

Using Stock Returns to Identify Government Spending Shocks*

Jonas D.M. Fisher
Federal Reserve Bank of Chicago
jfisher@frbchi.org

Ryan Peters
Federal Reserve Bank of Chicago
ryan.peters@chi.frb.org

September 16, 2009

Abstract

This paper explores a new approach to identifying government spending shocks which avoids many of the shortcomings of existing approaches. The new approach is to identify government spending shocks with statistical innovations to the accumulated excess returns of large US military contractors. This strategy is used to estimate the dynamic responses of output, hours, consumption and real wages to a government spending shock. We find that positive government spending shocks are associated with increases in output, hours, and consumption. Real wages initially decline after a government spending shock and then rise after a year. We estimate the government spending multiplier associated with increases in military spending to be about 1.5 over a horizon of 5 years.

Journal of Economic Literature Classification Numbers: E0, E30, E60

Keywords: Government spending shocks, fiscal policy, stock returns

*The views expressed herein are those of the authors and do not necessarily reflect those of the Federal Reserve Bank of Chicago or the Federal Reserve System. Without implicating, we thank Marco Bassetto, Jeff Campbell, Valerie Ramey, David Romer, Chris Sleet, François Velde and Sevin Yeltekin for helpful conversations and comments.

1 Introduction

A classic question in macroeconomics is: How does the economy respond to a change in government spending? The answer to this question clearly is important from a public policy perspective, particularly given current events. It is also important because competing macroeconomic models make different predictions, and hence answering the question is helpful for choosing between these models. The magnitude of the stakes has stoked much research. Yet, there is surprisingly little consensus on what the answer to the question is. The lack of consensus derives from the fact that the two main approaches to identifying government spending shocks lead to different answers and that both have important shortcomings. This paper seeks to avoid the shortcomings of the existing approaches and contribute to the debate in the literature by developing and applying a new measure of government spending shocks based on defense contractor stock returns.

One of the two main ways of identifying government spending shocks, often called the *narrative approach*, was developed by [Ramey and Shapiro \(1997\)](#). These authors argue that there are three dates when persistent increases in US military spending were anticipated. By viewing the events precipitating the spending as exogenous, these dates can be used to identify government spending shocks. [Ramey and Shapiro \(1997\)](#) find that output and hours rise and consumption and real wages fall after a spending shock, supporting the simple neoclassical view of the effects of government spending. These findings have been confirmed in subsequent work by [Burnside, Eichenbaum, and Fisher \(2004\)](#), [Cavallo \(2005\)](#), [Edelberg, Eichenbaum, and Fisher \(1999\)](#) and [Eichenbaum and Fisher \(2005\)](#). The principle advantages of the narrative approach are the likely exogeneity of the spending episodes and that it takes into account the anticipated nature of the government spending. However, there are some clear shortcomings. These include the small number of observations, that the episodes only involve spending increases, the need to assume that the spending is known with certainty, and that the selection of the dates is inherently subjective. Recently [Ramey \(2008\)](#) has expanded the set of dates, mitigating the first two shortcomings, but there is little that can be done about the other two.

The second main way of identifying government spending is associated with [Blanchard and Perotti \(2002\)](#). Essentially, this involves identifying a government spending

shock with the Choleski innovation to government spending in a vector-autoregression (VAR) with government spending ordered first. Many other papers have used this approach, including [Berndt, Lustig, and Yeltekin \(2009\)](#), [Fatas and Mihov \(2001\)](#), [Gali, Lopez Salido, and Valles \(2007\)](#), [Mountford and Uhlig \(2002\)](#), [Perotti \(2002\)](#), and [Perotti \(2007\)](#). The *government spending innovations approach* typically finds that output, hours, consumption and real wages all rise after a positive government spending innovation. These findings are usually interpreted as supporting models where sticky-prices play an important role. The principle advantages of this approach are that there are many observations, both positive and negative shocks, and it is not as subjective as the narrative approach. However, there is one important shortcoming. This is that, as the narrative approach highlights, there are many cases where patterns of government spending are anticipated before they are recorded in the data. [Ramey \(2008\)](#) shows how failure to take into account the anticipated nature of changes in government spending can lead to estimates of the response of consumption that are of the wrong sign.

The identification strategy explored in this paper avoids the shortcomings of both the narrative approach and the government spending innovation approach, at least in principle. The idea is simple, intuitive and builds on the existing approaches. Like the narrative approach, the starting premise is that there are intermittent disturbances to current and expected future US military spending that can be treated as exogenous. In the event of such a disturbance, current and expected future earnings of firms which specialize in producing services and equipment for the military will change. Forward looking agents incorporate expectations of sales into their valuation of the stocks of these firms, thereby affecting the returns to holding these stocks. This suggests it may be useful to use surprises in the returns of military contractors to identify government spending shocks.

Of course, not all variation in military contractor stock returns is due to changing expectations about military spending. Military contractors sell to the private sector and so are influenced by the state of the economy. There is also idiosyncratic variation in the stock of any single contractor due to success or failure in winning particular US or foreign military contracts. These issues are addressed in several ways. First, we identify innovations in government spending with innovations in military contractor stock returns in excess of the market as a whole. We assume these innovations are

orthogonal to variables representing the state of the economy. Second, we use the combined stock returns of multiple contractors. Third, we emphasize the returns of large contractors that specialize in producing for the military. The stock returns of these firms should be less driven by idiosyncratic contracting outcomes compared to small contractors.

While this identification strategy certainly has some drawbacks, which we discuss below, it should avoid the shortcomings of the existing approaches. Like the government spending innovation approach it increases the number of observations, includes increases and decreases in defense spending and is less subjective than the narrative approach. Importantly it is also immune to the Ramey critique of the government spending innovation approach. Finally, unlike the narrative approach it incorporates the inherent uncertainty involved in predicting the future path of defense spending. These advantages suggest that it is worth investigating the implications of using contractor stock returns to address the government spending question.¹

Our main results are based on the stock returns of military contractors that have ever appeared in an annual listing of the top 3 military contractors by value of primary contracts awarded in US government fiscal years from 1958 to 2007. A primary military contract is a contract directly with the military. The value of primary contracts in any given year is an imperfect measure of the defense business of a firm since firms will sub-contract pieces of a primary contract, and they will undertake sub-contract work from primary contracts of other firms. Still, they are a useful indicator of the magnitude of the defense business of any given contractor. To find our group of contractors we have compiled listings of the top 100 primary military contractors from 1958-2007 using a variety of government and private sources. The total sales of contractors ever in the top 3 follow the contours of aggregate defense spending and excess returns of these firms have superior explanatory power for defense spending and government spending as a whole.

We find that after a positive excess return innovation, output, hours and consumption initially are roughly constant for several quarters, and then all rise in a

¹We are not the first to consider stock returns of the defense industry in the context of government spending. [Berndt et al. \(2009\)](#) use the Fama-French “Guns” portfolio to help forecast government spending. We discuss this below. In private correspondence Valerie Ramey describes preliminary research into using defense industry stock returns to predict government spending. We call the portfolio of stocks used by Ramey “Guns+” and discuss it below as well.

hump-shaped pattern. The results for wages are less clear cut. There is an initial decline during the period when the other variables are essentially unchanged. After this initial decline, wages rise persistently over the period when output, hours and consumption are following their hump-shaped paths. We estimate the government spending multiplier associated with increases in military spending to be about 0.6 over a horizon of 5 years. Overall, the results are closer to the government spending innovation based findings than those based on the narrative approach, and they point toward models that depart from the simple neoclassical paradigm.

The remainder of the paper proceeds as follows. The next section describes the dynamics of sales and stock returns of the military contractors which underlie the analysis. In section three we describe and defend our identification strategy. Section four contains the main results, compares them to findings based on the narrative and government spending innovation approaches, and quantifies the contribution of government spending shocks to macroeconomic fluctuations. Section five contains some concluding remarks.

2 The Military Contractor Data

The stock returns used in this analysis are those for firms which specialize in munitions and related equipment, and which are ever ranked in the top 3 in value of primary contracts awarded in a given year. A prerequisite for determining these firms is to compile lists of top contractors. The first sub-section describes how we do this. This is followed by a description of the sales and stock returns of the firms which satisfy the top 3 criterion. The third sub-section briefly relates the excess returns to the military shocks described in [Ramey \(2008\)](#).

2.1 Compiling the Top 3 Military Contractors, 1958-2007

The main source for compiling the top contractor lists is the report “100 Companies Receiving the Largest Dollar Volume of Prime Contract Awards” published by the Department of Defense Directorate for Information Operations and Reports. The report presents summary data on the 100 companies and their subsidiaries receiving the

largest dollar volume of Department of Defense (DoD) prime contract awards during the fiscal year for each year, starting in fiscal year 1958 and continuing until 1995. These reports have been archived at the web site for the Defense Technical Information Center (www.dtic.mil). Unfortunately, not all of these reports are available. The missing years are 1980, 1984-1987. Starting in 1996, the reports are available online from the Department of Defense Personnel and Procurement Statistics web site (siadapp.dmdc.osd.mil). The reports have been produced from individual prime contracts, specifically DoD form DD350. Electronic copies of these forms are also available from the Department of Defense Personnel and Procurement Statistics.

For the missing years 1984-1987 we used lists published in *Defense Magazine*. Each year from 1983 to 1996 this publication contains an annual supplement documenting the structure and finances of the DoD. This supplement includes an extract of the “100 Largest Companies” report with just the parent company name and the volume of prime contracts for the company and its subsidiaries. This is enough information for our purposes.

For the missing year 1980 there is no summary data available. Fortunately, we can go to the primary source – the DD350 forms originally used to produce the “100 Largest Companies” reports. The 1981 report has a list of companies that have dropped off of the top 100 list since 1980 as well as a list of firms that are on the list in 1981 but not 1980. Using these lists, it is trivial to construct the top 100 list for 1980. Filling out the procurement numbers is only a matter of referencing the subsidiary lists from 1979 and 1981 and adding up sales from all subsidiaries. There is a chance that some subsidiaries have changed over the course of a year, but generally subsidiaries do not change that often and are not that large a share of total military sales.

On the top 100 lists are companies for which military contracts are not the main source of revenue. For example, AT&T was a major military contractor for many years, but this was a small portion of its business. We used two procedures to cull firms from the list to avoid including firms whose stocks returns may be strongly influenced by idiosyncratic industry developments with nothing to do with military spending. First, we study just those firms in the lists which have a clear military

focus as determined by their SIC code.² In practice this has little impact on the top 3, eliminating just one firm, an automobile parts manufacturer named Tenneco. Second, we exclude Boeing from the list. About half of Boeing’s business is commercial aircraft and this business has its own low frequency movements that have a significant impact on its sales and stock price, reducing its predictive power for future government spending.³

The list of companies that appear in the top 3 of our refined listings for at least one year, along with their SIC industry classification, and the years they appear in the top 3, is reported in Table 1. The table indicates that fourteen companies have been in the top 3 at one time or another. Due to mergers over the years these fourteen companies are currently just five. Martin and Marietta (which is not listed in the table) merged in 1961, and then the combined company merged with Lockheed in 1995. Grumman and Northrop merged in 1994. McDonnell Douglas merged with Boeing (and so drops from our top 3) in 1996. We refer to the list of firms in Table 1 as the Top 3 firms.

A significant fraction of revenues for all these firms derives from military contracting, but, as the SIC descriptions suggest, they are also involved in the US space program. Indeed many of these firms also appear at the top of listings of the value of space contracts awarded. For example, Lockheed Martin, Raytheon and Northrop Grumman are second, third and fourth on a listing of the top 50 space companies by sales in 2002 published by *Space News* (Boeing is first on this list.) This is convenient given that Ramey (2008) has used spending on the space program to augment the original Ramey-Shapiro war dates. So the stock returns of these companies should reflect anticipated space program spending as well and can be compared to the “space dates” compiled by Ramey (2008).

² The SIC codes we use are: 3482 – Small Arms Ammunition, 3483 – Ammunition, Except for Small Arms, 3484 – Small Arms, 3489 – Ordnance and Accessories, Not Elsewhere Classified (NEC), 3720 – Aircraft and Parts, 3721 – Aircraft, 3724 – Aircraft Engines and Engine Parts, 3728 – Aircraft Parts and Auxiliary Equipment NEC, 3730 – Ship and Boat Building and Repairing, 3760 – Guided Missiles and Space Vehicles and Parts, 3761 – Guided Missiles and Space Vehicles, 3764 – Guided Missile and Space Vehicle Propulsion Units and Propulsion Unit Parts, 3769 – Guided Missile Space Vehicle Parts and Auxiliary Equipment NEC, 3795 – Tanks and tank components, and 3812 – Search, Detection, Navigation, Guidance, Aeronautical Systems.

³United Technologies also has a large non-defense business. However, this business is closely aligned with the overall state of the economy which is accounted for by focusing on excess returns.

2.2 Top 3 Contractor Sales and Stock Returns

For the stock returns of our list of military contractors to be credible indicators of *total* military spending their sales should line up well with the total. To examine whether this is the case, consider Figure 1. This figure displays linearly detrended real military spending and three measures of defense industry sales. These measures are for ‘Top 3’, the firms in Table 1, ‘Guns,’ the firms in a subset of the industries listed in footnote 2, and ‘Guns+’, the firms in all the industries listed in footnote 2. The excluded industries in ‘Guns’ are 3720-3728, 3795 and 3812. We have included the sales of Guns and Guns+ since they are plausible alternative portfolios to identify government spending shocks with. The Guns category corresponds to the Fama-French definition of the ‘Guns industry’ in their 48 industry portfolio. [Berndt et al. \(2009\)](#) have used excess returns of these firms to help forecast future government spending. The Guns+ variable includes additional defense-like industries and seems to be a plausible perturbation to Guns. We obtain the sales data from COMPUSTAT.

Figure 1 indicates that Top 3 sales lines up very well with the main undulations of military spending, but Guns and Guns+ sales do less well. The contemporaneous correlation between sales and spending is .92, .34 and .43 for Top 3, Guns and Guns+ sales, respectively. The relatively poor alignment with military spending of Guns and Guns+ is because using SIC codes alone to select a defense industry portfolio leads to including firms having non-military businesses that dominate the overall dynamics of the firm. For example, Guns includes Brunswick, a large maker of recreational products including boats, and Taser, a maker of personal security devices. The very high degree of co-movement with military spending of Top 3 sales emerges because the military businesses dominate the dynamics of these firms. This strongly suggests stock returns for the Top 3 portfolio will be a better indicator of expectations about defense spending compared to stock returns for Guns or Guns+.

Now we describe the stock return series we use. Here we focus on the stock returns of the Top 3 firms. Our measure of contractor stock returns is simply the market-value-weighted sum of total returns on the individual stocks. By total returns we mean the return from both capital gains and dividends. We use the market value at the beginning of the quarter to weight monthly holding period returns and accumulate these returns over each quarter. Our overall market return indicator is the market-

value-weighted sum of total returns of all publicly traded firms. The source for these data is CRSP. We define *excess returns* as the difference between the Top 3 returns and the overall market return. Excess returns are very noisy and it is difficult to discern low frequency movements with them. Consequently, we focus on *accumulated* excess returns.

Before discussing the returns of the Top 3 as a whole, it is worth briefly considering the returns of the individual firms. These are displayed in Figure 2 along with vertical lines indicating the Ramey-Shapiro war dates. The Ramey-Shapiro war dates correspond to the Korean War (1950:3), the escalation of the Vietnam War (1965:1), and the Carter-Reagan military build-up (1980:1). There is also a vertical line corresponding to the quarter of the 9/11 terrorist attacks, which has been considered as a “war date” by Eichenbaum and Fisher (2005) and Ramey (2008). Accumulated excess returns are normalized to unity in the first year the firm is listed in CRSP so that values above unity reflect holding period returns in excess of the market from the first date of returns.

Two observations on Figure 2 are worth making. First, the dynamics of the individual returns are quite different, reflecting the idiosyncratic nature of defense contracting and the differences in behavior of the firms’ other business lines. Second, some of the firms make large excess returns over the sample. General Dynamics has almost always fluctuated above break-even returns, and since 1990 has done exceptionally well, so that the return from holding General Dynamics stock from 1947 to 2007 is more than 8 times the market. Northrop-Grumman has also done very well, with holding period returns exceeding 10 over many periods. Most of the other returns also do well.

Figure 3 displays accumulated excess returns of the Top 3 portfolio along with real military spending and the war dates. Accumulated excess returns are normalized to unity in 1947:1. The returns prior to 1958 are calculated using the list of firms in Table 1 since we have no information prior to 1958 to add to the list. We make four points about Figure 1.

First, as anticipated by Figure 2, the accumulated excess returns for the portfolio of Top 3 firms are large. Had one followed the investment strategy of buy and hold on the stocks of the firms in Table 1 starting in 1947:1 the accumulated excess return

would have been more than five times the overall market return by the end of 2007.⁴

Second, the excess returns closely match the Vietnam War date and the 9/11 date, but not the Korean War or Carter-Reagan dates. In the case of the Korean War we see that military spending increased sixfold between 1950 and 1953. Yet it was not until 1953 that excess return started to accumulate significantly. This delayed response of Top 3 stock returns is not puzzling after one realizes that, as was the case in World Wars I and II, excess profits tax legislation was enacted during the Korean War. This legislation was effective from July 1 1950 to December 31 1953. To the extent that “excess profits” were accurately determined by the legislation, the absence of significant excess returns is understandable. This kind of legislation has not been enacted since the Korean War, although there was an attempt during the first Iraq War.

Third, movements in excess returns correctly forecast the persistent declines in spending at the end of the Vietnam War and the end of the Carter-Reagan build-up. It also appears that agents were anticipating a persistent rise in military spending before the Carter-Reagan build-up was fully underway.

Fourth, there are some large movements in excess returns which are not followed by future changes in spending which would rationalize the movements. The best example of this is dynamics of excess returns in the 1990s and early 2000s. Following the first Iraq war, agents expected increases in military spending due to the success of certain weapons systems during that war. This counteracted expectations of diminished spending following the end of the Cold War. When the new spending failed to materialize there was a sharp drop in excess returns in the latter part of the 1990s, reflecting a general perception that post Cold War military spending would not rise to levels seen during the Reagan years. Anticipation of military spending under a George W. Bush administration partially offset these declines in returns, before 9/11 and then the second Iraq War led to widespread expectations of increased military spending that are reflected in large excess returns.

The last two points highlight key benefits, relative to the narrative approach, of

⁴A regression of the Top 3 returns on a constant and the overall market return over the sample 1947:1 2007:4 yields a constant equal to .012 (robust standard error .006) and slope coefficient equal to .98 (.09). Over the sample period beginning in 1957:3 we focus on below, the estimates are .009 (.006) and .95 (.09).

using military contractor stock returns to measure expectations of future military spending. The stock returns capture expectations of declines in spending, which the original narrative approach does not. They also incorporate unrealized expectations, which are largely ignored in the narrative approach.

2.3 Excess Returns and Ramey’s Military Shocks

To gain more insight into how contractor excess returns and the narrative approach are related, Figure 4 displays accumulated Top 3 excess returns along with the new military shock series described in Ramey (2008). This series is the result of a careful analysis of the historical record to determine when revisions to expectations of future military and space program spending may have occurred, and by how much those expectations changed. The series is zero on all dates except those determined to correspond to revisions in expectations. The series on these dates equals the expected present value of the revision in spending scaled by the nominal value of government spending in the previous period. Each plot in the figure displays a window of dates over the sample 1947-2007 where the windows encompass the entire sample.

The top-left plot shows the failure of excess returns to respond to the onset of the Korean War we have already discussed. The remainder of the plots indicate there are few examples where the stock returns obviously reflect the military shocks compiled in Ramey (2008) and in no case is there the sharp, persistent change in stock returns expected with discrete revisions to expectations. One case where the stock returns match the shock is 1980:1 (top-right), when excess returns and the shock series both spike. Still, there were movements in stock returns prior to 1980:1 which suggest the information that military spending would be rising was revealed over several quarters. In addition returns fall in the quarter after 1980:1. This suggests there was uncertainty about the magnitude of the future spending. There are spikes in returns on the dates corresponding to the two space program shocks in the Ramey (2008) series, 1957:4 (Sputnik) and 1961:2 (Moon mission), middle-left plot. But, these are also followed by subsequent declines, and, as evident in the other plots as well, the variation in the excess returns is just as large on dates where the shock series is zero as it is when the shock series is non-zero. In many cases, for example the two middle plots, excess returns do not qualitatively match the magnitude of the military shock

- smaller shocks have similar or greater excess returns associated with them than nearby larger shocks. Finally, there are examples where the timing in the shock series appears to be off. In the top-right plot the shock in 1981:1 appears to be anticipated by a quarter. In the lower-right plot the 2001:3 shock appears to be one quarter too early.

The overall impression from Figure 4 is that the shock series and the excess returns do not correspond well. This could reflect that the excess returns are a poor measure of expected future military spending. In our view it reflects the fact that the process of information revelation by which expectations about future military spending evolve is more nuanced than the stark determination indicative of the narrative approach. This is not to say one approach is definitively superior to the other, but that there is a strong case to be made for considering using excess returns to identify government spending shocks.

3 Identifying Government Spending Shocks with Stock Return Data

In this section we build a case for using the stock returns of firms on our top 3 military contractor list to identify government spending shocks and then describe how we go about doing so.

3.1 The Explanatory Power of Top 3 Excess Returns

A key test of whether our excess returns measure is useful for identifying government spending shocks is its explanatory power for government spending and other macro variables. Our procedure for evaluating explanatory power follows [Ramey \(2008\)](#). In particular, we regress variables to be predicted on current and lagged values of the indicator variable in question. Table 2 reports R^2 for these regressions using five indicator variables to predict military spending, overall government spending, GDP and non-durable and services consumption. The top panel is based on the sample 1948:1–2007:4 and the second panel on the sample 1957:3–2007:4, which corresponds to the sample we have data for compiling our Top 3 list. All the predicted variables

are log first differences after first dividing by population.

The five indicator variables include three log accumulated excess return series and two narrative based series. The excess return series correspond to portfolios of Top 3, Guns and Guns+ firms as defined above. We include the additional excess return series because they are natural stock return based alternatives to Top 3. The narrative series are the Ramey-Shapiro war dates plus the 9/11 date, ‘War Dates,’ and the new military shock series introduced in [Ramey \(2008\)](#), ‘Shocks.’

Consider the top panel of Table 2. This shows the clear superiority in explanatory power of the narrative variables. The explanatory power of the Shock series for military and overall government spending is particularly striking. The stock return variables perform substantially worse over this sample period. This performance was anticipated by our discussion of the Korean War episode and the role played by the excess profits tax during that time.

In the bottom panel of Table 2 we see that the explanatory power of Top 3, Guns, and Guns+ for military and total government spending is greater in the later sub-sample. It is substantially greater for Top 3 and Guns, and Top 3 is has somewhat better explanatory power for military spending compared to Guns+. There is only a slight improvement in the explanatory power for Guns. All three stock return indicators show modest improvements in explanatory power for the other variables in the table. The narrative series perform much worse once the Korean War episode is excluded. Both series have very poor explanatory power for military and overall government spending compared to the stock return indicators. The explanatory power for output and consumption of the war date variable is somewhat better than the stock variables, but it is worse for Shocks.

We conclude that for military and total government spending, Top 3 and Guns+ are clearly superior to Guns and the two narrative variables over the sample period we think the stock return series have a chance of being useful for identifying shocks to government spending. The Top 3 returns are marginally better than the Guns+ returns at explaining the government spending variables. This motivates our focus on the Top 3 returns for identifying government spending shocks.

3.2 Identification of Government Spending Shocks

We identify government spending shocks with the Choleski innovation to the log of the accumulated Top 3 excess return series in VARs including several macroeconomic variables representing the state of the economy, where the excess return variable is ordered last. Thus innovations to government spending are identified with innovations to Top 3 excess returns which are orthogonal to the state of the economy. Assuming that the macroeconomic variables in the VARs are a good representation of the state of the economy, this strategy should identify shocks to government spending well if (i) technological progress in production (costs) at the Top 3 firms evolves in the same way as in the rest of the economy, (ii) Top 3 mark-ups do not behave differently than in the rest of the economy, (iii) variation in sales of the Top 3 firms are dominated by shocks to defense spending. The third condition is verified by Figure 1.

The second two conditions are more questionable. One concern about condition (i) may be that the Top 3 firms are largely involved in producing capital equipment. Since relative prices of capital equipment have been declining over the sample, this suggests the production technology in the defense industry may be improving at a faster rate compared to the rest of economy. Condition (ii) might also be viewed as problematic given the large excess returns of the Top 3 and of political considerations which may have boosted profits in the defense industry. We address these concerns about (i) and (ii) by including linear trends in the VAR. By doing this we are essentially hoping that any remaining idiosyncratic variation in stock returns is overwhelmingly dominated by (exogenous) variation in demand for the military products of the Top 3 firms, rather than idiosyncratic technology or mark-up shocks.

We use quarterly data. The first left-hand-side observation in our VARs is 1959:1 and the last observation is 2007:4. In the VARs discussed below we use six lags. So the first lagged observation is 1957:3. This corresponds to the first quarter of the fiscal year 1958, the first year we have a listing of the top 100 primary military contractors. The starting date for the analysis represents a departure from the literature and so deserves some discussion. Previous work uses 1947:1 or 1948:1 as the first quarter of the analysis. We use the later starting date for several reasons. First, the quarter 1957:3 is the first for which we have information on the top 3 military contractors. Second, this date excludes the Korean War episode. We think this is ap-

appropriate given our previous discussion of the use of the excess profits tax during this war. We also think it is appropriate to exclude the Korean War episode because the increase in military spending associated with this episode appears to be a permanent shock. Military spending increases more than sixfold during the Korean War and is maintained at elevated levels thereafter. Later waves of per capita military spending appear to be temporary fluctuations around a permanently higher level. It therefore seems worthwhile to assess the impact of excluding the Korean War and focusing on a later sample.

A related argument for beginning the sample in the late 1950s is that this appears to be a watershed period in the economic history of the US. Essentially, it marks the beginning of the modern defense industry. President Eisenhower makes this point in his farewell address on January 17, 1961:

Our military organization today bears little relation to that known by any of my predecessors in peacetime, or indeed by the fighting men of World War II or Korea. Until the latest of our world conflicts, the United States had no armaments industry. American makers of plowshares could, with time and as required, make swords as well. But now we can no longer risk emergency improvisation of national defense; we have been compelled to create a permanent armaments industry of vast proportions.

From this perspective, it seems difficult to make the case to begin the sample any earlier than the late 1950s, if one wants to identify government spending shocks using variables related to military spending.

4 New Estimates of the Economy's Response to a Government Spending Shock

This section describes our main findings. We describe the responses of output, consumption, hours and wages to an identified government spending shock. After this we quantify the contribution of our identified shocks to macroeconomic fluctuations. Finally, we relate our estimated responses to those obtained using alternative identifications.

4.1 Main Findings

The dynamic responses to the shocks are estimated in VARs which, in addition to our Top 3 accumulated excess returns variable, include logs of per capita military spending and GDP, the log of the nominal gross three month treasury bill rate, plus one additional variable we wish to estimate the response of. The additional variable is one of the following: the logs of per capita hours worked, chain weighted consumption of nondurables and services, and total government spending, and real product wages. The population and hours measures are the ones described in [Francis and Ramey \(2008\)](#). The population measure is adjusted for changes in the demographic composition of the workforce and the hours measure is for the economy as a whole. The wage measure is private business wages deflated by the price deflator for the private business sector.⁵ The VARs include six lags and a linear time trend. We use six lags because the impulse responses are more precisely estimated compared to cases with fewer lags. The qualitative findings are similar with four or five lags. We include linear trends since that is the usual practise in the literature. Similar results also are obtained including variables in first differences. The qualitative nature of the results are also robust to including inflation and taxes in the VAR.

Figure 5 and 6 display our main findings. The blue lines are point estimates of responses to one standard deviation innovations to the accumulated excess return variable, that is our measure of government spending shocks. The green lines demarcate 68% posterior probability regions which contain the true responses with approximately 2/3 probability. This is the criteria advocated by [Sims and Zha \(1999\)](#) for assessing the plausibility of estimated impulse response functions. The units of the responses are percent deviations from the un-shocked path of the variable.

The responses in Figure 5 confirm our finding from Table 2 that the excess returns variable is a good forecaster of future military and total government spending. If this were not true we would have little confidence that the estimated responses for the

⁵The population and hours series were obtained from Valerie Ramey's website. The remaining variables were obtained from the Haver Analytics Database. The series names are: real military spending – GFDH, real total government spending – GH, real output – GDPH, nominal nondurable consumption – CN, real nondurable consumption, – CNH, nominal services consumption – CS, real services consumption – CSH, compensation in private business – LXBC, and deflator for private business – LXBI.

other variables reflect the effects of government spending shocks. A one standard deviation shock to excess returns is associated with hump-shaped positive spending responses, peaking after about three years. The peak response of military spending is about 2% and of total government spending is about 3/4%.

Figure 6 indicates that the hump-shaped government spending paths are associated with delayed hump-shaped positive responses of output, hours and consumption. In all three cases it takes about a little over a year before economic activity starts to rise. The response of wages is quite different. During the period when the activity variables are essentially unchanged, real wages drop by about .2%. Just prior to the expansion in activity, wages start to rise until after about two years the response is positive for the remainder of the response horizon of five years. However, given the probability bands encompass zero after six quarters, it is hard to rule out that the real wage returns to its pre-shock path, rather than rising above it.

4.2 Quantifying the Effects of Government Spending Shocks

There are three ways we quantify the effects of our identified government spending shocks. We focus on output, but the magnitudes are similar for consumption and hours. The first approach is to consider government spending multipliers. For this we calculate the accumulated responses of military spending, total government spending and output over the five year horizon used for the impulse responses. These are 30%, 10% and 3%, respectively. The average values over our estimation sample of military and total government spending's nominal shares of output are 6.7% and 20%. It follows that the military spending multiplier is about $(.03/.3) \times (1/.067) = 1.5$. That is, one dollar of extra military spending translates to 1.5 dollars of output. Using total government spending the multiplier is the same.

The second way we quantify the effects of government spending is to consider forecast error decompositions. Since the responses in Figure 6 are relatively small, the identified spending shocks explain under 10% of the forecast error variance of output, hours and consumption over a 5 year horizon.

Still, there is the possibility that spending shocks have larger effects in particular periods. The third way we quantify the effects of spending shocks is to consider

historical decompositions during the three war date episodes in our sample, the Vietnam War, the Carter-Reagan build-up and post 9/11. Figure 7 displays the actual log levels of total government spending and output and the values of these variables predicted by assuming only the estimated excess return shocks are realized. In both cases the paths predicted as of the information one period prior to the last date before the plots begin have been subtracted out.

Figure 7 shows that during the Vietnam War the contribution of government spending to the deviation from output was quite substantial for about three years. During the “growth recession” of 1967 it appears as if government spending prevented the economy from entering a full-blown recession. The impact of the Carter-Reagan build-up was relatively small, raising the level of output by a maximum of just .68%, as output was tumbling by over 5% relative to trend. After 9/11 the effects of the increase in spending appear more significant, having a noticeable impact during the nascent recovery in 2003 and 2004.

4.3 Comparison with Alternative Identification Strategies

We now address how our findings compare to those obtained using the main alternative identification strategies. Since our sample period begins later than is typical in the literature, we need to re-estimate the impulse responses of interest with the alternative identifications. We focus on the original war date and statistical innovation based approaches, and omit from the analysis estimates based on the military shock series created by Ramey (2008). With our sample period we were unable to obtain estimates of the responses of output and hours which are significantly different from zero with the military shock series. For the two alternative identification strategies we do consider, we study specifications we view to be representative of how these approaches have been applied in the literature.

For the narrative approach, we use the estimation procedure described in Eichenbaum and Fisher (2005). They estimate VARs with current and lagged values of a dummy variable equal to zero everywhere except the war dates. Impulse responses are identified as the simulated response of the system to the onset of a war episode. Eichenbaum and Fisher (2005) weight the war dates by their eventual magnitude. We assign a value of 1 for the Vietnam War date and 1/3 to the Carter-Reagan date and

the 9/11 date. These values are consistent with the ones considered by [Eichenbaum and Fisher \(2005\)](#) and [Ramey \(2008\)](#). The variables in the VAR are the same as in our baseline specification, except we drop the excess return variable. As in recent applications of this approach, for example [Perotti \(2007\)](#), we use four lags of the endogenous variables and six lags of the war dummy.

For the government spending innovation based approach we consider a single VAR ordered as total government spending, output, consumption, hours, wages and the interest rate. We identify the government spending shock with the Choleski innovation to government spending. As in recent applications of this approach we estimate the VAR using four lags, c.f. [Perotti \(2007\)](#) and [Ramey \(2008\)](#).⁶

The results we obtain are displayed in Figures 8 and 9. These are formatted in the same way as Figures 5 and 6. The striking feature of these figures is how closely the responses resemble our baseline estimates, at least qualitatively. Output, hours and consumption all rise, after a delay. Wages initially fall during the delay in the responses of the quantity variables and rise thereafter, eventually turning positive. The main departure from the baseline is that the decline in wages is very short-lived with the government spending innovation approach. Another difference is the timing of the delay in the sustained increases in output, consumption and hours. Still, when viewed overall these estimates line up quite well with our baseline responses.

The findings for the innovation-based approach are not that different from previous results reported in the literature. So these findings do not appear to be sensitive to the inclusion of the Korean War period. However, the results for the narrative approach are different from those emphasized in the literature. The narrative approach is usually associated with declines in both consumption and wages. An obvious question is whether our particular implementation of the narrative approach underlies the differences in our findings. It does not. We have confirmed that if we extend our sample back to 1948 then consumption and wages fall.⁷

That our findings can be reconciled with the narrative approach suggests that

⁶Unlike our baseline VARs, including six lags of the endogenous variables with the alternative identifications lowers the precision of the estimates. The qualitative features of the responses are similar with six lags.

⁷We are not the first to notice this sensitivity to sample period selection. [Perotti \(2007\)](#) finds the sign of the response of consumption under the narrative approach depends on whether the Korean War date is in the sample.

the critique of the government spending innovation approach made by [Ramey \(2008\)](#) may not be serious in practise. This is consistent with recent work by [Ravn and Mertens \(2009\)](#). These authors develop an empirical procedure for handling anticipated government spending shocks in a VAR setting and show that results based on the government spending innovation identification strategy are not substantially affected by implementing their procedure.

5 Conclusion

In this paper we have introduced a new way to measure shocks to government spending and have used it to revisit the question of how the US economy responds to such a shock. We have emphasized that our new measure, based on the stock returns of large military contractors, avoids the main shortcomings of the available alternatives. Our new measure predicts that, after a delay, output, hours and consumption all rise following an identified positive shock to government spending. Real wages fall during the period of delayed response of the quantity variables and then start to rise. We have confirmed the most widely used alternative approaches to identifying government spending shocks lead to quite similar responses, once the sample period is restricted to exclude the Korean War.

We are encouraged by the fact that the three different approaches, which involve different identifying assumptions, lead to similar answers once the Korean War is excluded from the sample, since we think there are good reasons for doing so. Therefore, it seems reasonable to conclude that the weight of the evidence is starting to point to the view that the response of the economy to a government spending shock is inconsistent with the simple neoclassical growth model. The exact set of assumptions needed to reconcile theory with findings such as those reported in this paper is an important task for future research.

We conclude by emphasizing the kind of lesson we think readers should take from an empirical analysis of government spending shocks which focuses on military spending. As we emphasize in the introduction, there are two reasons to be interested in the findings reported here. One reason is that it is interesting from a public policy perspective to know how government spending, including military spending, affects

the economy. The second reason is that knowledge of the response of the economy to a particular kind of shock is useful for distinguishing between alternative economic models.

In our view, the main advantage of focusing on military spending is that it addresses a very specific kind of fiscal experiment that is easy to replicate in a model: an exogenous change in government spending where that spending is a pure drain on production. While advantageous in this respect, it does mean the analysis is probably not useful for understanding other, policy relevant, kinds of government spending shocks. So it is important to keep in mind that the empirical findings of this paper should not be used to understand how the economy responds to, among other possibilities, shocks to spending on transfers, education, health, or infrastructure.

References

- Berndt, A., H. N. Lustig, and S. Yeltekin (2009). Does the us government hedge against fiscal risk? Carnegie-Mellon University manuscript.
- Blanchard, O. and R. Perotti (2002). An empirical characterization of the dynamic effects of changes in government spending and taxes on output. *Quarterly Journal of Economics* November, 1329–1368.
- Burnside, C., M. Eichenbaum, and J. D. Fisher (2004). Fiscal shocks and their consequences. *Journal of Economic Theory* 115, 89–117.
- Cavallo, M. (2005). Government employment expenditure and the effects of fiscal policy shocks. Federal Reserve Bank of Chicago Working Paper 2005-16.
- Edelberg, W., M. Eichenbaum, and J. D. Fisher (1999). Understanding the effects of a shock to government purchases. *Review of Economic Dynamics* 2(1), 166–206.
- Eichenbaum, M. and J. D. Fisher (2005). Fiscal policy in the aftermath of 9/11. *Journal of Money, Credit and Banking* 37(1), 1–22.
- Fatas, A. and I. Mihov (2001). The effects of fiscal policy on consumption and employment: Theory and evidence. CEPR Discussion Paper No. 2760.
- Francis, N. and V. A. Ramey (2008). A century of work and leisure. *American Economic Journal: Macroeconomics*. Forthcoming.
- Gali, J., D. Lopez Salido, and J. Valles (2007). Understanding the effects of government spending on consumption. *Journal of the European Economic Association* 5, 227–270.
- Mountford, A. and H. Uhlig (2002). What are the effects of fiscal policy shocks. CEPR Discussion Paper No. 3380.
- Perotti, R. (2002). Estimating the effects of fiscal policy in oecd countries. CEPR Discussion Paper No. 3380.
- Perotti, R. (2007). In search of the transmission mechanism of fiscal policy. *NBER Macroeconomics Annual*.

- Ramey, V. (2008). Identifying government spending shocks: It's all in the timing. University of California, San Diego Manuscript.
- Ramey, V. A. and M. Shapiro (1997). Costly capital reallocation and the effects of government spending. *Carnegie Rochester Conference on Public Policy*.
- Ravn, M. O. and K. Mertens (2009). Anticipation of fiscal policy. University of Southampto manuscript.
- Sims, C. A. and T. Zha (1999). Error bands for impulse responses. *Econometrica* 67(5).

Table 1: Firms that are Ever Among the Top 3 US Military Contractors

Primary Contractor	SIC Industry	Years in Top 3
General Dynamics	Ship and Boat Building, Repairing	1958-1962, 1965-1971, 1974-1974, 1978-1992
Grumman	Aircraft	1970-1972
Lockheed Aircraft, Lockheed (1977-1994), Lockheed Martin (1995-)	Guided Missiles, Space Vehicle, Parts	1958-1977, 1983-1987, 1992-2007
Martin, Martin Marietta (1961-1993)	Guided Missiles, Space Vehicle, Parts	1959-1962, 1993
McDonnell Aircraft, McDonnell Douglas (1966-1996)	Aircraft	1963-1964, 1966, 1968-1969, 1971-1996
Northrop, Northrop Grumman (1994-)	Search, Detection, Navigation, Guidance, Aeronautical Systems	1975, 1992, 1994-2007
Raytheon	Search, Detection, Navigation, Guidance Aeronautical Systems	1988-1991, 1996, 1998-2000
United Aircraft, United Technologies (1969-)	Aircraft, Parts	1958, 1966-1968, 1975-1982

Table 2: Explanatory Power (R^2) of Indicators of Government Spending

Variable	Top 3	Guns	Guns+	War Dates	Shocks
1948:1–2007:4					
Military Spending	.07	.05	.06	.26	.58
Government Spending	.08	.02	.03	.24	.56
Output	.03	.03	.03	.08	.07
Consumption	.05	.06	.06	.08	.12
1957:3–2007:4					
Military Spending	.21	.07	.17	.06	.09
Government Spending	.11	.04	.11	.02	.06
Output	.08	.06	.06	.09	.03
Consumption	.05	.06	.05	.09	.03

Note: The reported R^2 's are based on regressions of the predicted variable on the current value and six lags of the indicator variable. Variables are log first differences after first dividing by population. The first three indicators are log accumulated excess returns. 'Top 3' corresponds to the firms in Table 1, 'Guns' corresponds to firms with SIC codes included in footnote 2 excluding 3720-3728, 3795 and 3812, and 'Guns+' is the returns of all the firms with SIC codes listed in footnote 2. The war dates are the original Ramey-Shapiro dates plus the 9/11 date. The 'Shocks' is the Ramey (2008) military shocks variable described in the text.

Figure 1: Detrended Defense Industry Sales and Military Spending, 1958-2007

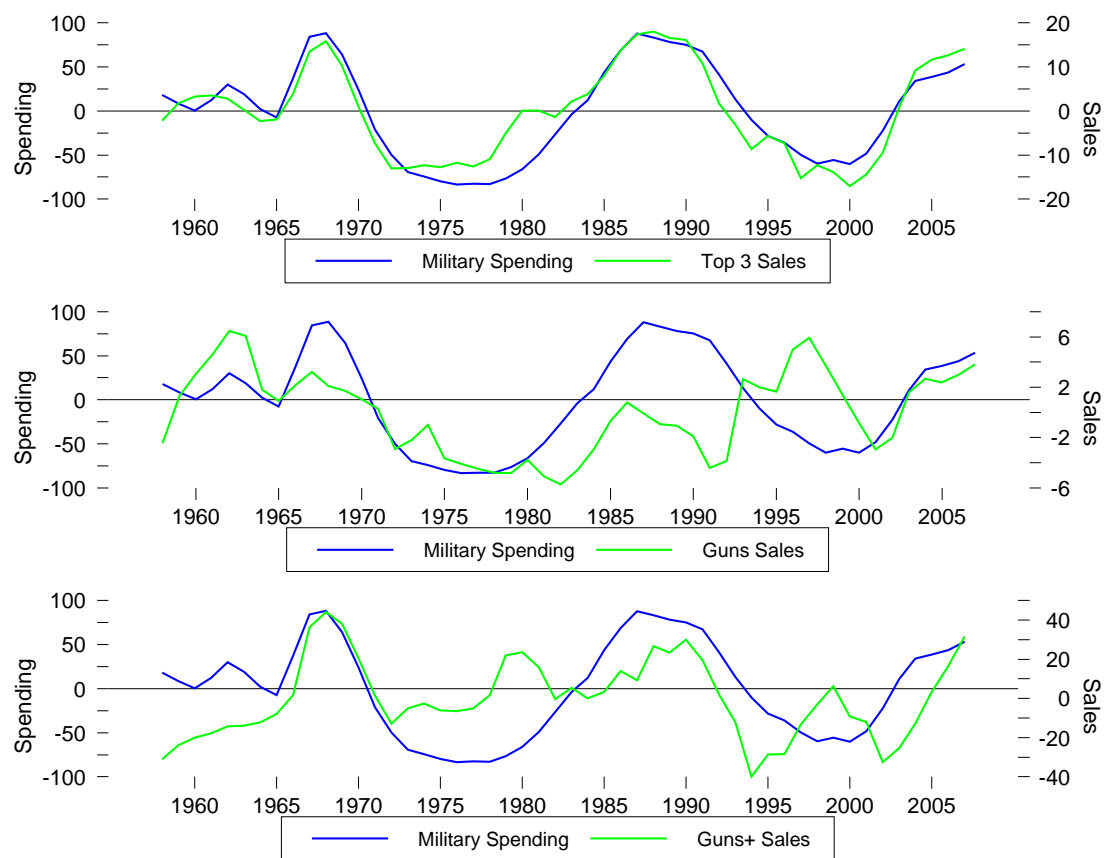


Figure 2: Accumulated Excess Returns of the Top 3 Contractors, 1947-2007

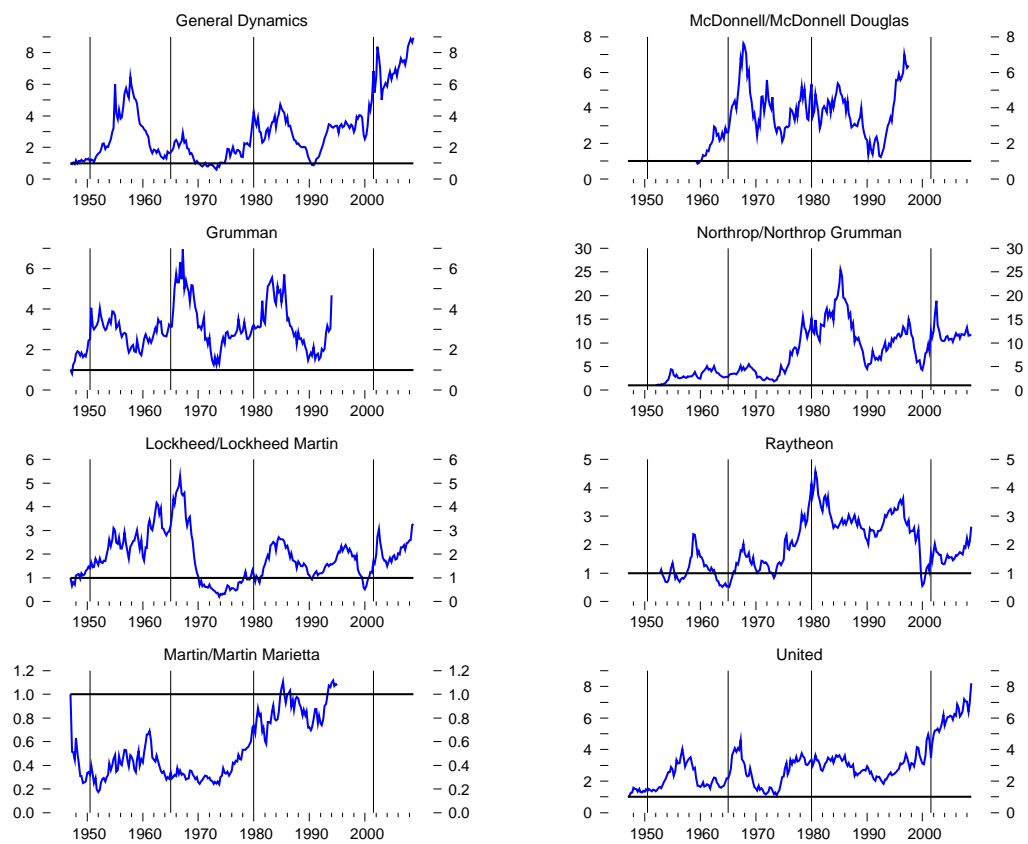


Figure 3: Accumulated Excess Returns and Military Spending, 1947-2007

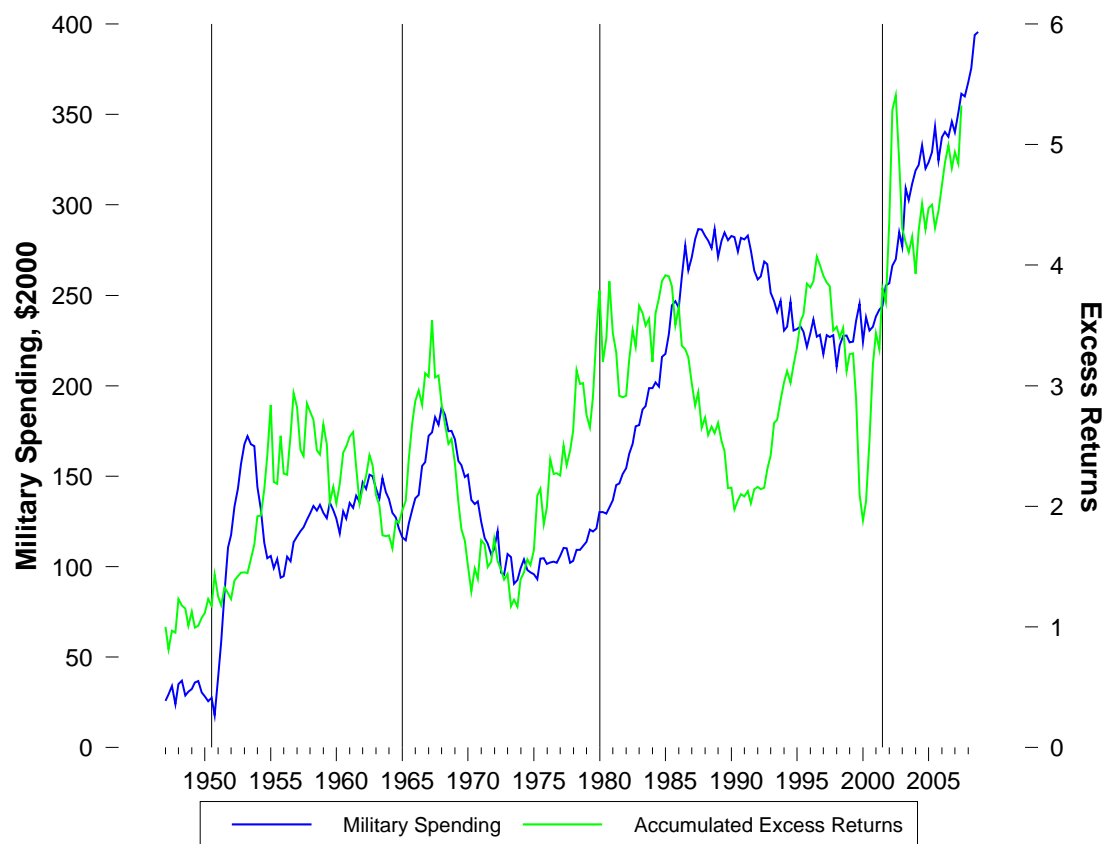
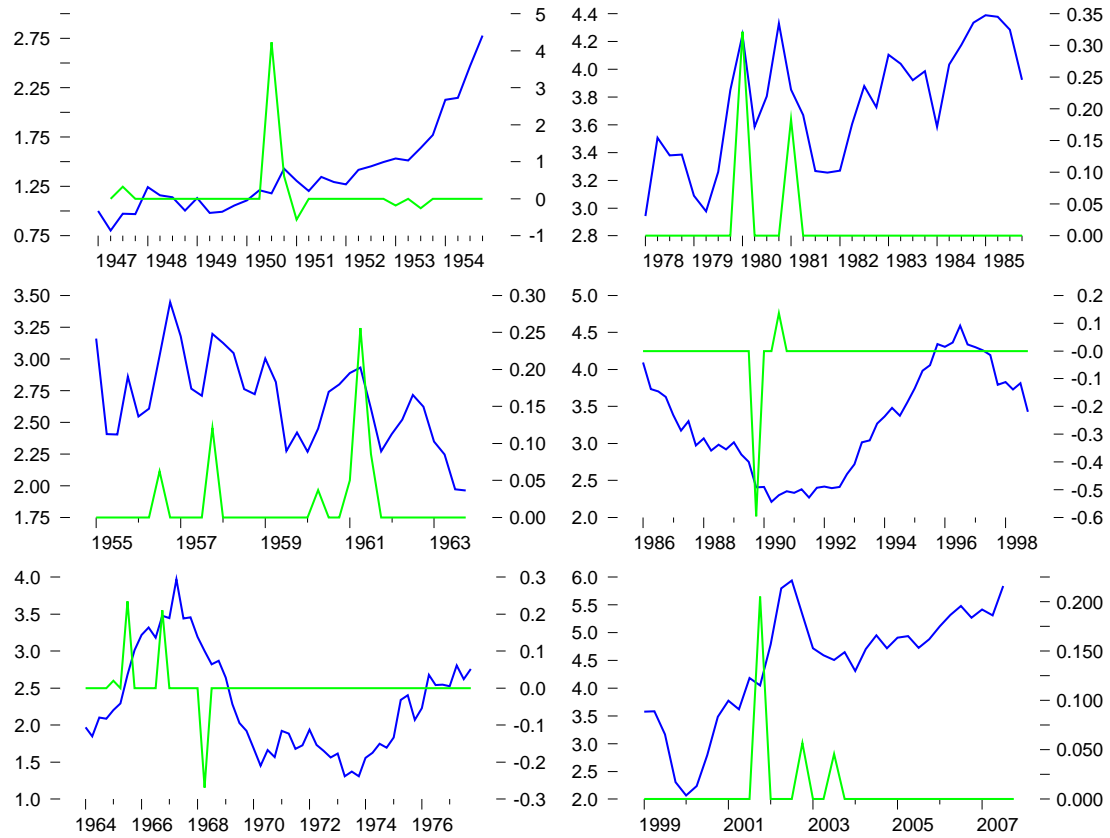
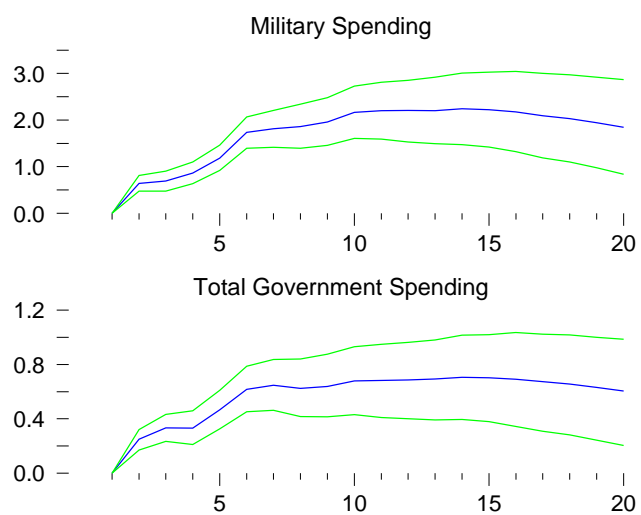


Figure 4: Excess Returns and Ramey (2008)'s Government Spending Shocks, 1947-2007



Note: Blue Line (left scale) is accumulated excess returns, Green Line (right scale) is Ramey (2008) Military Shocks.

Figure 5: Effects of Government Spending Shocks Using Excess Returns



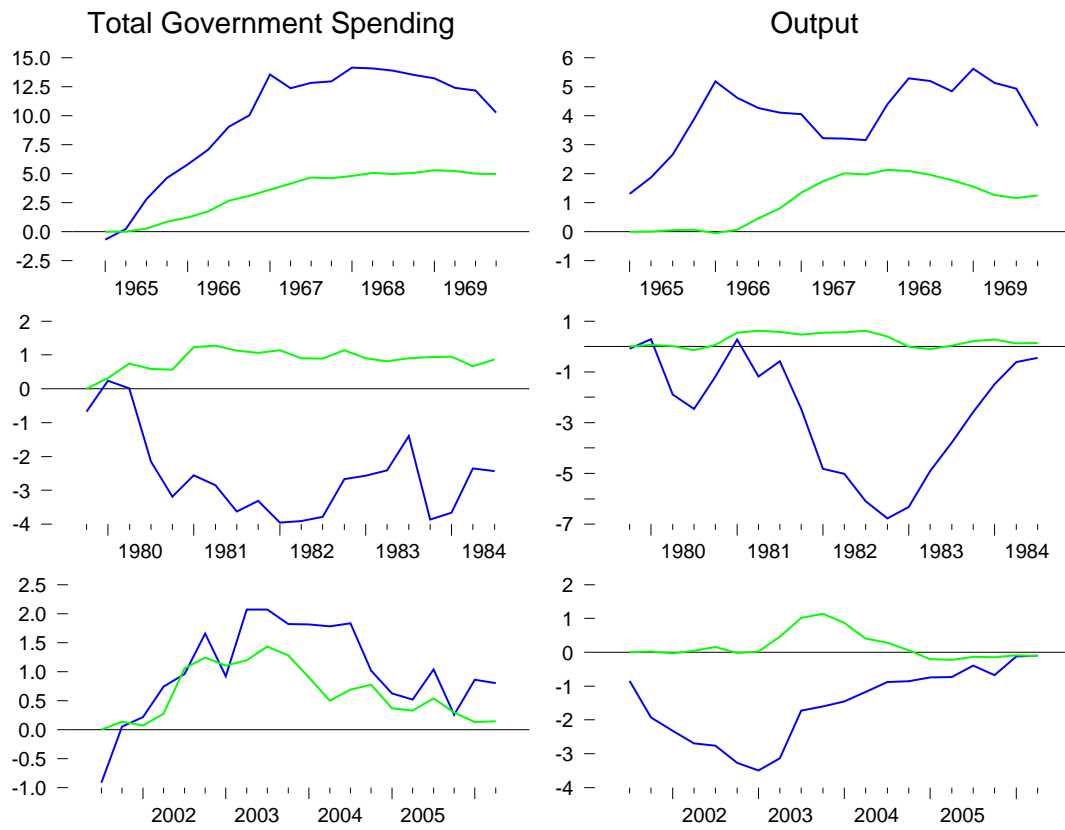
Note: Blue lines – point estimates, green lines – 68% posterior probability bands.

Figure 6: Effects Of Government Spending Shocks Using Excess Returns



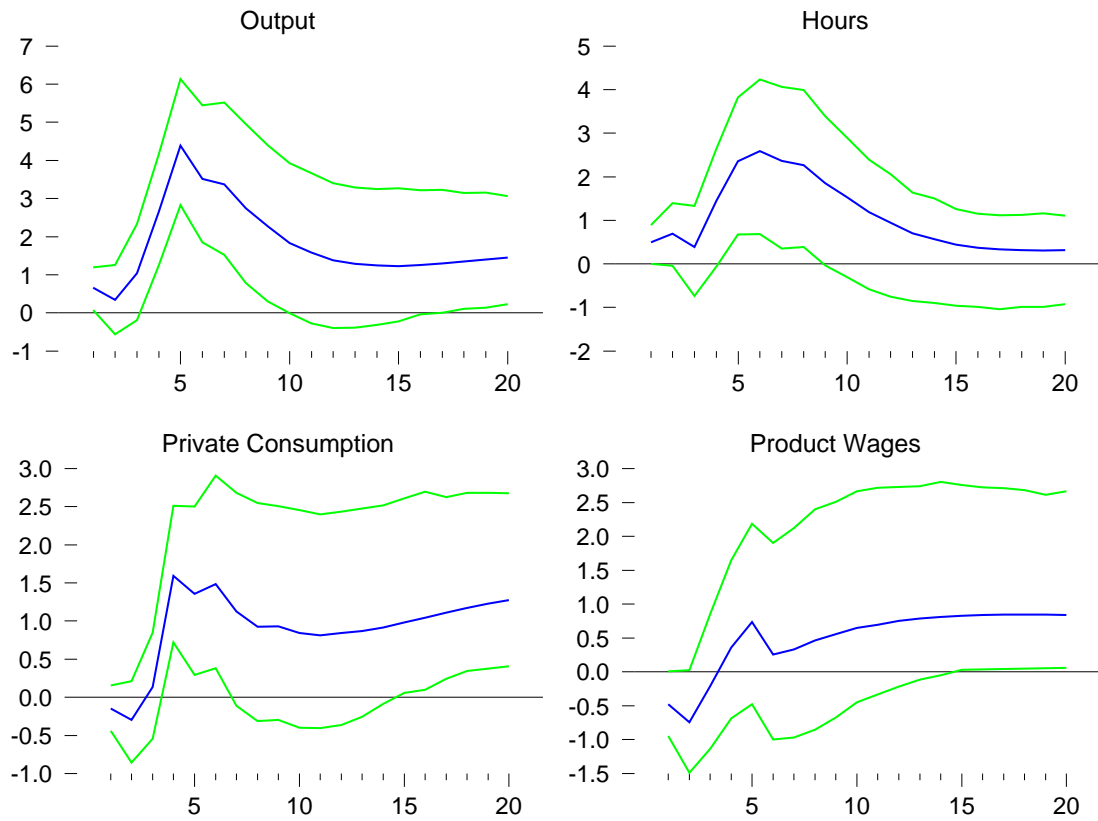
Note: Blue Lines – Point Estimates, Green Lines – 68% Posterior Probability Bands.

Figure 7: Historical Effects of Government Spending Shocks



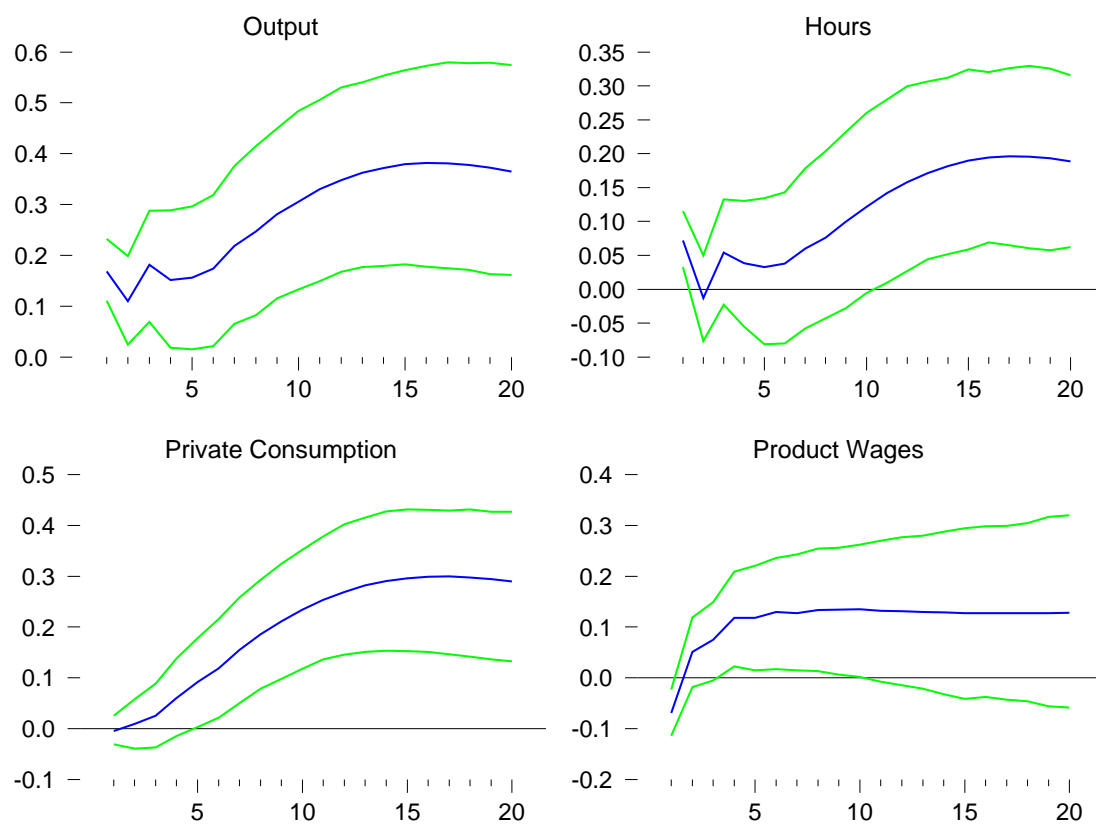
Note: Blue Lines – Actual deviation from trend, Green Lines – Deviation from trend predicted by spending shocks.

Figure 8: Effects of Government Spending Shocks Using War Dates



Note: Blue lines – point estimates, green lines – 68% posterior probability bands.

Figure 9: Effects of Government Spending Shocks Using Spending Innovations



Note: Blue lines – point estimates, green lines – 68% posterior probability bands.