This finding is consistent with the results of Leduc and Wilson (2013) and Brunet (2020) and confirmed exactly by the Department of Commerce's Government Transaction Methodology Paper:[21]

"*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.*"

In other words, early stage production associated with long procurement contracts is recorded at an aggregate level in inventories until the delayed payment-on-delivery. The value of completed and paid contract work is then moved from inventories to G. We can observe the delay between defense contract awards and payment directly from our data.

[21] Many thanks to Gillian Brunet for redirecting us to that document.

3.3 Time-to-build or Production Smoothing?

We decompose the early response of inventories to a defense news shock into time-to-build and production smoothing. We already have contract level evidence of a long time-to-build, but at the onset of a military build-up, contractors should also presumably change their expectations about future government demand. Even if contractors lack resources to forecast government demand, federal agencies are required by the FAR to provide procurement forecasts each quarter.[22] If contractors anticipate winning future contracts, they might decide to increase production today to smooth convex adjustment costs or reduce future delivery times. We do not take a stance on the exact mechanism here. We consider a recent example of this type of behavior.

Example 3.1 (Lockheed Martin in 2022)**.** In the context of an ongoing military conflict between Russia and Ukraine, new military tests in North Korea, and escalating tension in relationship between China and Taiwan, US-based contractors have modified expectations about future defense spending. In particular, the largest American defense contractor, Lockheed Martin, decided in October 2022 to increase production of HIMARS (High Mobility Artillery Rocket System). When asked about this decision, CEO Jim Taiclet responded as follows:[23]

"*The company has met with its long lead supply chain and spent $65 million — which will eventually be paid back by the US government — to fund parts in advance, shortening the time needed to manufacture the rocket system. That was without a contract or any other memo or whatnot back from the government. We just went ahead and did that because we expected it to happen. So those parts are already being manufactured now*".

In order to measure production smoothing, we first provide a formal definition.

Definition 3.1 (Production Smoothing of Defense Industries)**.** We define production smoothing $\Delta(h)$ as the effect of a defense news shock on inventories, orthogonal to shocks to newly awarded contracts (i.e., defense procurement obligations). In particular, production smoothing is the impulse response of inventories to a defense news shock conditional on zero shocks to defense procurement obligations (i.e. orthogonalized IRF):

$$\Delta(h) = \mathbb{E}_t[\text{Invt}_{t+h}|Z_t = 1, \varepsilon_t^O = 0] - \mathbb{E}_t[\text{Invt}_{t+h}|Z_t = 0, \varepsilon_t^O = 0], \tag{3}$$

where Invt_t is changes in aggregate inventories as from the NIPA, Z_t is a defense news shock, and ε_t^O is a shock to defense procurement obligations.

[22]See e.g., Agency Recurring Procurement Forecasts.

[23]Find the associated article on *Breaking Defense*. here.

We estimate production smoothing using the following tri-variate VAR using quarterly data from 1948Q1 to 2015Q4:

$$\begin{bmatrix} 1 & 0 & 0 \\ -\alpha & 1 & 0 \\ -\beta_{\text{News}} & -\beta_{\text{Oblg}} & 1 \end{bmatrix} \cdot \underbrace{\begin{bmatrix} Z_t \\ O_t \\ \text{Invt}_t \end{bmatrix}}_{\boldsymbol{X}_t^3} = \underset{3\times 3}{\boldsymbol{B}_3(L)} \cdot \boldsymbol{X}_t^3 + \boldsymbol{\varepsilon}_{3,t}$$

where $\mathbf{B}_3(L)$ is a polynomial in the lag operator. The parameter α captures the contemporaneous effect of a defense news shock on obligations, while β_{News} and β_{Oblg} capture the contemporaneous effect of shocks to news and obligations on inventories.

By including our aggregate series for defense procurement obligations O_t, we are able to calculate the effect of defense news shocks on inventories which is independent of the effect of shocks to newly awarded contracts. Figure 7 shows the impulse response function to a defense news shock estimated using the above tri-variate VAR as well as the total response of inventories estimated in a bi-variate VAR without obligations.

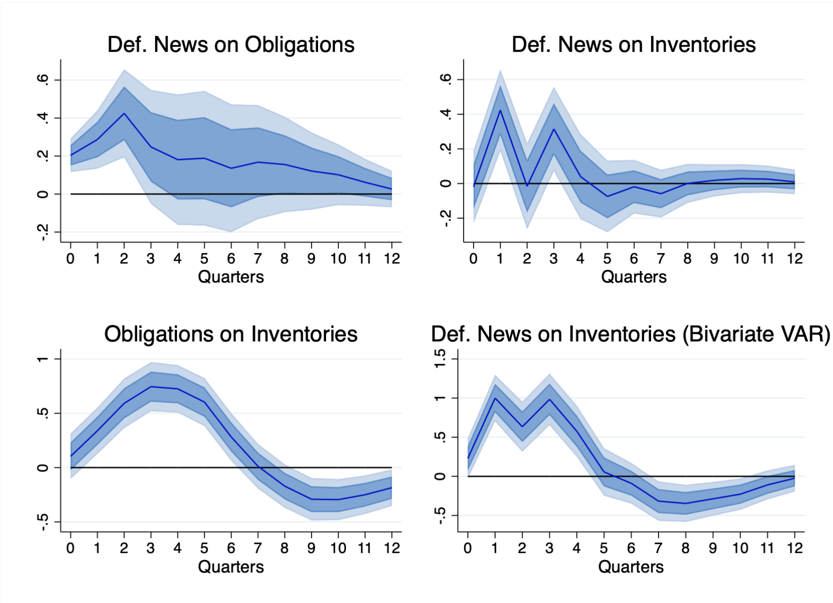


Figure 7: **(Orthogonalized) IRFs from Tri/Bi-Variate VARs.** *Variables are divided by potential GDP and include a linear time trend. Sample goes from 1948Q1 to 2015Q4. Confidence Bands are 68% and 90%. Values are rescaled by the peak response of Inventories to a defense news shock from the bivariate VAR which excludes defense procurement obligations.*

The top-left panel of Figure 7 shows the positive response of defense procurement obligations to a defense news shock. This indicates that new contracts start being awarded as soon as a defense news shock occurs. This confounds the effects of news (i.e., anticipation) with the effects of newly awarded contracts which show up in G with delay. In the bottom-left panel of the figure, we show the effect of shock to obligations ε_t^O, on inventories, orthogonal to defense news. The effect is positive and significant. Additionally, the top-right panel reports production smoothing, or the response of inventories to a defense news shock which is orthogonal to newly awarded contracts. The positive and significant estimates of $\Delta_(h)$ at horizons 1 and 3 suggest that production smoothing plays a role in the response of inventories. The bottom-right panel shows the IRF of inventories to a defense news shock without including obligations in the VAR, i.e. bivariate VAR.

For interpretability, we rescale the IRFs by the peak response of inventories to a defense news shock in the bivariate VAR occurring at horizon 1. Since the horizon 1 response of inventories to a defense news shock in the tri-variate VAR is slightly more than 0.4 it means that roughly 40% of the response of inventories at horizon 1 comes from production smoothing, while the residual part (gap between bottom-right and top-right responses) originates from the effects of newly awarded contracts, i.e. time-to-build production. Intuitively, this can be seen by shrinking the tri-variate VAR into a bivariate one by plugging obligations into the equation of inventories:

$$\begin{bmatrix} 1 & 0 \\ -\left(\beta_{\text{News}} + \alpha \cdot \beta_{\text{Oblg}}\right) & 1 \end{bmatrix} \cdot \underbrace{\begin{bmatrix} Z_t \\ \text{Invt}_t \end{bmatrix}}_{\boldsymbol{X}_t^2} = \underset{2\times 2}{\boldsymbol{B}_2(L)} \cdot \boldsymbol{X}_t^2 + \boldsymbol{\varepsilon}_{2,t}$$

Notice that the impact effect of a defense news shock on inventories is the combination of production smoothing (β_{News}) and the effect of a shock to new contracts on inventories triggered by the news ($\alpha \cdot \beta_{\text{Oblg}}$). Basically, without controlling for new contracts, defense news shock capture both production smoothing and the time-to-build, while augmenting the VAR with new contracts allows us to tell-apart the two effects.

4 Threats to Identification

In this section, we argue that government spending shocks identified using standard VAR methods do not capture early-stage production associated with newly awarded federal procurement contracts during a military build-up. In most macroeconomic studies, researchers are interested in the economic effects of government spending from the moment funds are contractually obligated and contractors begin reacting. In this setting, the actual transfer of cash is not the main focus. Given our results from the previous section, we expect that VAR-identified government spending shocks are capturing transfers of cash rather than obligation of funds.

We provide an illustrative example of this problem around the outbreak of the Korean War.

Example 4.1 (Korean War)**.** Figure 8 plots quarterly defense procurement obligations, shocks to government spending, inventories, and some other key variables between 1950Q1 and 1953Q3. In the summer of 1950 (Q3), we observe a large defense news shock associated with the outbreak of the Korean War (bottom-left panel). However, the VAR-identified shock to NIPA's measure of G does not spike until 2-3 quarters later (bottom-middle panel). Similarly, G has a slow positive response (top-left panel) following the defense news.

On the other hand, defense procurement obligations, constructed using our obligation data, react almost immediately to the shock. In other words, the DoD begins awarding defense procurement contracts at the onset of the war event (top-middle panel). As expected, we also observe quick increases in inventories starting 1950Q4 (top-right panel) and in defense production (bottom-right panel). We proxy defense production by average hours of production workers in the aircraft industry.[24]

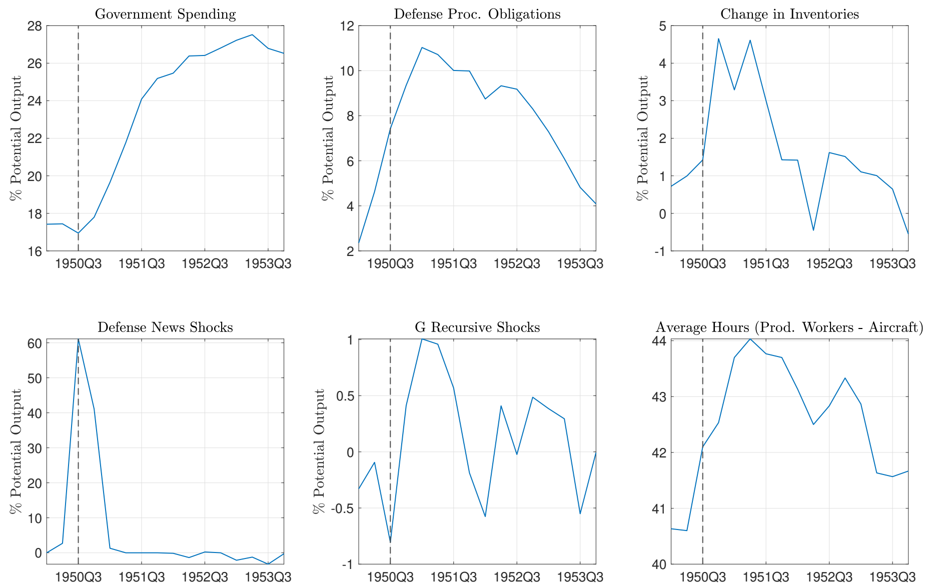


Figure 8: **Illustrative Example: Response to the Korean War**

Building on this intuition, we show using the entire sample that VAR-identified

[24]In Appendix C, we provide more details as to why we believe this variable is a good proxy for defense production.

shocks to defense procurement obligations Granger-cause VAR-identified shocks to government spending.[25] Moreover, recursive shocks to obligations cannot be predicted by Ramey (2011)'s defense news variable. This result is robust to the exclusion of the Korean War in the sample. Table 1 summarizes the results of these Granger causality tests.

Predicted	*Predictor*	*F*	*Pvalue*	*Korea*
Recursive Shocks	Obligation Shocks	15.62	0.0%	Yes
Recursive Shocks	Obligation Shocks	5.13	0.0%	No
Obligation Shocks	Recursive Shocks	2.17	3.0%	Yes
Obligation Shocks	Recursive Shocks	1.66	10.9%	No
Defense News Shocks	Obligation Shocks	0.42	91.0%	No
Defense News Shocks	Obligation Shocks	0.33	95.3%	Yes

Table 1: **Predictability of Recursive Shocks via Obligations**: *Granger Causality test is a Wald test on the 8 lags of the predictor while controlling for 4 lags of the predicted variable.*

In the top panel, we show that shocks to defense procurement obligations predict recursively identified government spending shocks. On the other hand, the second panel shows a much weaker relationship in the other direction, especially when you omit the Korean War from the sample. Our results are consistent with Ramey (2011), who shows that war dates Granger-cause recursive shocks to G.

However, obligations and defense news shocks are not significantly predictive of each other. In particular, the bottom panel shows that defense news shocks do not predict shocks to obligations. This indicates that early economic effects of newly awarded contracts, which are missed by recursive shocks to G, can be captured using both defense news shocks or shocks to obligations. Next, we show that the difference in fiscal multipliers estimated using the two methods is almost entirely explained by a difference in capturing the early response of inventories.

Following Ramey (2016), we estimate cumulative fiscal multipliers using LP-IV with both VAR-identified shocks to G and narratively identified defense news shocks. We use the following estimation equation:[26]

[25] We obtain shocks to obligations by ordering defense procurement obligations first in the VAR.

[26] More technical details on LP-IV are available in Stock and Watson (2018).

$$\sum_{h=0}^{H} y_{t+h} = \gamma_H + \mathcal{M}(H) \cdot \underbrace{\sum_{h=0}^{H} g_{t+h}}_{\text{instrument with Shock}_t} + \text{lags}_t + \varepsilon_{t+h}, \tag{4}$$

where $\mathcal{M}(H)$ is the cumulative government spending multiplier at horizon H, y_t is GDP at time t, g_t is government spending at time t, Shock_t is an exogenous instrument for cumulative government spending, and lags_t contains lagged values of the shock, government spending, consumption, investment, hours worked and 3 months T-Bill rate. We rescale nominal variables by potential output. The narrative method sets Shock_t equal to the defense news shock variable, while the recursive method is equivalent to setting Shock_t equal to G.

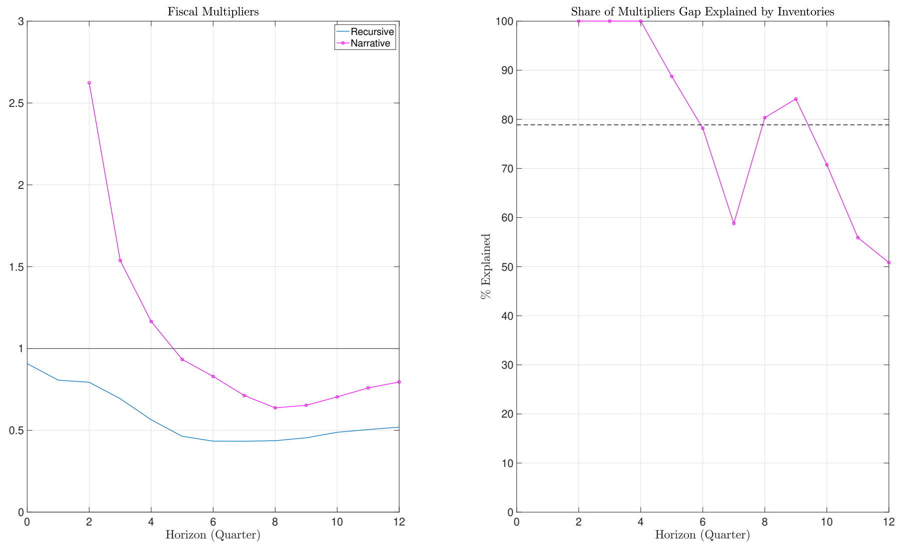


Figure 9: **Cumulative Fiscal Multipliers and Multiplier-Gap**. *Sample goes from 1947Q1 to 2015Q4. Left-panels: cumulative fiscal multipliers (point estimates). Right Panels: share of multipliers-gap explained by the differential response of inventories (dashed black line is the average of the response). Share is capped at 100% and is calculated only when the multiplier gap is finite and positive.*

The top-left panel of Figure 9 shows that the recursive method delivers uniformly lower point estimates of the multiplier relative to the narrative method. To investigate how much of the difference in multipliers (*multiplier gap*) can be explained by a differential response in inventories, we propose the following measure:

$$\frac{\hat{\mathcal{M}}(H)^{Narrative}_{Inventory} - \hat{\mathcal{M}}(H)^{Recursive}_{Inventory}}{\hat{\mathcal{M}}(H)^{Narrative}_{GDP} - \hat{\mathcal{M}}(H)^{Recursive}_{GDP}}$$

which computes the proportion of the multiplier gap that is explained by differences in the inventory multiplier. More specifically, the inventory multiplier is estimated by replacing y_t with inventories at time t in equation (4). Figure 9 shows that on average, 80% of the multiplier gap can be explained by the differential response of inventories.

5 Comparison of Multipliers with Brunet (2020)

Finally, we estimate multipliers using our new quarterly measure of defense procurement obligations. We compare our results to the recent estimates from Brunet (2020), who also uses a measure of government spending which is better aligned with the timing of obligated funds. In particular, Brunet (2020) estimates multipliers by regressing cumulative changes in GDP on cumulative changes of *Budget Authority*, which tracks defense spending when it is authorized, before funds are dispersed from the Treasury:

$$\sum_{k=0}^{H} \frac{GDP_{t+h} - GDP_{t-1}}{GDP_{t-1}} = \mathcal{M}(H) \cdot \sum_{k=0}^{H} \frac{BA_{t+h} - BA_{t-1}}{GDP_{t-1}} + lags + \varepsilon_{t+h}$$

where BA_t is Budget Authority in year t. We report these estimates in Table 2 for two post WWII samples which either include or do not include the Korean War.

Sample	Horizon (Years)	0	1	2	3	4
Post WWII Sample	Budget Authority	1.76 (4.08)	1.51 (2.73)	1.30 (1.63)	1.28 (1.29)	
	Def. Proc. Oblig.	4.84 (1.71)	3.96 (3.01)	1.61 (3.94)	1.27 (3.18)	1.17 (2.61)
Post Korean War	Budget Authority	1.83 (4.54)	1.84 (3.56)	1.72 (2.25)	1.67 (1.66)	
	Def. Proc. Oblig.	1.44 (2.85)	1.7 (2.46)	1.15 (1.99)	1.09 (2.04)	1.14 (2.09)

Table 2: **Multiplier Comparison:** t-statistics reported in parentheses below the multipliers' point estimates. Budget Authority samples go from 1948 to 2016 and from 1955 to 2016 (annual frequency). Samples using defense procurement obligations go from 1948Q1 to 2015Q4 and from 1954Q1 to 2015Q4 (quarterly frequency).

On the other hand, we estimate multipliers as suggested in Section 4, instrumenting the cumulative change in the NIPA-measured government spending, G, with defense procurement obligations (one step LP-IV).[27] We plot our estimates of cumulative fiscal multipliers in Figure 10 for the two sample periods from Brunet (2020).

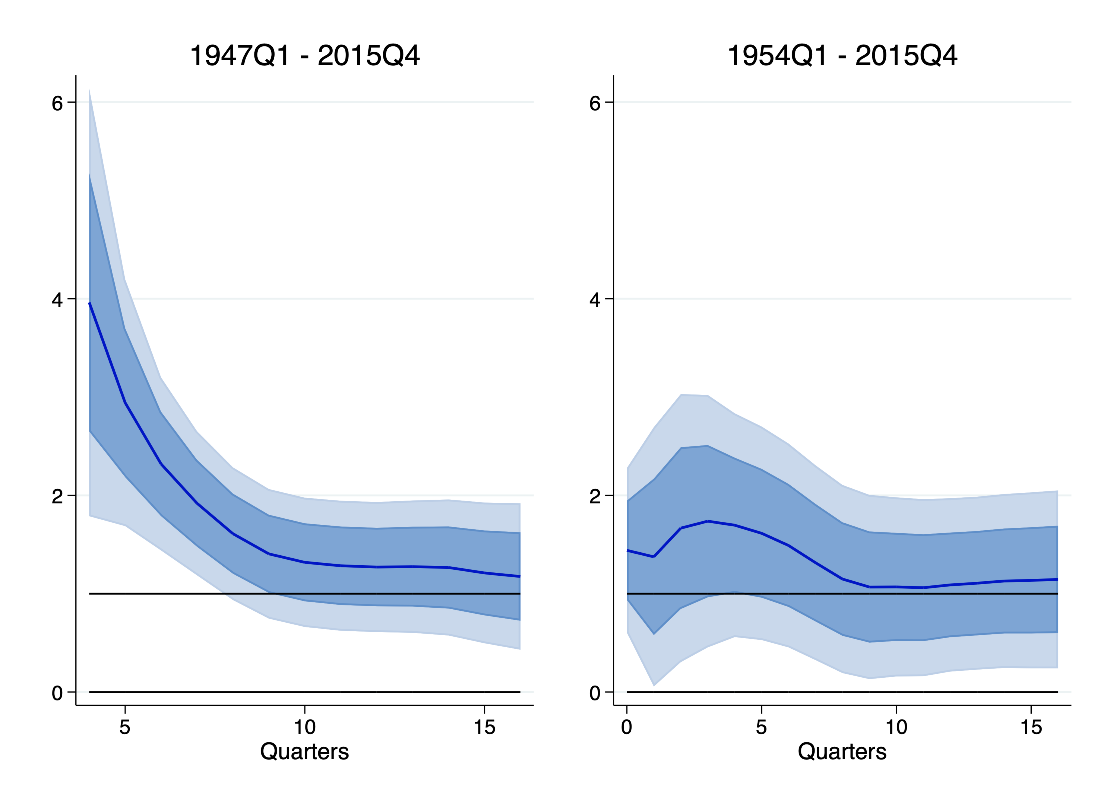


Figure 10: **Cumulative Fiscal Multipliers via Shocks to Defense Procurement Obligations.** *Standard errors are two-stage-least-squares robust standard errors and bands are the 68% and 90% confidence levels. Dark horizontal lines referring to the values of zero and one. Data are quarterly.*

Notice how in both sample periods, the multipliers are higher at short horizons and smaller at longer horizons. This is a consequence of anticipation effects: GDP increases even before G moves.[28] We also report the point estimates of the multipliers in Table 2 for different years. This is done to facilitate the comparison with the results of Brunet (2020).

In the top panel of the table, it is clear that analysis on both Post WWII samples deliver similar results, particularly at the three-year horizon. On the

[27]Recall that this is equivalent in population to ordering defense procurement obligations first in a VAR.

[28]When the Korean war is in the sample, anticipation is so strong that the multiplier is infinite until horizon 4.

contrary, results slightly differ from each other when the Korean war is excluded from the sample. In fact, our point estimates are smaller than Brunet (2020), even if they both remain above one. Finally, our multipliers tend to be more statistically significant at longer horizons, which is likely due to the fact that our analysis is carried out at quarterly frequency rather than an annual frequency.

Despite minor discrepancies between our results and those of Brunet (2020), our obligations-based method also delivers point estimates for the multiplier which are greater than one. Nevertheless, we note three important differences in our methodology, without taking a stand on the relative effects of each. First of all, defense procurement obligations is a quarterly variable which captures the whole universe of newly awarded defense procurement contracts, while Budget Authority is an annual variable which captures authorizations for defense-spending and is broader than procurement spending. Secondly, our LP-IV multiplier is interpretable as the ratio of the IRFs of GDP and G following a shock to defense procurement obligations, and is therefore a spending multiplier. On the contrary, Brunet (2020) regresses cumulative changes of GDP on cumulative changes of Budget Authority. Since Budget Authority does not map directly to NIPA defense spending (i.e., changes in Budget Authority are not necessarily changes in NIPA G), their estimates cannot be directly interpreted as a spending multiplier. Thirdly, contemporaneous changes in non-defense spending are not captured in the Budget Authority measure.

6 Conclusion

In this work, we show that the early response of GDP relative to G following a defense news shock can be explained by a robust increase in inventories. More specifically, we find in the panel of manufacturing industries that this increase is driven by defense industries building up inventories in response to war events. Although federal contractors increase inventories for few reasons, the main reason is increased production associated with newly awarded defense procurement contracts.

Using novel data at the procurement contract level, we provide evidence of a long time-to-build associated with defense contracts. Given the payments-on-delivery mandate of the Federal Acquisition Regulation, this generates a delay between contract awards (obligations) and payments (spending). We are the first to directly document a delay of 2-3 quarter between defense procurement obligations and payments. Since NIPA measures government spending in terms of payments, this delay provides a simple accounting explanation for GDP's immediate reaction to defense news shocks relative to G.

In principle, part of the response of inventories might not be attributed to newly awarded defense procurement contracts. We refer to these effects as production smoothing. By controlling for contract awards, we can measure the relative effects of this channel compared to the time-to-build channel associated with actual contract awards. We find that time-to-build production dominates in the quarter of the shock, but production smoothing makes up about 40% of the total response of inventories at larger horizons.

These findings have important implications on the identification of government spending shocks. We show that recursive identification of government spending shocks using standard VAR methods leads to underestimates in the estimated response of inventories relative to narratively identified defense news shocks. The differential response in inventories virtually accounts for the entire difference in estimates of the aggregate fiscal multiplier using these two identification strategies. This difference can also be explained in terms of payment delays in G. In particular, we verify that shocks to federal procurement obligations Granger-cause VAR-identified shocks to government spending.

Finally, we calculate government spending multipliers using defense procurement obligations. We compare our estimates with the results from Brunet (2020), who also uses a measure of federal funds at the time of obligation. Our results are similar and both deliver estimates of the multiplier greater than one. In particular, we estimate a 4-year multiplier of 1.17.

More details on the theoretical macroeconomic implications of our results are left to future work.

Bibliography

[AGM20] Alan J Auerbach, Yuriy Gorodnichenko, and Daniel Murphy. "Local Fiscal Multipliers and Fiscal Spillovers in the USA". en. In: *IMF Economic Review* 68.1 (Mar. 2020), pp. 195–229. ISSN: 2041-4161, 2041-417X. DOI: 10.1057/s41308-019-00102-3. URL: http://link.springer.com/10.1057/s41308-019-00102-3 (visited on 09/23/2020).

[BB21] Carola Conces Binder and Gillian Brunet. "Inflation expectations and consumption: Evidence from 1951". en. In: *Economic Inquiry* (Nov. 2021), ecin.13051. ISSN: 0095-2583, 1465-7295. DOI: 10.1111/ecin.13051. URL: https://onlinelibrary.wiley.com/doi/10.1111/ecin.13051 (visited on 01/17/2022).

[BC94] Mark Bils and Jang-Ok Cho. "Cyclical Factor Utilization". In: *Journal of Monetary Economics* 33 (1994), pp. 319–354.

[BD91] Peter J. Brockwell and Richard A. Davis. *Time Series: Theory and Methods*. en. Springer Series in Statistics. New York, NY: Springer New York, 1991.

[BEA17] BEA. *Concepts and Methods of the U.S. National Income and Products Account*. Nov. 2017.

[BEF04] Craig Burnside, Martin Eichenbaum, and Jonas D.M. Fisher. "Fiscal shocks and their consequences". en. In: *Journal of Economic Theory* 115.1 (Mar. 2004), pp. 89–117. ISSN: 00220531. DOI: 10.1016/S0022-0531(03)00252-7. URL: https://linkinghub.elsevier.com/retrieve/pii/S0022053103002527 (visited on 01/17/2022).

[BP02] Olivier Blanchard and Roberto Perotti. "An empirical characterization of the dynamic effects of changes in government spending and taxes on output". en. In: *The Quarterly Journal of Economics* 117.4 (Nov. 2002), pp. 1329–1368.

[BP17] Nadav Ben Zeev and Evi Pappa. "Chronicle of A War Foretold: The Macroeconomic Effects of Anticipated Defence Spending Shocks". en. In: *The Economic Journal* 127.603 (Aug. 2017), pp. 1568–1597. ISSN: 0013-0133, 1468-0297. DOI: 10.1111/ecoj.12349. URL: https://academic.oup.com/ej/article/127/603/1568-1597/5068886 (visited on 04/23/2021).

[BR11] Robert J. Barro and Charles J. Redlick. "Macroeconomic Effects From Government Purchases and Taxes *". en. In: *The Quarterly Journal of Economics* 126.1 (Feb. 2011), pp. 51–102. ISSN: 0033-5533, 1531-4650. DOI: 10.1093/qje/qjq002. URL: https://academic.oup.com/qje/article-lookup/doi/10.1093/qje/qjq002 (visited on 08/31/2021).

[Bru20] Gillian Brunet. "When Does Government Spending Matter? Evidence from a New Measure of Spending". en. In: *Working Paper* (May 2020), p. 41.

[Cox+21] Lydia Cox et al. "BIg G". In: *Working Paper* (Jan. 2021).

[DG17] Bill Dupor and Rodrigo Guerrero. "Local and aggregate fiscal policy multipliers". en. In: *Journal of Monetary Economics* 92 (Dec. 2017), pp. 16–30. ISSN: 03043932. DOI: 10.1016/j.jmoneco.2017.07.007. URL: https://linkinghub.elsevier.com/retrieve/pii/S0304393217300855 (visited on 09/23/2020).

[EEF99] Wendy Edelberg, Martin Eichenbaum, and Jonas Fisher. "Understanding the Effects of a Shock to Government Purchases". en. In: *Review of Economic Dynamics* 2.1 (Jan. 1999), pp. 166–206.

[EF05] Martin Eichenbaum and Jonas Fisher. "Fiscal Policy in the Aftermath of 9/11". en. In: *Journal of Money, Credit, and Banking* 37.1 (2005), pp. 1–22. ISSN: 1538-4616. DOI: 10.1353/mcb.2005.0005. URL: http://muse.jhu.edu/content/crossref/journals/journal_of_money_credit_and_banking/v037/37.1eichenbaum.pdf (visited on 01/17/2022).

[Fer12] John G. Fernald. "A Quarterly, Utilization-Adjusted Series on Total Factor Productivity". en. In: *Federal Reserve Bank of San Francisco, Working Paper Series* (Apr. 2012), pp. 01–28. DOI: 10.24148/wp2012-19. URL: https://www.frbsf.org/economic-research/publications/working-papers/2012/19/ (visited on 02/02/2022).

[FM01] Antonio Fatas and Ilian Mihov. "The Effects of Fiscal Policy on Consumption and Employment: Theory and Evidence". en. In: (2001). Available at SSRN 267281 (2001), p. 39.

[FP10] Jonas D.M. Fisher and Ryan Peters. "Using Stock Returns to Identify Government Spending Shocks". en. In: *The Economic Journal* 120.544 (May 2010), pp. 414–436. ISSN: 0013-0133, 1468-0297. DOI: 10.1111/j.1468-0297.2010.02355.x. URL: https://academic.oup.com/ej/article/120/544/414-436/5089580 (visited on 08/31/2021).

[GCW21] Andres Gonzalez-Lira, Rodrigo Carril, and Michael S Walker. "Competition under Incomplete Contracts and the Design of Procurement Policies". en. In: *Working Paper* (Jan. 2021), p. 104.

[GK10] Robert Gordon and Robert Krenn. "The End of the Great Depression 1939-41: Policy Contributions and Fiscal Multipliers". en. In: (Sept. 2010), w16380. DOI: 10.3386/w16380. URL: http://www.nber.org/papers/w16380.pdf (visited on 10/02/2021).

[Ham18] James D. Hamilton. "Why You Should Never Use the Hodrick-Prescott Filter". en. In: *The Review of Economics and Statistics* 100.5 (Dec. 2018), pp. 831–843. ISSN: 0034-6535, 1530-9142. DOI: 10.1162/rest_a_00706. URL: https://direct.mit.edu/rest/article/100/5/831-843/58479 (visited on 04/05/2021).

[HP09] Karen Horowitz and Mark Planting. *Concepts and Methods of the U.S. Input-Output Accounts*. en. Apr. 2009.

[Jor05] Òscar Jordà. "Estimation and Inference of Impulse Responses by Local Projections". en. In: *American Economic Review* 95.1 (Feb. 2005), pp. 161–182. ISSN: 0002-8282. DOI: 10.1257/0002828053828518. URL: https://pubs.aeaweb.org/doi/10.1257/0002828053828518 (visited on 09/22/2020).

[LW13] Sylvain Leduc and Daniel Wilson. "Roads to Prosperity or Bridges to Nowhere? Theory and Evidence on the Impact of Public Infrastructure Investment". en. In: *NBER Macroeconomics Annual* 27.1 (2013), pp. 89–142.

[MP20] José Luis Montiel Olea and Mikkel Plagborg-Møller. "Local Projection Inference is Simpler and More Robust Than You Think". en. In: *Working Paper* (Dec. 2020), p. 45.

[NR11] Christopher Nekarda and Valerie Ramey. "Industry Evidence on the Effects of Government Spending". en. In: *American Economic Journal: Macroeconomics* 3.1 (Jan. 2011), pp. 36–59. ISSN: 1945-7707, 1945-7715. DOI: 10.1257/mac.3.1.36. URL: https://pubs.aeaweb.org/doi/10.1257/mac.3.1.36 (visited on 09/23/2020).

[OW20] Ali Ozdagli and Michael Weber. "Monetary Policy through Production Networks: Evidence from the Stock Market". en. In: (May 2020), p. 79.

[Per07] Roberto Perotti. "In Search of the Transmission Mechanism of Fiscal Policy [with Comments and Discussion]". en. In: *NBER Macroeconomics Annual* 22 (Jan. 2007), pp. 169–249. ISSN: 0889-3365, 1537-2642. DOI: 10.1086/ma.22.25554966. URL: https://www.journals.uchicago.edu/doi/10.1086/ma.22.25554966 (visited on 09/23/2020).

[Per14] Roberto Perotti. “Defense Government Spending Is Contractionary, Civilian Government Spending Is Expansionary”. en. In: *NBER Working Paper* (May 2014), w20179. DOI: 10.3386/w20179. URL: http://www.nber.org/papers/w20179.pdf (visited on 03/31/2022).

[Ram11] Valerie Ramey. “Identifying Government Spending Shocks: It’s all in the Timing*”. en. In: *The Quarterly Journal of Economics* 126.1 (Feb. 2011), pp. 1–50. ISSN: 0033-5533, 1531-4650. DOI: 10.1093/qje/qjq008. URL: https://academic.oup.com/qje/article-lookup/doi/10.1093/qje/qjq008 (visited on 09/22/2020).

[Ram16] Valerie Ramey. “Macroeconomic Shocks and Their Propagation”. In: *Handbook of macroeconomics*. Vol. 2. Elsevier, 2016, pp. 71–162.

[Ram89] Valerie Ramey. “Inventories as Factors of Production and Economic Fluctuations”. en. In: *American Economic Review* 79.3 (June 1989), pp. 338–354.

[Ram91] Valerie Ramey. “Nonconvex Costs and the Behavior of Inventories”. en. In: *Journal of Political Economy* 99.2 (Apr. 1991), pp. 306–334. ISSN: 0022-3808, 1537-534X. DOI: 10.1086/261752. URL: https://www.journals.uchicago.edu/doi/10.1086/261752 (visited on 11/17/2021).

[RR04] Christina D. Romer and David H. Romer. “A New Measure of Monetary Shocks: Derivation and Implications”. en. In: *THE AMERICAN ECONOMIC REVIEW* 94.4 (2004), p. 64.

[RS98] Valerie Ramey and D Shapiro. “Costly capital reallocation and the effects of government spending”. en. In: *Carnegie-Rochester Conference on Public Policy* 48.1998 (1998), pp. 145–194.

[RZ18] Valerie Ramey and Sarah Zubairy. “Government Spending Multipliers in Good Times and in Bad: Evidence from US Historical Data”. en. In: *journal of political economy* (Mar. 2018), p. 52.

[STV19] Frank Smets, Joris Tielens, and Jan Van Hove. “Pipeline Pressures and Sectoral Inflation Dynamics”. en. In: *SSRN Electronic Journal* (2019). ISSN: 1556-5068. DOI: 10.2139/ssrn.3346371. URL: https://www.ssrn.com/abstract=3346371 (visited on 01/25/2022).

[SW18] James H. Stock and Mark W. Watson. “Identification and Estimation of Dynamic Causal Effects in Macroeconomics Using External Instruments”. en. In: *The Economic Journal* 128.610 (May 2018), pp. 917–948. ISSN: 0013-0133, 1468-0297. DOI: 10.1111/ecoj.12593. URL: https://academic.oup.com/ej/article/128/610/917-948/5069563 (visited on 05/03/2021).

[WY20] Johannes F. Wieland and Mu-Jeung Yang. “Financial Dampening”. en. In: *Journal of Money, Credit and Banking* 52.1 (Feb. 2020), pp. 79–113. ISSN: 0022-2879, 1538-4616. DOI: 10.1111/jmcb.12681. URL: https://onlinelibrary.wiley.com/doi/10.1111/jmcb.12681 (visited on 08/26/2022).

Appendices

A Details on Defense Procurement in the Data

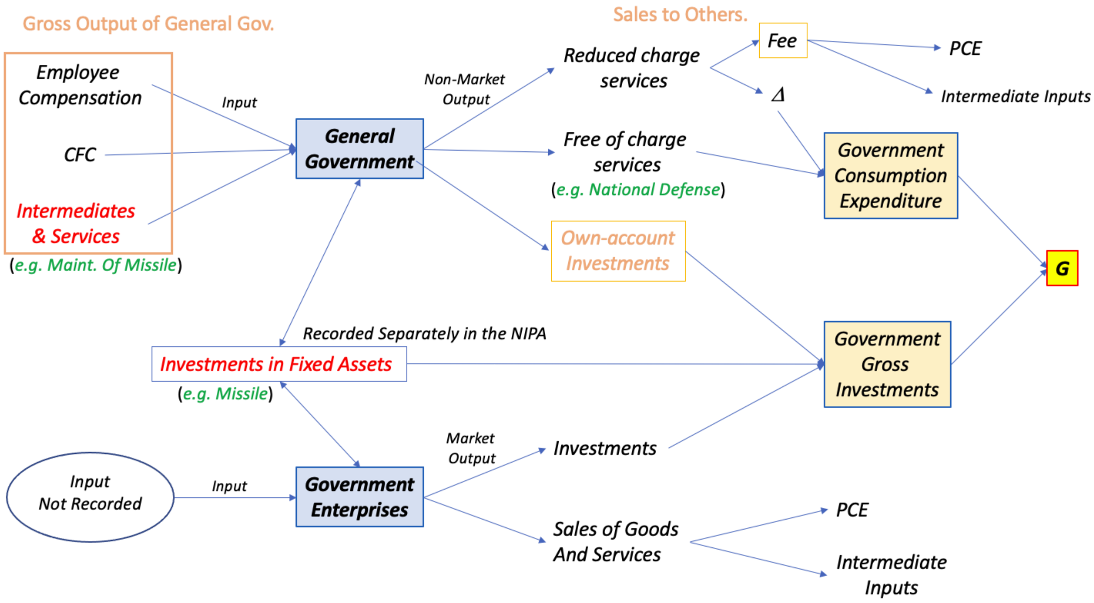
In this section of the Appendix we outline the details about measurement of defense procurement spending.

Section A.1 clarifies the accounting origin in the NIPA of outlays which refer to the purchase of goods. Section A.2 shows how we calculate the 2 to 3 quarters delay between defense procurement obligations and spending. Section A.3 uses contract level data from the 2000 to rationalize the existence of a time delay. Section A.4 illustrates how we construct the quarterly time series of defense procurement obligations. Section A.5 uses data from the 2000 on contracts' opportunities (i.e. contract level *solicitations*) to show that it is unlikely that contracts awards are anticipated by more than a quarter.

A.1 Accounting Origin of Procurement in the NIPA

In this section we provide further details on the accounting origin of public procurement contracts in the NIPA tables. Figure 11 summarizes the accounting of G, according to Chapter 9 of BEA (2017), which explains how the NIPA record all the entries of G. It highlights in red the two entries which contain public procurement spending: (i) Intermediate Goods and Services and (ii) Investment in Fixed Assets.

Figure 11: Accounting of G - Summary



CFC means "Cost of Fixed Capital" and measures depreciation of government assets. PCE means Personal Consumption Expenditure, the NIPA measure of Consumption which absorbs reduced charge services from the government (e.g. tuition fees from public universities). Own Account Investment is own resources reinvested in the public capital stock.

First of all, notice that G is made of two components, consumption and investments:

$$G = \underbrace{\text{Government Consumption Expenditure}}_{G^C} + \underbrace{\text{Government Gross Investments}}_{G^I}$$

Government Consumption Expenditure: Government consumption originates from the gross output of the government after deducting (i) Sales to Other Sectors and (ii) Own-Account Investments:

$$G^C = \underbrace{\text{Compensations} + \text{CFC} + \text{Intermediates and Services Purchased}}_{\text{Gross Output of General Government}} - \ldots$$
$$- \text{Own Account Investments} - \text{Sales to Others}$$

When a general government entity (e.g. DoD) decides to purchase goods and/or services, they are mainly accounted as Intermediates which eventually end up in G as government consumption.

Government Gross Investments: The government also makes three types of investments. Firstly, the General Government makes Own-Account-Investments, which are deducted from the gross output of general government in order to account them as investments. Secondly, both the General Government and Government Enterprises make investments in fixed assets. Investment in Fixed Assets contain other purchases from the private sector.

Example A.1 (Purchasing a Missile)**.** To clarify this point, consider the case of the government purchasing a new set of guided missiles:

1. The missile is accounted as Equipment in the Federal Defense category of Change in Government Fixed Assets and therefore contributes to G as part of Government Gross Investments.

2. Installation, Maintenance, Quality Control and other services related to the Missile are accounted as Intermediate Goods and Services Purchased (input of production).

3. The missiles and the related services are used to produce a non-market output of production, namely, national defense.

The production of the missile shows up in business inventories as long as the contractor supplying the missile delivers it to the government. Once delivered, inventories decrease and government investment increase. Notice that the reduction in inventories and the corresponding increase in G is a zero-sum game which does not increase GDP (recall that GDP in the US is constructed as the sum of final demand). GDP increases while production takes place and is recorded as inventories. In absence of time-to-build, inventories do not increase and the purchase of the item by the government directly shows up in G. For instance, this is the case of the Installation, Maintenance, Quality Control and other services related to the missile purchased by the Government.

Finally, Figure 12 provides an example of official accounting table of G, namely NIPA Table 3.10.5A, taken from BEA (2017).

A.2 Time Mismatch Between Obligations and Payments

In Section A.2 we show how we construct a proxy for defense procurement spending using data from the NIPA. We now show how we construct the defense procurement obligation proxy. Recall that obligations arise when the DoD awards new contracts while spending reflect government outlays, that is, payment to contractors.

We observe obligations through two data sources, discussed below.

Figure 12: NIPA Table 3.10.5A - Example

CHAPTER 9: GOVERNMENT CONSUMPTION EXPENDITURES AND GROSS INVESTMENT

Table 9.1—Government Consumption Expenditures and Gross Investment and Government Gross Output
[2012, billions of dollars]

Government consumption expenditures and gross investment	**3,158.6**
Consumption expenditures	**2,544.2**
Gross output of general government	3,036.7
Value added	2,028.6
Compensation of general government employees	1,592.5
Consumption of general government fixed capital	436.1
Intermediate goods and services purchased	1,008.1
Durable goods	72.6
Nondurable goods	296.9
Services	638.6
Less: Own-account investment	73.2
Sales to other sectors	419.4
Gross investment	**614.4**
Structures	282.4
Equipment	142.8
Intellectual property products	189.2

Don't enter GDP as Sum of Value Added

Enter GDP as sum of Final Demand

Business Condition Digest The periodical *Business Conditions Digest*, available on Fraser at this link, provided Business Cycle Indicators, among which a list of Defense Indicators. The original source of the data was the Department of Defense, Office of the Assistant Secretary of Defense (seasonal adjustment implemented by BEA). In particular, we use Series 525, "*Defense Prime Contracts Awards*". This series was firstly collected by Valerie Ramey for her papers: Ramey (1989) and Ramey (1991). We are grateful to her for providing the data.

The periodical was issued monthly from October 1961 until March 1990, however data is available from January 1951 to November 1988.

Business Condition Digest was discontinued in March 1990, unfortunately data on prime contracts stopped being recorded on it from December 1988 (all year 1989 is missing). Most business indicators on *Business Condition Digest* were moved to another monthly periodical, namely the *Survey of Current Business*. Prime award contracts (series 525) was preserved and moved to Appendix C on *Business Cycle Indicators* (section 2.4: other important economic measures/government activities). Data is available as scanned versions of the *Survey of Current Business* at this link. For some reason, data starting from 1991 did not report values of prime contracts awards for months belonging to the 4th quarter of the year. We believe this is a systematic omission which makes the source of data less reliable for this time period.

Finally, due to resources re-organization at the BEA, the *Business Cycle Indicators* section was discontinued and prime award contracts were no longer disclosed to public, as specified in the joint November-December 1995 issue. Therefore data is not available.

To summarize, we have reliable monthly data on prime contract awards from January 1951 to November 1988. We plot defense prime contract awards versus defense procurement spending in the top-left panel of Figure 6 in Section 3 of the paper. Notice that in order to match the quarterly frequency of procurement spending, obligation data is aggregated by quarters. Moreover, since NIPA data are annualized (their quarterly averages return their yearly values), we do the same for obligation data to allow for a comparison between the two series.

Firstly, it is clear from the graph that obligations lead spending. Secondly, notice that obligation data tends to be more noisy than spending data. The main reason for this is that large contracts are often awarded and then terminated a few months later for convenience or due to litigation with a losing offeror (this is also highlighted in Auerbach, Gorodnichenko, and Murphy (2020)). Moreover, obligations are more lumpy than payments which get smoothed over the duration of a contract. In order to account for this, we use a simple MA smoother (red line in the graph). We then provide a quantitative assessment of the delay by looking at the lead-lag correlation map between the growth rates of smoothed obligations and the growth rates of spending (see top-right panel).[29]

Overall, we find a positive correlation between the two series which increases when obligations are delayed (top-right quadrant of the lead-lag correlation map). In particular, correlation spikes when obligations are delayed by 2, 5 and 8 quarters. Results are robust to a different approach which looks at the lead-lag correlation between year-to-year quarterly changes ($\nabla_4 \, x_t = (1-L^4)x_t = x_t - x_{t-4}$) of original -i.e. not smoothed - obligations and spending. In this case, the spikes happen at 2 and 5 quarters.

Federal Procurement Data System Next generation: On September 26th 2006, the Federal Funding and Accountability Act is passed by congress as a first step towards a more transparent procurement system, which allows full disclosure of information involving federal contracts. The transparency effort by FFATA culminates in 2019 with the opening of a public website, USASpending.gov, which discloses information on all federal procurement contracts.[30] Data from USASPending.gov is pulled from FPDS-NG, the Federal Procurement Data System Next Generation, which actually includes the whole universe of procurement contracts. FPDS is the system used by government contracting officers to officially input data on awarded contracts in the government-wide system. Data from FPDS can be downloaded from USASpending.gov. The data spans 2000Q4 to the present

[29]We look at growth rates ($\Delta_1 \, x_t = (1-L)^1 x_t = x_t - x_{t-1}$) to cope with the non-stationarity of the series.

[30]More information on the history of USASpending.gov can be found here.

with a caveat: contract data awarded before the beginning of the construction of the database could have been lost or not recorded. We collect data on all defense procurement contracts on FPDS between 2000Q4-2020Q3.

We again compare obligations and spending in Figure 13.

Figure 13: Accounting Mismatch - January 2000 onward

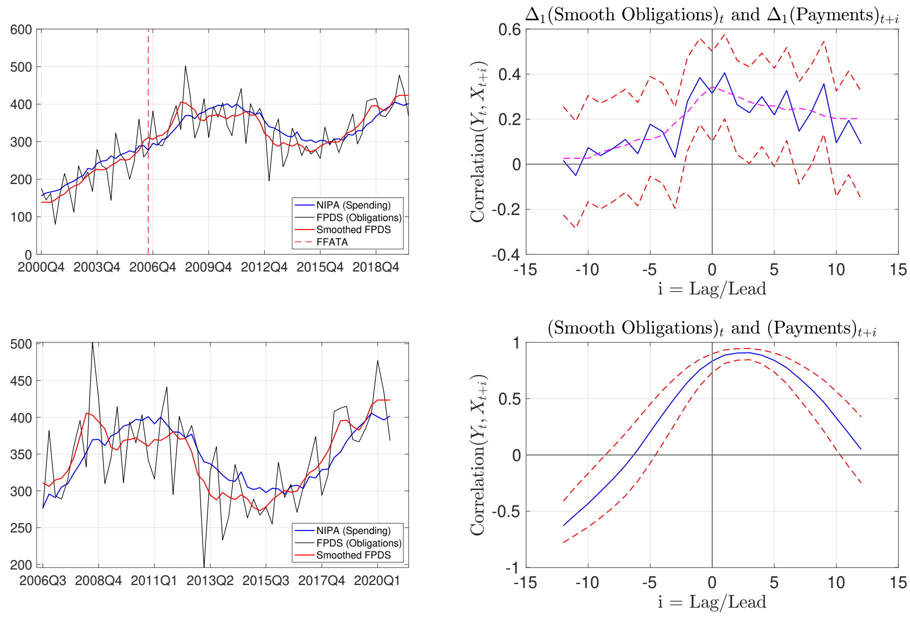


Figure 13's notes: the FPDS measure of obligation (blue line) is constructed by: (i) summing the daily data to obtain quarterly data; (ii) annualizing the data as done by NIPA; (iii) de-seasonalize data using Brockwell and Davis (1991)'s method (the Matlab code can be found here).

The top-right panel plots again the lead-lag correlation between the growth rates of (smoothed) obligations and the growth rates of spending. The highest correlation is recorded when obligations are delayed by 1 quarter. Once again, the results are robust to looking at the lead-lag correlation of year-to-year quarterly changes between original obligations and spending. In this case, the peak occurs from 0 to 2 quarters.

Before the signing of the FFATA, 2006, obligations seem under-reported relative to spending, thus inducing a downward bias in the estimation of the accounting time-mismatch. We have two possible explanations for this counter-intuitive result: firstly, this might be a consequence of missing contract modifications awarded before the introduction of FPDS; secondly, those modifications could have been classified before FFATA made most of contracts available to the public. In fact, FPDS-obligation data catches up and starts leading spending after the signing of the FFATA. Moreover, we show in Figure 14 that the share of large contracts (top 1%, 5% and 10%) out of all procurement spending stabilizes after 2006, indicating that large classified contracts are not showing up in FPDS.

We take this into account and we repeat the analysis only on those quarters following the signing of the FFATA (bottom panel of Figure 13). We observe a single clear spike in the lead-lag correlation, which indicates that obligations are delayed by 3 quarters relative to payments.[31]

Summary of Time Mismatch We summarize the time delay between obligations and spending in Table 3

Table 3: Summary of Time Mismatch Between Spneding and Obligations

Period	*Data Source*	*Correlation Spike Delay (Quarters)* Δ_1	∇_4
1951M1 to 1988M11	BCD	2-5	2-5
2000M10 to 2020M9	FPDS	1	2
2006M1 to 2020M9	FPDS	3	-

These results suggest that the accounting delay between beginning of production (award date) and the first payment (outlay) is on average between 2 to 3 quarters. Notice also that the time delay seems to shorten over time, when we use FPDS data.

Overall, our results are consistent with anecdotal evidence that government payments happen once every 180 days.[32]

A.3 Rationalizing the Time Mismatch

In this section we rationalize the existence of an aggregate time-mismatch between defense procurement obligations and spending. In particular, we provide both theoretical and empirical micro-level evidence of the time mismatch.

Duration of Defense Procurement Contracts Firstly, a necessary condition for the existence of an accounting mismatch is the long duration of contracts. If

[31]The peculiar non-trending sinusoidal-wave shape of the data referring this period allows us to directly look at the correlation between the two series in levels. The super-positioning of waves which happen when we shift one series back and forth in time, allows to observe a single clear spike which refers to the exact period when the two series overlap. The correlation spikes when obligations are delayed by 3 quarters.

[32]We confirm this timeline in discussions with several federal procurement contractors.

contracts were less than 90 days in duration, then payments would be processed in the same quarter as the award date.

We use FPDS data pulled from USASpending.gov from 2000Q4 to 2020Q3 to construct the distribution of duration of defense government contracts.

In this context, contracts have two main types: (i) single transaction and (ii) multiple transaction.[33] We calculate the duration of a single transaction contract from the award date to the end of work. The award date almost always indicates the start of work associated with a contract. To calculate the duration of multiple transaction contracts, we take the oldest contract modification end date and subtract from it the "new-action" award date.[34] Table 4 shows contract durations without distinguishing between single and multiple transaction contracts.

Table 4: (Log)Duration of Defense Contracts

Stats		*Unweighted*		*Weighted (by Obligation)*	
		Duration (days)	*Log-Duration*	*Duration (days)*	*Log-Duration*
Percentiles	1%	0	0	0	0
	5%	0	0	46	3.85
	10%	0	0	193	5.27
	25%	3	1.39	514	6.24
	50%	20	3.04	1519	7.33
	75%	126	4.84	2962	7.99
	90%	377	5.93	4844	8.49
	95%	794	6.68	5464	8.61
	99%	2584	7.86	6887	8.84
Mean		173.03	3.09	1988.02	6.94
Std.		485.32	2.14	1746.81	1.57
Min.		0.00	0	0.00	0
Max.		7300.00	8.89	7300.00	8.89

Table 4: defense contracts are identified by reporting DoD as funding/awarding agency. Data is taken from FPDS, all defense contracts from 2000Q1 to 2020Q1. Sample is restricted to those contracts for which the entire history of transactions (from the first new contract to the last modification) are available. Number of contracts in the sample is about 17 millions. Almost 5 thousands contracts with duration more than 20 years (7,300 days) are eliminated from the sample.

The median contract duration is 20 days and 90% of contracts have duration

[33] Transactions which refer to the same contracts are pooled together through a unique contract identifier field in FPDS: contract_unique_key_identifier.

[34] This is possible because FPDS reports both the beginning and the end of the PoP (Period of Performance).

less than one year. These results are in line with the findings of Cox et al. (2021) and suggest that contracts have a short duration.[35] However, this measure does not take into account the size of contracts, as larger contracts might have longer duration. The right columns of Table 4 report the distribution of the contracts' (log)duration, weighted by the total obligation amount. The weighted distribution can be interpreted as the duration distribution of a \$1 of spending in defense procurement. The following remark characterizes the mean and median of this distribution.

Remark A.1 (Median/Mean Duration of \$1)**.** The median duration of \$1 of defense procurement spending is 4.16 years. The mean duration of \$1 of defense procurement spending is 5.44 years.

Notice that after weighting, the shape of the distribution drastically changes. This suggests that procurement spending is characterized by a small number of large and long-duration contracts. We confirm this in Figure 14, which plots the share of total spending of the largest 1%, 5% and 10% of contracts. We find that the largest 10% of contracts account for 95% of total spending, on average. Similarly, the top 1% of contracts accounts for roughly 80% of total spending on average.

To summarize the results of this section: (i) large contracts make the bulk of defense spending and (ii) large contracts have long duration.

Partial Delivery Payments: Now, we want to rationalize the observed aggregate time delay. We do so by assuming there exists a "representative large contract" which follows a specific delayed payment schedule consistent with partial delivery payments.

Firstly, consider an example of a top 5% defense contract from FPDS. For instance, on December 22nd, 2015, the Department of Defense (DoD) awards a new "multi-transactions" contract to L-3 Communications Corporation.[36] The contract has a duration of two years and involves the reparation and maintenance of some aircraft components and accessories.[37] At the time of obligation, this contract has several components, denoted child contracts, and 24 contract modifications. Each modification represents a new child contract with its own duration.[38]

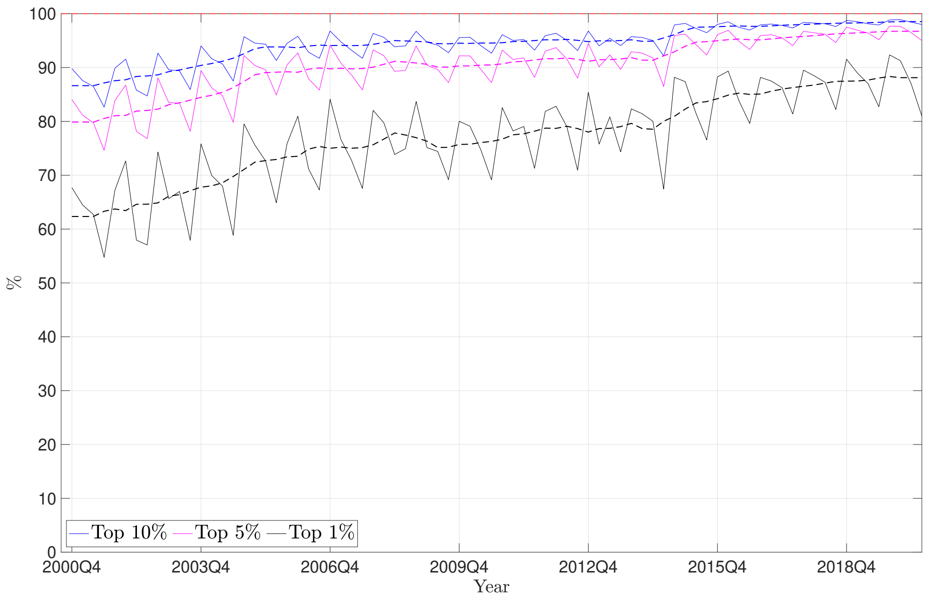
[35] They use a sample from 2001 to 2018 and find a median duration of defense contracts of 26 days and 90% of them have duration less than a year.

[36] See contract here.

[37] Duration is measured as the number of days between the Period of Performance (PoP) end date and the PoP start date.

[38] Modifications can have two types: (i) uni-lateral (e.g., administrative actions which obligate new funds for the specific contract) or (ii) bi-lateral (e.g., change to the original orders or additional work).

Figure 14: Large Contracts Share of Total Procurement Spending



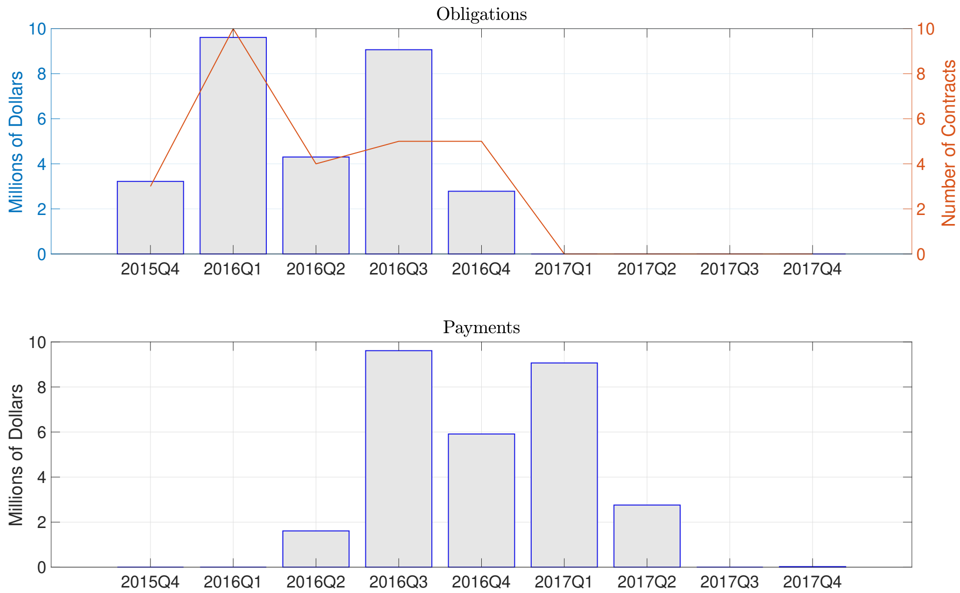
In the top panel of Figure 15, we show on the left axis the amount of dollars obligated every quarter by this contract, and on the right axis the number of (child) contracts signed every quarter. The bottom panel shows the corresponding payment schedule which assumes that payments are disbursed once every 180 days, by uniformly spreading the initial amount of obligated funds over a contract duration.[39]

For instance, the first new child contract, signed in December 2015, lasts 375 days and obligates almost \$3 million by the DoD. The payment schedule assumes that the contractors start producing the parts to be replaced immediately with partial delivery and partial reimbursement after 180 days, Therefore, the contractor is paid \$1.5 million in June 2016. Finally, in December 2016, the period of performance ends and the DoD pays to the contractor the remaining half of obligated funds.

Notice that payments look like a delayed version of obligations for this particular contract. The choice of 180 days delay between payments is consistent with our estimates for the average time mismatch between defense obligations and payments found earlier. The assumption of uniform production and payments is standard and consistent with the work of Auerbach, Gorodnichenko, and Murphy (2020). In general, contractors are often incentivized to distribute production as-

[39]This assumption is also made in Auerbach, Gorodnichenko, and Murphy (2020).

Figure 15: Example of a Contract's Obligation and Payment Schedule



sociated with an obligation throughout the whole duration of the contract.[40] In the data, cost-overruns and delays are common (see e.g., Gonzalez-Lira, Carril, and Walker (2021)).

Therefore, consider a representative contract with a structure similar to the one just analyzed: few new child contracts followed by several modifications. Overall, the contract lasts 48 months - consistent with the median weighted duration of defense contracts (see Table 4) and is characterized by payments disbursed once every 6 months (for a total of 8 payments). If we denote by P_t the total payments to contractors at time t and by O_t the amount of obligations, it is easy to show that the mapping between spending and obligations is given by the following equation:

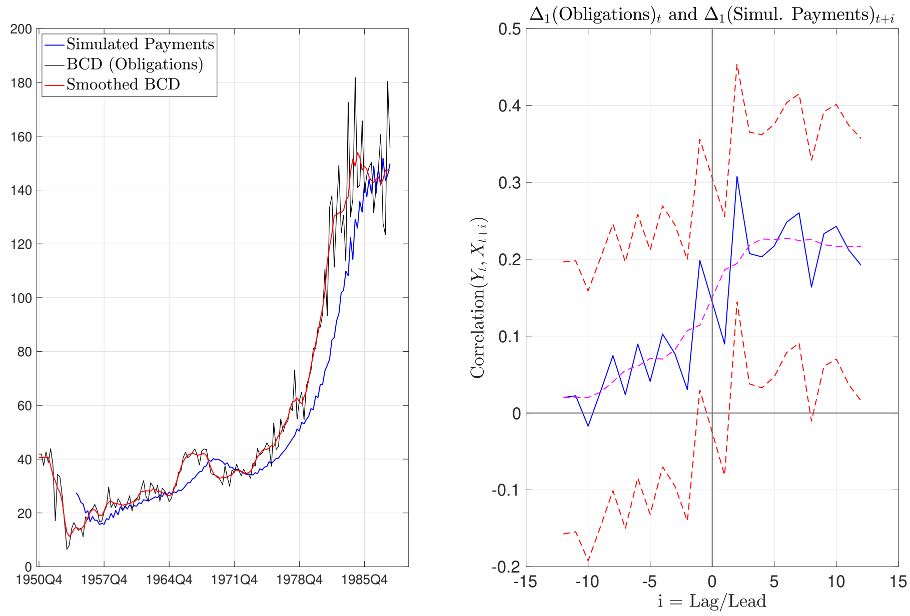
$$P_t = \frac{1}{8} \cdot \sum_{j=1}^{8} O_{t-6 \cdot j}. \tag{5}$$

We take the obligation data from BCD and feed it into Equation (5) to construct a time series of simulated payments. The left panel of Figure 16 plots BCD defense obligations data and the so constructed payments.

Despite the simplicity of the payments data generating process given by Equation (5), the simulated payments data approximate quite well the actual ones

[40]Consider a simple firm optimization problem with convex adjustment costs.

Figure 16: BCD Obligations and Simulated Payments



shown in Figure 6 of Section 3 of the paper. Similarly, the right panel shows the lead-lag correlation map between the growth rates of (smoothed) obligations and simulated payments. Notice that the results are very similar to the ones obtained using real spending data.[41]

A.4 Construction of Quarterly Defense Obligation

We show here how we construct the time series of defense procurement obligations.

We face two main challenges: (i) we have obligations data only from 1951 to 1989 and from 2000 onward; (ii) obligations are very lumpy because contracts also get cancelled and we want to focus on obligations which turn into actual production.

i. We turn BCD and FPDS monthly data into quarterly annualized data (sum monthly observations and multiply by 4).

ii. We apply the standard Brockwell and Davis (1991) filter to deaseasonalize the data.

[41]We highlight that by allowing time varying number of payments (here 8) and payments delays (here 6 months), we can improve by far the approximation to the actual data. Here, we preferred to keep things simple in the interest of brevity and clarity.

iii. We construct a time trend which takes value of 1 in 1947Q1. Denote it by t.

iv. We predict obligations using 4 leads and lags and contemporaneous defense procurement spending, as well as time trends t and t^2.

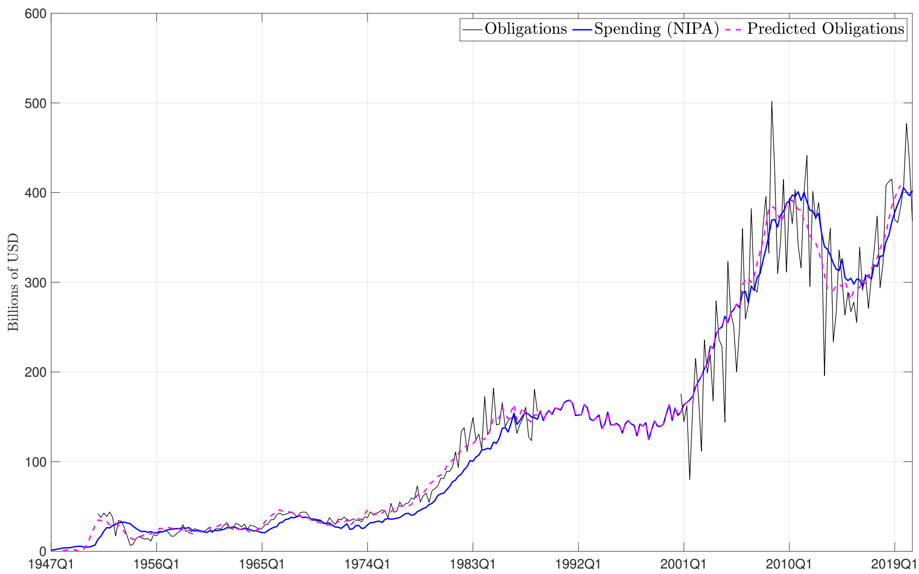
v. We construct obligations from 1951Q1 to 1988Q4 using the predicted values from the previous regression. We use the estimated coefficients and the values of defense procurement spending from 1947 to 1951 to extrapolate obligations for those years.

vi. We predict obligations from year 2006 onward (FFATA introduction) in the same way. We use the predicted values to be our new series of obligations for those years. Since defense procurement (smoothed) obligations and spending overlap from 2000 to 2006, we use actual defense procurement spending for those years.

vii. From 1989 to 2000 we use defense procurement spending to proxy obligations which turn into actual production.

Figure 17 plots obligations, defense procurement spending, obligations and the predicted ones.

Figure 17: Quarterly Defense Obligation and Spending



A.5 What Goes On Before Contract Awards?

Although public procurement contracts are awarded at a highly decentralized level (i.e., by over 69 federal agencies, 209 sub-agencies), all contracting officers are required to abide by the guidelines proposed in the Federal Acquisition Regulation (FAR). The FAR is a set of principles and procedures intended to organize and guide the procurement process across all federal agencies. In this section, we focus on the publicizing requirements associated with procurement contracts, depicted in Figure 18.

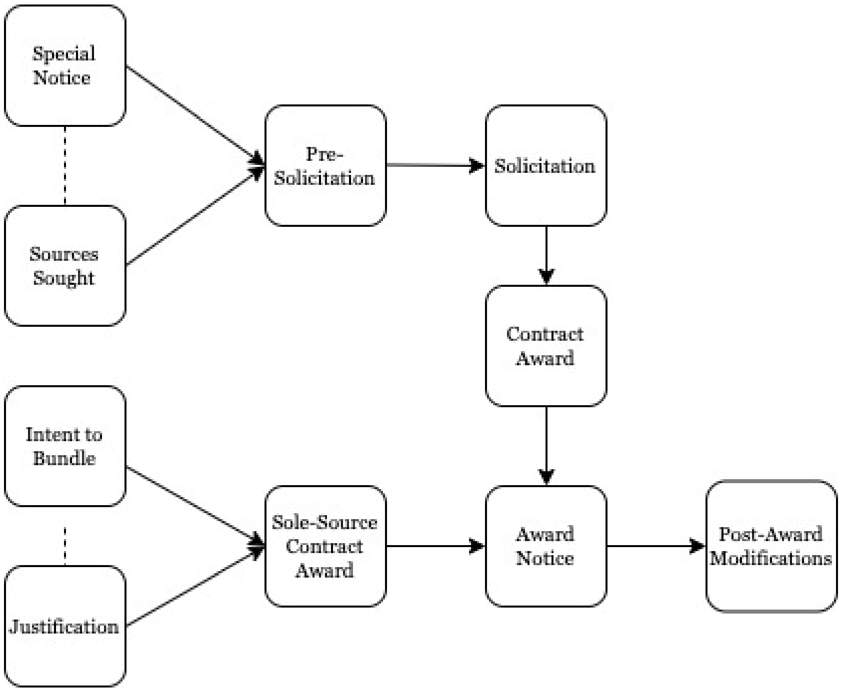


Figure 18: **Timeline of the Procurement Process**
Notice prior to contract award step occur on average within the quarter. Source: `beta.sam.gov` daily files.

In particular, FAR Part 5 (*Publicizing Contract Actions*) requires that contractors publicize contract opportunities with the goal of increasing competition, broadening industry participation, and assisting small businesses in obtaining contracts. Since October 1, 2001, contract actions with an expected value of over $25,000 must be publicized in an online and easy-to-access government platform, which we refer to as *Contract Opportunities*. Contract actions below the threshold might still be posted to increase visibility. On the other hand, FAR allows for exemptions to the requirement above the threshold when the posting might "compromise national security" or when the posting is "not in the government's interest". The result is that many contracts which are awarded are never solicited.

When the regulation applies, Contract Opportunity notices are posted publicly at beta.sam.gov and include award notices such as solicitations, pre-solicitations, or other pre-award and post-award actions. We describe the types of contract notices below.[42]

Special Notice Agencies use Special Notices to announce important pre-award events such as business fairs, long-run procurement forecasts, or pre-award conferences and meetings. Special Notices might also refer to "Requests for Information" (RFI) or draft solicitations.

Sources Sought Agencies post Sources Sought Notices in order to seek possible sources for a project. As discussed in FAR 7.3, the Sources Sought notice is not a solicitation for work or a request for proposal. Agencies typically use Sources Sought notices to collect industry feedback on key contracting strategy decisions and to perform market research on firm capabilities.

Pre-Solicitation Agencies post a pre-solicitation to notify vendors that a solicitation may follow. Potential vendors might then express interest in the contract by adding themselves to the Interested Vendors List. Government agencies use pre-solicitations to determine the number of qualified vendors to perform the desired work. Contracting officers can also use pre-solicitations to gather information on interested suppliers and determine if a set-aside for a small business might be applicable.

Intent to Bundle Requirements Agencies post "Intent to Bundle Requirements (DoD-Funded)" (ITB) whenever awarding actions are funded solely by the DoD. ITB supports the requirements in Section 820 of the Fiscal Year (FY) 2010 National Defense Authorization Act (NDAA) for contracting officers to post a notice of intent to use contract bundling procedures 30 days prior to releasing a solicitation or placing an order - if a solicitation is not required.

Solicitation Agencies post a solicitation to clearly define government requirements for a potential contract so that businesses can submit competitive bids. A "Request for Proposal" (RFP) is the most common type of solicitation used by

[42] Gonzalez-Lira, Carril, and Walker (2021) also provides a useful description and analysis of the publicizing requirements for Federal Procurement and the effects of information diffusion via public notices. We thank Andres Gonzales-Lira for directing us to the General Services Administration Technical Documentation for the FedBizOpps (FBO) website, whose information is now migrated to Contract Opportunities.

federal agencies. The solicitation also sets conditions and requirements for contractor proposals and includes the government's plan for evaluating submissions for potential award.

Combined Synopsis/Solicitation Agencies post a combined synopsis/solicitation when a contract is open for bids from eligible vendors. The Synopsis/Solicitation includes specifications for the product or service requested and a due date for the proposal, as well as the bidding procedures associated with the solicitation.

Award Notice Agencies post an award notice when they award a contract in response to a solicitation. Federal agencies may choose to upload a notice of the award to make aware other interested vendors of the winning bid. Note that the requirement guidelines for posting the award notice vary based on the agency and the solicitation.

Justification Agencies are required to post a justification in order to obtain approval to award a contract without posting a solicitation as required by FAR 41 U.S.C. 253(c) and 10 U.S.C. 2304(c). Under certain conditions, agencies are authorized for contracting without full and open competition. The Department of Defense, Coast Guard, and National Aeronautics and Space Administration are subject to 10 U.S.C. 2304(c). Other executive agencies are subject to 41 U.S.C. 253(c). Contracting without providing for full and open competition or full and open competition after exclusion of sources is a violation, unless permitted by one of the exceptions in FAR 6.302.

Sale of Surplus Property Agencies post a sale of surplus property notice when they wish to sell federal real estate properties for public use. These properties are typically made available for public use to state and local governments, regional agencies, or nonprofit organizations to state and local governments. Public uses for properties are those that are accessible to and can be shared by all members of a community, and include community centers, schools and colleges, parks, municipal buildings and many more.

A.5.1 Timing of Solicitations

Although there is still uncertainty about the contract award when a pre-award notice is posted, firms might still take action in anticipation of the award. This might occur if a firm wants to become more competitive in the bidding process or predicts a contract win with a high probability. In addition, some pre-award notices justify the lack of competition in a sole-sourced contract proposal. In this case, the contractor might even predict a contract award with full certainty.

We argue from the data that any anticipatory behavior is likely to occur at a frequency higher than the frequency of aggregate analysis in this work. In other words, almost all information about contract opportunities is revealed to contractors within the quarter of the contract award. We summarize this finding by notice type in Table 5 and plot the distribution of pre-award notice lags in Figure 19.

Table 5: **Average Lag Between Pre-Award Notices and Award Date**
Based on matched notices between FPDS and Contract Opportunities.

Notice Type	*Avg Lag in Days*	*Proportion of Notices*
Justification / Fair Opportunity	87	1.2%
Other	54	62.5%
Special Notice	41	2.1%
Pre-Solicitation	28	14.6%
Sources Sought	21	4.1%
Solicitation/Contract Solicitation	16	15.5%
TOTAL	43	

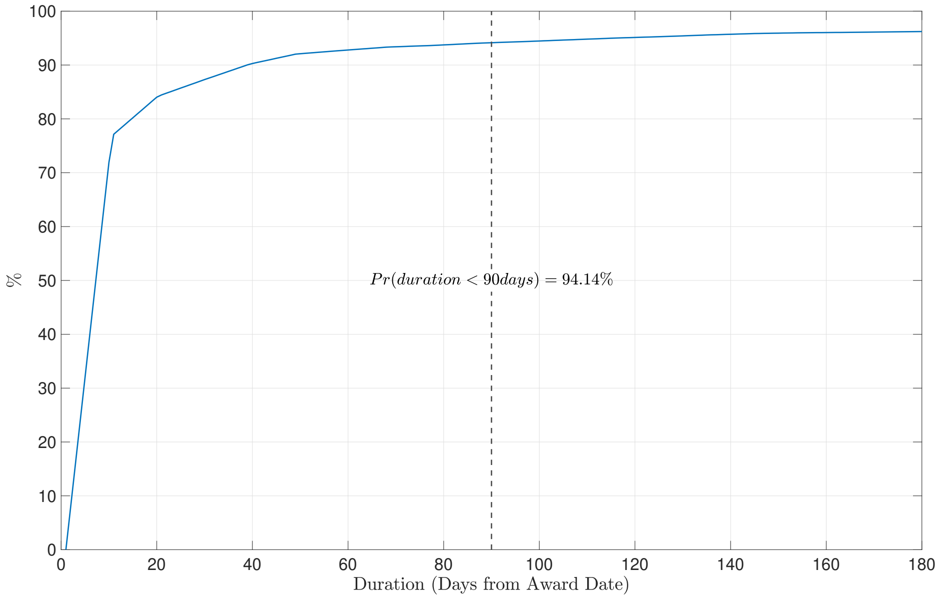
B Details on Industry Level Analysis

In this section of the Appendix we go over (i) the details of the construction of industry weights θ_i, which show up in equation (2) (see Appendix B.1) and (ii) the robustness checks that we perform for the industry level analysis of inventories (see Appendix B.2).

B.1 Construction of Industry Weights

We combine information from the Make and Use table with 60plus sectors from 1963 to 1996 and following the indications in the BEA's Horowitz and Planting (2009), we derive (i) direct requirement industry-by-industry matrices A_t and (ii) direct sales from the private sectors to the government. We use these two elements to construct our industry weights.

Figure 19: Empirical CDF of (Weighted) Solicitation Lag



Weighted duration between Contract Solicitation and Award dates measured in days. Dark dashed line represents 1 quarter (90 days). The Empirical CDF is estimated using Gaussian Kernel Density.

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

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

where t denotes a year, n is the number of manufacturing sub-industries that we consider, G is the federal general government and and the 0 subscript in the vector's name refers to the number of input-output orders of connection included (since the vector accounts only for direct sales to the government, the subscript is zero). Moreover, $\text{SALES}_{i\to G,t}$, where i denotes a generic manufacturing subsector, includes government gross investments, which show up as final uses in the Use tables.

The time-average values of $\boldsymbol{\gamma}_{0,t}$ are reported by industry in the third column of Table 6.

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

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

We construct a vector of direct and first-order indirect sales to the government shares of industry weights as follows:

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

Notice that the generic i-th element of $\boldsymbol{\gamma}_{1,t}$ is given by:

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

The time-average values of $\boldsymbol{\gamma}_{1,t}$ are reported by industry in the fourth column of Table 6.

Similarly, we construct direct, first and second order indirect sales to the government, shares of total output as:

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

The time-average values of $\boldsymbol{\gamma}_{2,t}$ are reported by industry in the fifth column of Table 6. Finally, we construct our industry weights θ_i as:

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

The weights are reported in the last column of Table 6.

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

Table 6: **Industry Weights.** Last column divides $\theta_{2,i}$ by the maximum value (i.e. the one of Other Transportation Equipment Manufacturing). In the paper, the weights θ_i that we use are the ones which include second order degree of connections, normalized (last column).

B.2 Robustness Checks of Industry-Level Analysis

Recall from Section 2 of the paper, that we carry out two robustness checks when we estimate the differential response of defense industries during military build-ups. In particular, we perform two placebo experiments when we estimate Equation (2), we (i) use war dates but we randomly shuffle the weights and (ii) we replace war dates with monetary policy shocks. The estimated average response (α_h) and the differential response of defense industries (β_h) are reported in the top and bottom panels of Figure 20, respectively.

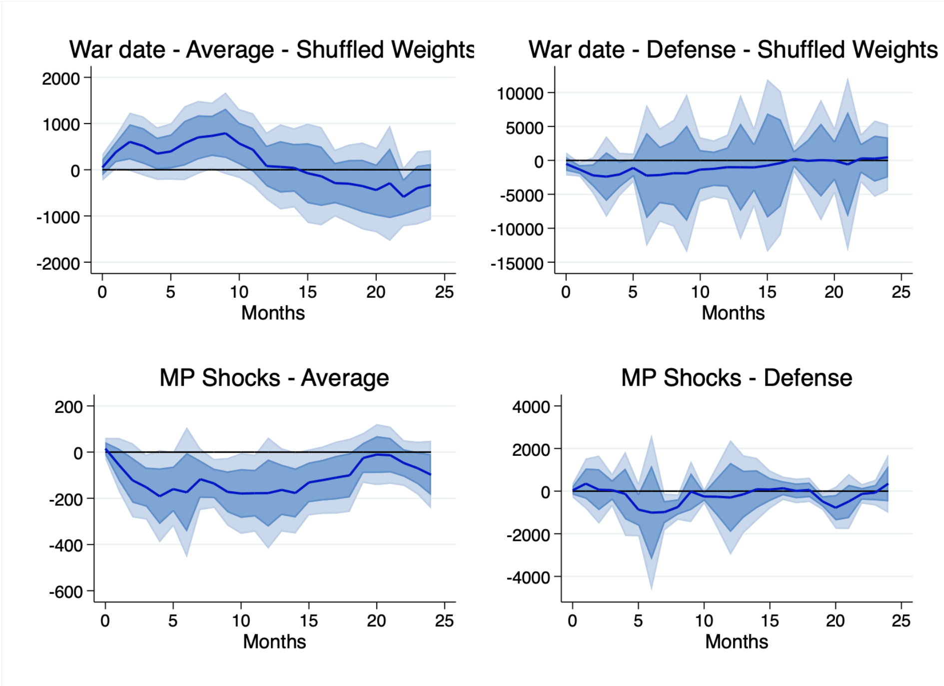


Figure 20: **Response of Sectoral Inventories to War Events (Robustness).**
Same as in Figure 3.

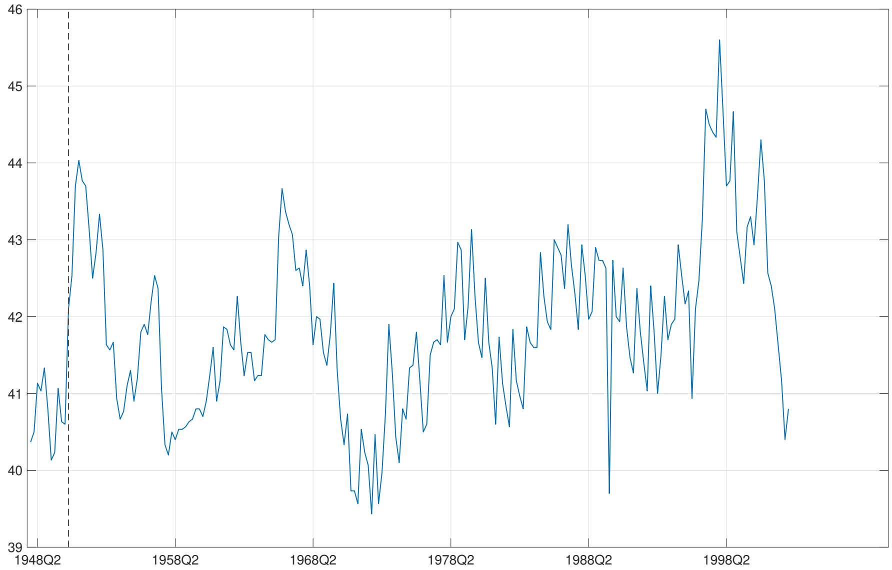
C Tracking Defense Industrial Production

In Section 4 we use use *Average Hours of Production Workers* of the *Aircraft industry*, showed in Figure 21, to keep track of the "*defense production machine*". We now explain the reasons behind that choice.

Aircraft Industry: We choose the Aircraft industry for two reasons: (i) great data availability (monthly data from BLS discontinued series starting from 1939) and (ii) high dependency on government purchases (see Nekarda and Ramey (2011)).

Hours-per-Worker: In general, measuring industrial activity requires the adoption of proxies for actual industrial output. In particular, in the case of the aircraft industry we don't observe the exact number of aircraft produced every month nor their percentage of completion. However, we have three variables which proxy industrial production: (i) average weekly hours of production workers, (ii) number of production workers (i.e. employment) and (iii) their product, namely total hours

Figure 21: Average Hours of Production Workers in the Aircraft Industry



worked. The first one is a measure of intensity of production, while the other two are stock variables measuring the extensive margin of production.

In order to understand which one is more suitable to measure changes in production's regime, consider the outbreak of the Korean war. Defense manufacturers foresaw a period of high demand of weapons by the government and needed to adjust production accordingly. The first sensible thing to do was to increase production given the predetermined level of capital and labor inputs. For instance, increasing production required extra use of electricity to run machinery in the assembly lines as well as a higher number of shifts with longer duration for production workers. By consequence, hours per worker increased straight away. Only over time, contractors expanded production by widening their stock of capital and workers, overcoming problems related capital immobility (see Ramey and Shapiro (1998)) and labor market frictions. As contractors expanded their production facilities and hired new production workers, intensity of production went back to normal.

This example highlighted two facts. Firstly, intensity of production of manufacturing industries is a good indicator of switches in the production regime. Secondly, intensity of production leads employment and other stock variable which moves more slowly. This intuition is consistent with Bils and Cho (1994), who find that hours per worker lead employment and the business cycle. Moreover, they emphasize how hours-per-worker co-moves with other relevant but unobserved measures of intensity of production.[43] Along these lines, Fernald (2012) suggest to use hours of production workers to proxy other unobserved measures of intensity of production; according to them, a cost-minimizing firm operates on all margins simultaneously, both observed (i.e. hours per worker) and unobserved (i.e. labor effort and workweek of capital).

In what follows, we show that (i) hours-per-worker in the aircraft industry leads employment and (ii) employment drives the dynamics of total hours, overshadowing those very informative lumpy changes in hours-per-worker.

In light of all this, *hours-per-worker is the most suitable variable to timely measure changes in production.*

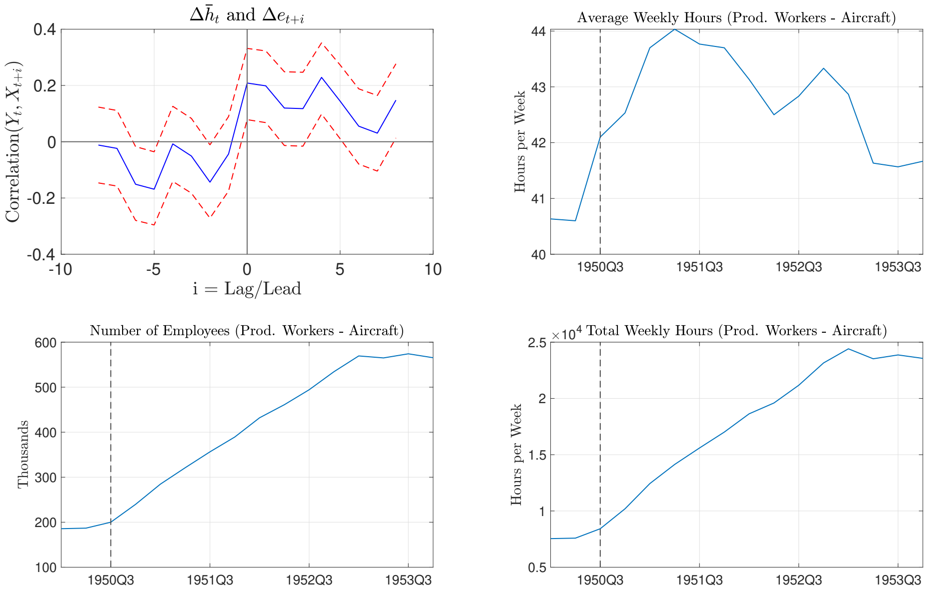
Hour per Worker, Employment and Total Hours in the Aircraft Industry: Figure 22 shows in its top-left panel the lead-lag correlation map between changes in average hours of production workers and changes in the number of production workers in the Aircraft industry. Clockwise from the top-right panel we show the time series of average hours of production workers ($\bar{h}_t$), number of production workers (e_t) and total hours of production workers ($\bar{h}_t \cdot e_t$) around the onset of the Korean war for the Aircraft industry (i.e. 1950Q3).

Firstly, from the lead-lag correlation map it is apparent that average hours of production workers lead employment. This is consistent with the findings of Bils and Cho (1994). Secondly, notice that the dynamics of total hours is dominated by employment, not by average hours per worker. Therefore, if we gauged industrial production by simply looking at the dynamics of total hours we would conclude that the response of the Aircraft industry at the outbreak of the Korean war was mild and slow. On the contrary, average hours per production worker anticipated the peak response of employment and total hours of production, signaling that the defense production machine had already fired up at the onset of the war.

We further clarify what is happening by breaking down the change in total hours into two components, one which accrues to changes in hours worked (intensive margin) and one which accrues to changes in number of workers (extensive

[43]They find that (i) "looms hours" are strongly related to hours per worker in the US textile industry and (ii) electricity use of manufacturing industries and hours worked per week co-moves. On the contrary, they find that the relationship between their measures of capital utilization and the number of production workers is much weaker.

Figure 22: Average Hours of Production Workers Vs Production Workers - Aircraft Industry



margin):

$$H_t = \bar{h}_t \cdot e_t$$
$$\frac{\partial H_{t+h}}{\partial z_t} = \underbrace{\frac{\partial \bar{h}_{t+h}}{\partial z_t} \cdot e_{t+h}}_{\text{Intensive Margin}} + \underbrace{\frac{\partial e_{t+h}}{\partial z_t} \cdot \bar{h}_{t+h}}_{\text{Extensive Margin}}$$

where z_t is a defense news shock. We break down the dynamic response of total hours to the Korean war using the previous expression:

$$(H_{1950Q3+h} - H_{1950Q2}) = \underbrace{\left(\bar{h}_{1950Q3+h} - \bar{h}_{1950Q2}\right) \cdot e_{1050Q3+h}}_{\text{Intensive Margin}} + \underbrace{\left(e_{1950Q3+h} - e_{1950Q2}\right) \cdot \bar{h}_{1050Q3+h}}_{\text{Extensive Margin}}$$

with $h = 0, 1, ..., H$. The breakdown is showed in Table 7: Notice that the dynamic of Total hours, H_t is dominated by the Extensive margin. Therefore, using total hours would overshadows the early change in hours-per-worker which is a clear signal that contractors were responding to the shock already in the third quarter of 1950.

Table 7: Breakdown Total Hours - Korean War

Date	$\bar{h}_t$	e_t	H_t	$H_{1950Q3+h} - H_{1950Q2}$	Int. Margin	Ext. Margin	Int. Margin (%)	Ext. Margin (%)
1950Q2	40.60	186.83	7585.43	0.00	0.0	0.0	-	-
1950Q3	42.10	200.00	8420.00	834.57	300.0	554.3	35.9%	66.4%
1950Q4	42.53	239.70	10195.24	2609.81	463.4	2248.6	17.8%	86.2%
1951Q1	43.70	284.57	12435.56	4850.13	882.2	4270.9	18.2%	88.1%
1951Q2	44.03	321.00	14134.70	6549.27	1102.1	5907.8	16.8%	90.2%
1951Q3	43.77	356.37	15596.98	8011.55	1128.5	7419.9	14.1%	92.6%
1951Q4	43.70	389.27	17010.95	9425.52	1206.7	8846.3	12.8%	93.9%
1952Q1	43.13	432.00	18633.60	11048.17	1094.4	10574.9	9.9%	95.7%
1952Q2	42.50	461.07	19595.33	12009.90	876.0	11654.9	7.3%	97.0%
1952Q3	42.83	494.30	21172.52	13587.08	1103.9	13169.8	8.1%	96.9%
1952Q4	43.33	534.37	23155.89	15570.46	1460.6	15059.8	9.4%	96.7%
1953Q1	42.87	569.43	24409.71	16824.28	1290.7	16400.8	7.7%	97.5%

C.1 Delay in the Fed's Defense Industrial Production Index

Notice that the Board of Governors of Federal Reserve System constructs a monthly real index of industrial production of manufacturing equipment in defense industries.[44]

The Fed makes clear that such defense production index is *mainly obtained from Production-hours (i.e. total hours) from the BLS. Hours are then used to infer output.* However, we have just seen that the dynamics of total hours worked is delayed relative to average hours worked. In fact, we now show that hour-per-worker in the Aircraft industry lead defense production as measured by the Fed.

In particular, we study the lead-lag correlation map between each labor margin and defense procurement obligations, production - as measured by the Fed - and spending. Figure 23 plots the results.

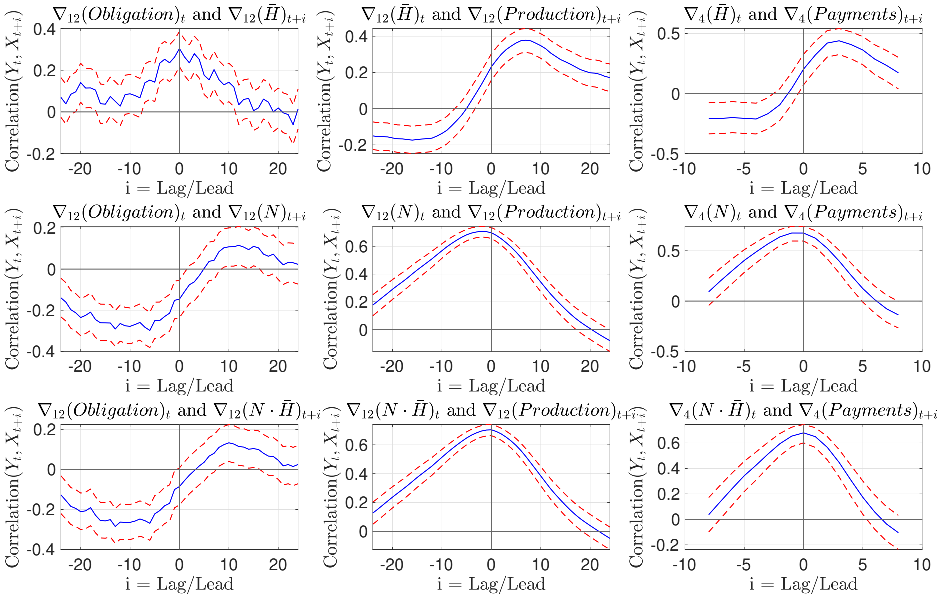
Firstly, looking at the first row, average hours of production workers in the Aircraft industry (intensive margin) appear to: (i) co-move with obligations, (ii) lead industrial output by 8 months (2 quarters) and (iii) lead payments by 4 quarters.

From the second row we notice that the number of production workers (extensive margin) appear to: (i) lag behind obligations (the delay is about 3 quarters), (ii) co-move with the production index and (iii) co-move with payments.

Finally, the third row shows that total hours of production workers co-move with industrial production as measured by the Fed, confirming the fact that the

[44] Data is available from 1947 to present at monthly and quarterly frequency, both seasonally adjusted and not. It can be downloaded at this link. Detailed information on the Real Index of Industrial Production of Manufacturing Equipment in Defense sector is available at this link. In particular, the underlying industries used for the construction of the series are discussed in these two tables: (i) market structure (Equipment); (ii) Industry Group (defense and space).

Figure 23: Lead-Lag Correlation Graph - Defense Industrial Production



Defense procurement spending is constructed as discussed in Section 3 and therefore tracks payments to contractors (sample from 1947Q1). Defense procurement obligations come from the original series from Business Condition Digest, discussed in Appendix A.2 and track new contract awards (monthly data from 1951M1 to 1988M11). Defense Production is the monthly seasonally adjusted index constructed by the Fed (data available from 1947M1). Hours and employment data come from the BLS discontinued data series on production workers data (available from 1939M1 to 2003M12).

Federal Reserve adopts total hours to construct the defense production variable. Moreover, the maps of total hours and employment are basically identical, confirming our previous finding that the dynamics of employment drives the one of total hours.

To summarize, we have showed that the Fed measures defense production using total hours of production workers. However, the dynamics of total hours is dominated by employment, which is a delayed measure of production which overlooks the ability of producers to ramp-up production by using more intensively their input of production (i.e. capital utilization and average hours worked). In fact, the Fed's measure lags behind defense procurement obligations but co-moves with spending, confirming that the delay which characterizes employment ends up contaminating their production index. In light of this, we believe that using average hours of production workers in the Aircraft industry is the best choice to timely measure defense production.

C.2 The effects of Military Build-Ups on Defense Production

We now calculate the IRFs of average hours of production workers in the Aircraft Industry, our proxy for defense production, to (i) defense news shocks shocks and (ii) defense procurement obligations. Results are shown in Figure 24.

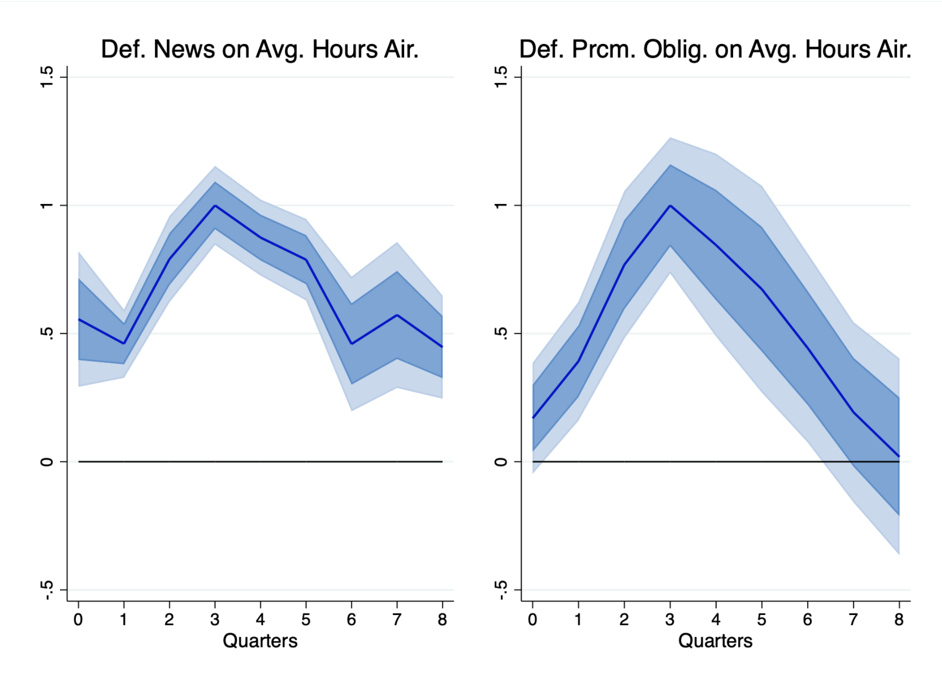


Figure 24: **Effects of Military Build-ups on Defense Production** *IRFs are obtained via lag-augmented local projections. Sample goes from 1947Q1 to 2002Q4. Standard errors are heteroskedasticity robust. Confidence bands are 90% and 68%.*

Notice that in response to either defense news or newly awarded contracts, defense production quickly ramps up.

D Breaking Down the Response of GDP

In this section we break down the response of GDP to a defense news shock in all of its four components. We do so by exploiting the linearity of the OLS estimates which are used to construct the IRFs via local projections.

First of all, we calculate the IRF of GDP to a defense news shock by regressing GDP on defense news shocks and four lags of investment, government spending, net-export, consumption total hours worked in the private sector, the 3-months

T-Bill rate, defense news shocks and a linear time trend. As usual all nominal variables are divided by nominal potential GDP (we take real potential GDP from Ramey and Zubairy (2018) and multiply it by the GDP price deflator). We group this set of lagged variables and the time trend into matrix X_t. The IRF of GDP is then obtained as:

$$\mathrm{GDP}_{t+h} = \theta_h^{\mathrm{GDP}} \cdot \mathrm{News}_t + X_t \cdot \boldsymbol{\beta}^{\mathrm{GDP}} + \varepsilon_{t+h} \quad h = 0, 1, ..., 8.$$

The estimated IRF of GDP is showed in the left panel of Figure 1 in the paper.

We repeat this procedure for all the four components of GDP and we estimate:

$$\begin{aligned} \mathrm{G}_{t+h} &= \theta_h^{\mathrm{G}} \cdot \mathrm{News}_t + X_t \cdot \boldsymbol{\beta}^{\mathrm{G}} + \varepsilon_{t+h}^{G} \quad h = 0, 1, ..., 8 \\ \mathrm{C}_{t+h} &= \theta_h^{\mathrm{C}} \cdot \mathrm{News}_t + X_t \cdot \boldsymbol{\beta}^{\mathrm{C}} + \varepsilon_{t+h}^{C} \quad h = 0, 1, ..., 8 \\ \mathrm{I}_{t+h} &= \theta_h^{\mathrm{I}} \cdot \mathrm{News}_t + X_t \cdot \boldsymbol{\beta}^{\mathrm{I}} + \varepsilon_{t+h}^{I} \quad h = 0, 1, ..., 8 \\ \mathrm{NX}_{t+h} &= \theta_h^{\mathrm{NX}} \cdot \mathrm{News}_t + X_t \cdot \boldsymbol{\beta}^{\mathrm{NX}} + \varepsilon_{t+h}^{NX} \quad h = 0, 1, ..., 8 \end{aligned}$$

Given that: (i) the IRF is obtained via OLS and the OLS is linear in the outcome and (ii) all equations have the same set of regressors X_t and (iii) GDP is obtained as the exact sum of its four components, it is easy to show that:

$$\hat{\theta}_h^{\mathrm{GDP}} = \hat{\theta}_h^{\mathrm{G}} + \hat{\theta}_h^{\mathrm{C}} + \hat{\theta}_h^{\mathrm{I}} + \hat{\theta}_h^{\mathrm{NX}} \quad \text{for all } h = 0, 1, ..., 8.$$

where the ˆ denotes the OLS estimate. Therefore, we can analytically break down the IRF of GDP to a defense news shock in its four components, plotted in Figure 25.

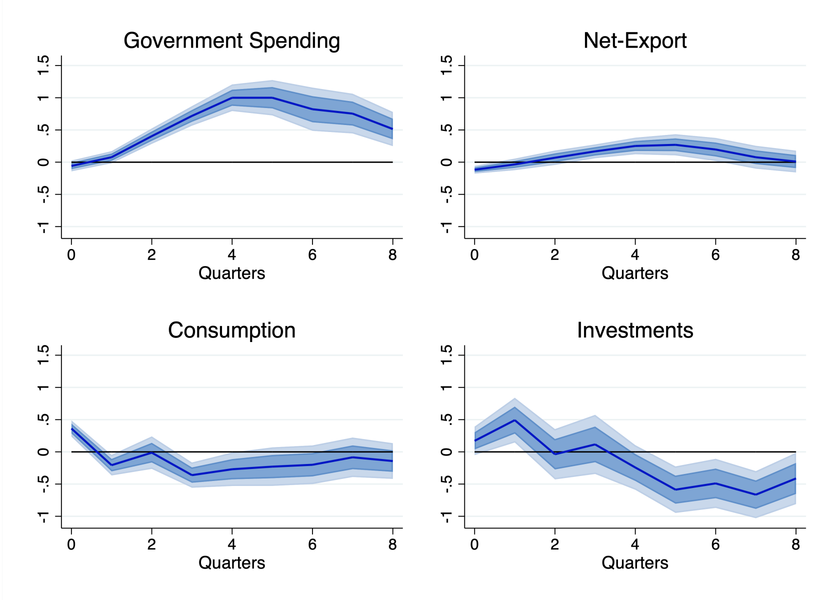
Figure 25 shows that aggregate consumption at horizon 0 and aggregate investment at horizon 1 drive the early increase in GDP after a defense news shock.

Consumption and Investment: Finally, we study what component of consumption and investment drive their early response. We do so by estimating the IRFs using again lag-augmented local projections and the same set of variables used before. However, we now split investment and consumption in two components. Investment is divided into inventories and fixed investment (i.e. residential plus non-residential). Consumption is divided as durable and the summation of non-durable and service. As usual, variables are taken in nominal terms, divided by the GDP price deflator times real potential output (Gordon and Krenn (2010) transformation).

The IRFs of these four components of consumption and investments to a defense news shocks are showed in Figure 26.

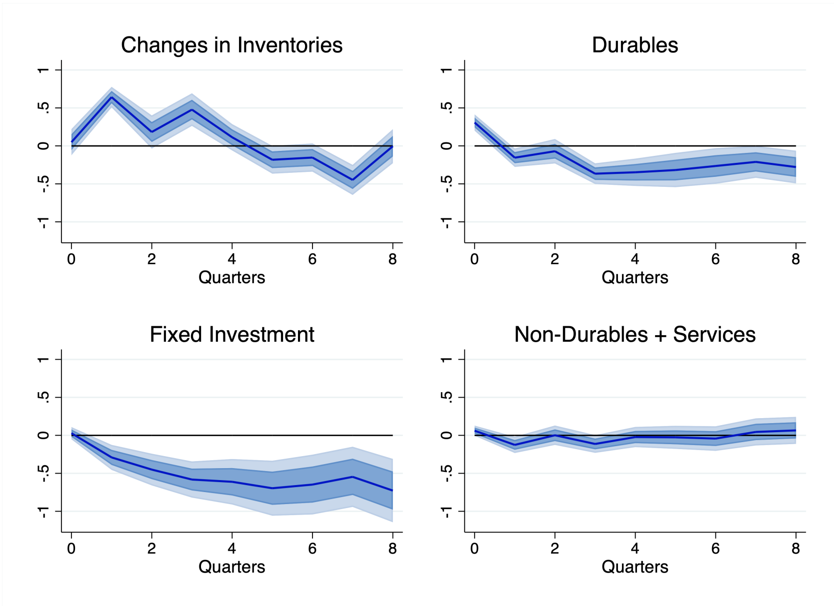
It is clear that the horizon 0 response of consumption originates from durable while the horizon 1 response of investment is driven by inventories.

Figure 25: Response of GDP Components to a Defense News Shock



See notes of Figure 1.

Figure 26: Response of Consumption and Investment to a Defense News Shock



See notes of Figure 1.

E Results without the Korean War

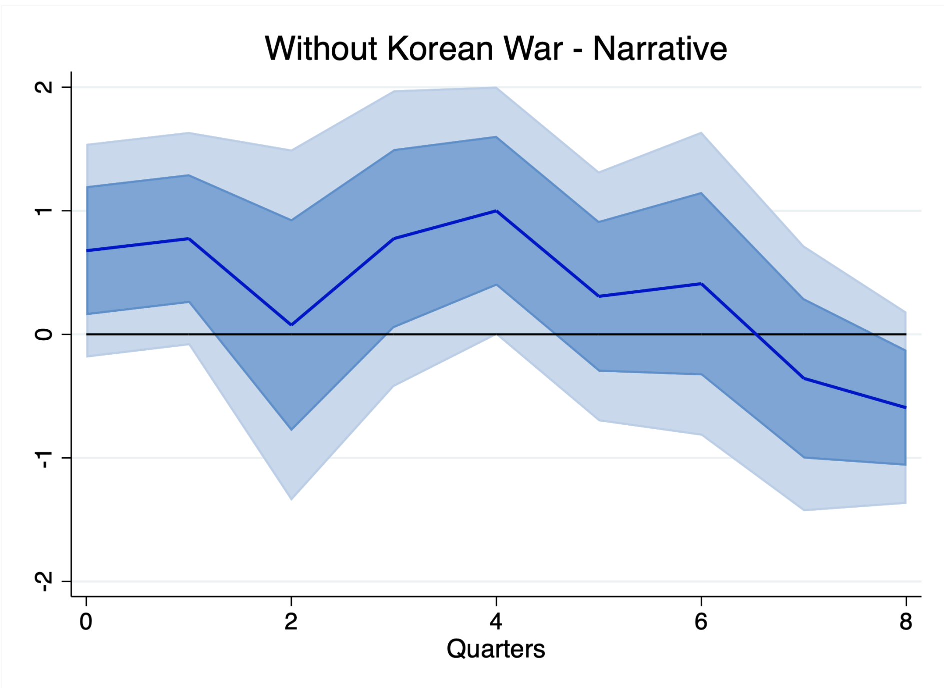


Figure 27: **Response of Inventories to a Defense News Shock**: *Sample goes from 1947Q1 to 2015Q4. All the rest is identical to Figure 5*

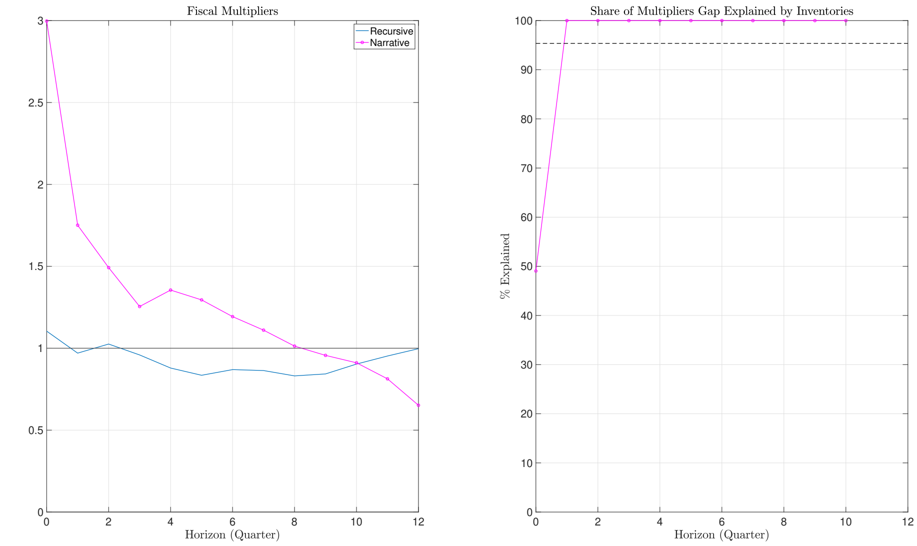


Figure 28: **Cumulative Fiscal Multipliers and Multiplier-Gap (Robustness).** *Sample goes from 1954Q1 to 2015Q4. All the rest is identical to Figure 9.*

F Results for Recursively Identified Shocks

Recursive shocks are constructed by ordering G first in a VAR which uses the same set of variables described to estimate our baseline Equation (1).

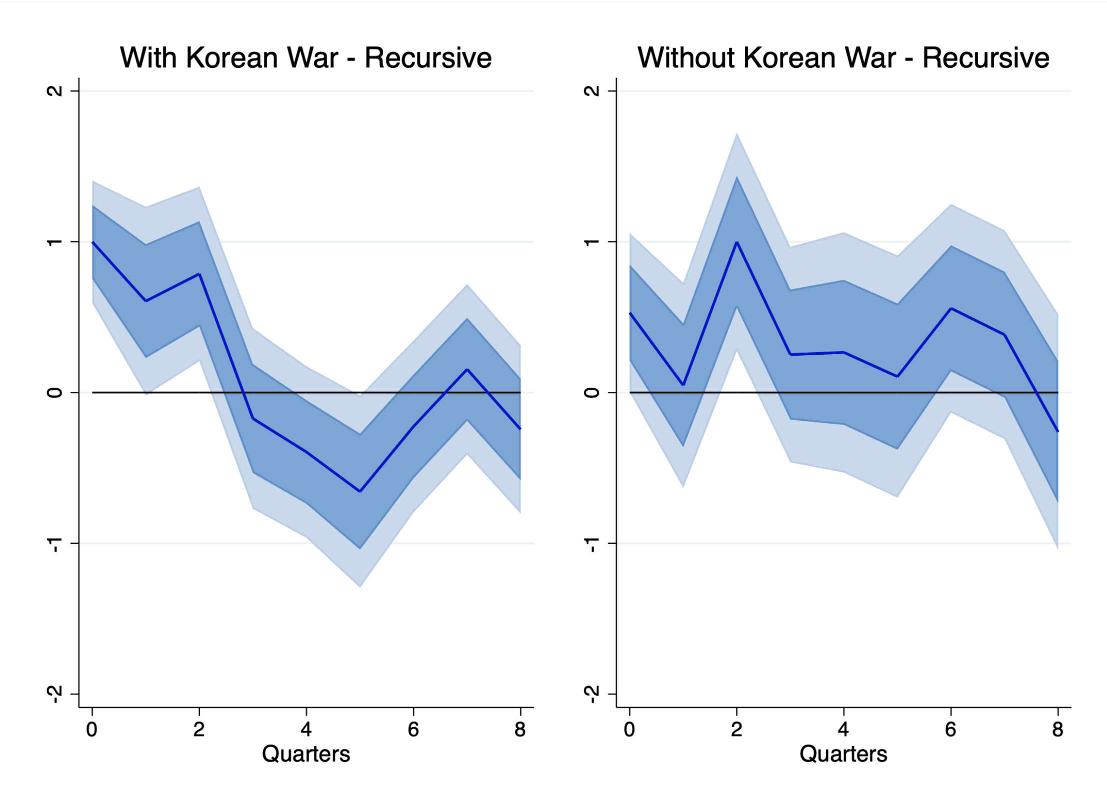


Figure 29: **IRF of Inventories to a Recursive Shock.** *Sample goes from 1948Q1 (1954Q1) to 2015Q4 with (without) the Korean war. All the rest is identical to Figure 2.*

G Obligations Multipliers - Robustness

In this section we show that the cumulative fiscal multipliers obtained by instrumenting the cumulative change in G with defense procurement oblitations, illustrated in Figure 10 os Section 5 are robust to different sub-samples.

Firstly, in light of the importance of the Korean war raised by the fiscal policy literature, we also consider a sample which starts from 1954Q1. Secondly, Eichenbaum and J. Fisher (2005) highlights that fiscal policy in the aftermath of 9-11 was deficit-financed, unlike previous military build-ups which were predominantly financed via taxation. Therefore, we consider a sample which goes from 1947Q1 to 2000Q4.

Results are shown in Figure:

The left panel shows the full-sample results. The middle panel removes the Korean war while the right panel removes the years after the 2000, since government spending in the (see). In all cases we have that the initial response is higher and then the multiplier decreases. This is consistent with the results obtained with defense news shock showed in Figure 9, which reflect the presence of anticipation effects.[45]

[45]In fact, $dG \approx 0$ while $dGDP > 0$ which implies $\mathcal{M} \to \infty$.