Identifying Government Spending Shocks: It's All in the Timing

By

Valerie A. Ramey

University of California, San Diego
National Bureau of Economic Research

First draft: July 2006
This draft: May 2008

Abstract

Do shocks to government spending raise or lower consumption and real wages? Standard VAR identification approaches show a rise in these variables, where as the Ramey-Shapiro narrative identification approach finds a fall. I show that a key difference in the approaches is the timing. Both professional forecasts and the narrative approach shocks Granger-cause the VAR shocks, implying that the VAR shocks are missing the timing of the news. Simulations from a standard neoclassical model in which government spending is anticipated by several quarters demonstrate that VARs estimated with faulty timing can produce a rise in consumption even when it decreases in the model. In additional findings, I show that the measure of government spending used is crucial, since even the standard VAR ceases to produce Keynesian results when defense spending is used.

This paper expands upon my March 2005 discussion at the Federal Reserve Bank of San Francisco conference "Fiscal and Monetary Policy." I wish to thank Susanto Basu, Garey Ramey, participants at the 2006 NBER Summer Institute, participants in seminars at UCSD, Stanford, Berkeley, Duke and the New York Federal Reserve for helpful comments. Thomas Stark of the Federal Reserve Bank of Philadelphia generously provided unpublished forecasts. I gratefully acknowledge financial support from National Science Foundation grant SES-0617219 through the NBER.

I. Introduction

How does the economy respond to a rise in government purchases? Do consumption and real wages rise or fall? The literature remains divided on this issue. VAR techniques in which identification is achieved by assuming that government spending is predetermined within the quarter typically find that a positive government spending shock raises not only GDP and hours, but also consumption and the real wage (or labor productivity) (e.g. Rotemberg and Woodford (1992), Blanchard and Perotti (2002), Fatás and Mihov (2001), Mountford and Uhlig (2002), Perotti (2005), Pappa (2005), Caldara and Kamps (2006), and Galí, López-Salido, and Vallés (2007)). In contrast, analyses using the Ramey-Shapiro (1998) "war dates" typically find that while government spending raises GDP and hours, it lowers consumption and the real wage (e.g. Ramey and Shapiro (1998), Edelberg, Eichenbaum, and Fisher (1999), Burnside, Eichenbaum, and Fisher (2004), and Cavallo (2005)). Moreover, event studies such as Giavazzi and Pagano's (1990) analysis of fiscal consolidations in several European countries and Cullen and Fishback's (2006) analysis of WWII spending on local retail sales generally show a negative effect of government spending on private consumption.

Whether government spending raises or lowers consumption and the real wage is crucial for our understanding of how government spending affects GDP and hours. It is also important for distinguishing macroeconomic models. Consider first the neoclassical approach, as represented by papers such as Aiyagari, Christiano and Eichenbaum (1992) and Baxter and King (1993). A *permanent* increase in government spending financed by nondistortionary means creates a negative wealth effect for the representative household. The household optimally responds by decreasing its consumption and increasing its labor supply. Output rises as a result. The increased labor supply lowers the real wage and raises the marginal product of capital in the short run. The rise in the marginal product of capital leads to more investment and capital accumulation, which eventually brings the real wage back to its starting value. In the new steady-state, consumption is lower and hours are higher. A *temporary* increase in government spending in the neoclassical model has less impact on output because of the smaller wealth effect. Depending on the persistence of the shock, investment can rise or fall. In the short run, hours should still rise and consumption should still fall.[1]

The new Keynesian approach seeks to explain a rise in consumption, the real wage, and productivity found in most VAR analyses. For example Rotemberg and Woodford (1992) and Devereux, Head and Lapham (1996) propose models with oligopolistic (or monopolistic) competition and increasing returns in order to explain the rise in real wages and productivity. In the Devereux et al model, consumption may rise only if returns to specialization are sufficiently great. Galí, López-Salido, and Vallés (2006) show that only an "ultra-Keynesian" model with sticky prices, "rule-of-thumb" consumers, and off-the-labor-supply curve assumptions can explain how consumption and real wages can rise when government spending increases. Their paper makes clear how many special features the model must contain to explain the rise in consumption.

This paper reexamines the empirical evidence by comparing the two main empirical approaches to estimating the effects of government spending: the VAR approach and the Ramey-Shapiro narrative approach. After reviewing the set-up of both approaches and the basic results, I show that a key difference appears to be in the timing. In particular, I show that the Ramey-Shapiro dates Granger-cause the VAR shocks, but not vice versa. Thus, big increases in military spending are anticipated several quarters before they actually occur. I show this is also true for

[1] Adding distortionary taxes or government spending that substitutes for private consumption or capital adds additional complications. See Baxter and King (1993) and Burnside, Eichenbaum, and Fisher (2004) for discussions of these complications.

several notable cases of non-defense government spending changes. I then simulate a simple neoclassical model to demonstrate that failing to account for the anticipation effect can explain all of the differences in the empirical results of the two approaches.

I also question the use of overall government spending in VARs. I show that most nondefense spending is state and local spending on potentially productive activities such as education. When shocks to defense spending rather than overall spending are identified using a standard VAR, I find that the Keynesian effects on consumption and real wages disappear.

II. Fluctuations in Government Spending

This section reviews the trends and fluctuations in the components of government spending. One key difference between the two approaches is whether the shocks identified are shocks to all of government spending or to defense spending alone. As we will see, this may make a difference.

Figure 1 shows the paths of real defense spending per capita and total real government spending per capita in the post-WWII era.[2] The lines represent the Ramey-Shapiro (1998) dates, including the Korean War, the Vietnam War, and the Soviet invasion of Afghanistan, augmented by 9/11. These dates will be reviewed in detail below. The major movements in defense spending all come following one of the four military dates. Korea is obviously the most important, but the other three are also quite noticeable. There are also two minor blips in the second half of the 1950s and the early 1960s.

Looking at the bottom graph in Figure 1, we see that total government spending shows a significant upward trend over time. Nevertheless, the defense buildups are still distinguishable

[2] Per capita variables are created using the entire population, including armed forces overseas.

after the four dates. The impact of the Soviet invasion of Afghanistan has a delayed effect on total government spending, because nondefense spending fell.

Some have argued that the Korean War was unusual large, and thus should be excluded from the analysis of the effects of government spending. To put the Korean War in context, Figure 2 shows the log of real per capita government spending since 1889. The Korean War, which looked so large in a post-WWII graph, is dwarfed by the increases in government spending during the two world wars. The post-9/11 spending would not be noticeable if the line were not drawn.

Figure 3 returns to the post-WWII era and shows defense spending, nondefense federal spending, and state and local spending as a fraction of GDP (in nominal terms). The graph shows that relative to the size of the economy, each military buildup has become smaller over time. Federal nondefense spending is a minor part of government spending, hovering around two to three percent of GDP. In contrast, state and local spending has risen from around five percent of GDP in 1947 to over twelve percent of GDP now. Since state and local spending is driven in large part by cyclical fluctuations in state revenues, it is not clear that aggregate VARs are very good at capturing shocks to this type of spending. For example, California dramatically increased its spending on K-12 education when its tax revenues surged from the dot-com boom in the second half of the 1990s.

The graphs suggest that defense spending is a major part of the variation in government spending around trend. To quantify the importance of defense spending, I estimate a variance decomposition of government spending using a simple VAR. I include the log of real total government spending, defense spending and either state and local spending or federal nondefense. All variables are in real per capita terms. Four lags and a trend are included in the

VAR. Table 1 shows the variance decomposition of total government spending at various horizons for different orderings of the variables. The table shows that no matter the ordering, shocks to defense spending account for 80 to 90 percent of the forecast error variance of total government spending for horizons of four to twenty quarters. Thus, the variance decomposition suggests that shocks to defense spending accounts for almost all of the unforeseen changes in total government spending.

What kind of spending constitutes nondefense spending? Figure 4 shows annual spending for some key functions, as a percent of total government spending. Spending by function is available annually only since 1959. The category of education, public order (which includes police, courts and prisons), and transportation expenditures has increased from 30 percent of total government spending to around 50 percent.

The standard VAR approach includes shocks to this type of spending in its analysis (e.g. Blanchard and Perotti (2002)). Such an inclusion is questionable for several reasons. First, the biggest part of this category, education, is driven in large part by demographic changes, which can have many other effects on the economy. Second, these types of expenditures would be expected to have a positive effect on productivity and hence would have a different effect than government spending that has no direct production function impacts. As Baxter and King (1993) make clear, the typically neoclassical predictions about the effects of government spending change completely when government spending is “productive.” Even the “other” category shown has questionable impacts, since it includes “other economic activities” such as spending on natural resources, housing and health. Thus, including these categories in spending shocks is not the best way to test the neoclassical model versus the Keynesian model.

Some of the analyses, such as Eichenbaum and Fisher (2005) and Perotti (2007), have tried to address this issue by using only "government consumption" and excluding "government investment." Unfortunately, this National Income and Product Account distinction does not help. As the footnotes to the NIPA tables state: "Government consumption expenditures are services (such as education and national defense) produced by government that are valued at their cost of production….Gross government investment consists of general government and government enterprise expenditures for fixed assets." Thus, since teacher salaries are the bulk of education spending, they would be counted as "government consumption."

In sum, three conclusions emerge from this review of the data. First, while nondefense spending accounts for most of the trend in government spending, fluctuations in defense spending account for almost all of the fluctuations in total government spending relative to trend. Second, most nondefense spending is done by state and local governments, not by the federal government, so it is not clear that aggregate VARs are very good at capturing shocks to this type of spending. Third, much of nondefense expenditures consists of spending that may impact the productivity of the economy, and thus should not be included in analyses of the pure effects of government spending.

III. Identifying Government Spending Shocks: VAR vs. Narrative Approaches

A. The VAR Approach

Blanchard and Perotti (2002) have perhaps the most careful and comprehensive approach to estimating fiscal shocks using VARs. To identify shocks, they first incorporate institutional information on taxes, transfers, and spending to set parameters, and then estimate the VAR. Their basic framework is as follows:

$$Y_t = A(L,q)Y_{t-1} + U_t$$

where Y_t consists of quarterly real per capita taxes, government spending, and GDP. Although the contemporaneous relationship between taxes and GDP turns out to be complicated, they find that government spending does not respond to GDP or taxes contemporaneously. Thus, their identification of government spending shocks is identical to a Choleski decomposition in which government spending is ordered before the other variables. When they augment the system to include consumption, they find that consumption rises in response to a positive government spending shock. Galí et al (2007) use this basic identification method in their study which focuses only on government spending shocks and not taxes. They estimate a VAR with additional variables of interest, such as real wages, and order government spending first. Perotti (2007) uses this identification method to study a system with seven variables. [3]

B. The Ramey-Shapiro Narrative Approach

In contrast, Ramey and Shapiro (1998) use a narrative approach to identify shocks to government spending. Because of their concern that many shocks identified from a VAR are simply anticipated changes in government spending, they focus only on episodes where *Business Week* suddenly began to forecast large rises in defense spending induced by major political events that were unrelated to the state of the U.S. economy. The three episodes identified by Ramey and Shapiro were as follows:

[3] See the references listed in the introduction to see the various permutations on this basic set-up.

Korean War

On June 25, 1950 the North Korean army launched a surprise invasion of South Korea, and on June 30, 1950 the U.S. Joint Chiefs of Staff unilaterally directed General MacArthur to commit ground, air, and naval forces. In the July 1, 1950 issue, *Business Week* wrote: "We are no longer in a peacetime economy. Even if the Communists should back down in Korea, we have had a warning of what can happen any time at all or in any of the Asiatic nations bordering on the USSR. The answer will be more money for arms." (p. 9). By August 1950, *Business Week* was predicting that defense spending would more than triple by fiscal year 1952. After early UN victories, *Business Week* in October 1950 predicted a quick end to hostilities in Korea, but a continuing defense spending increase. It predicted a somewhat faster pace of spending after China entered the war on November 9, 1950, but pointed out that it would take at least six months to translate defense programs into men and material.

The Vietnam War

Despite the military coup that overthrew Diem on November 1, 1963, *Business Week* was still talking about defense cuts for the next year (November 2, 1963, p. 38; July 11, 1964, p. 86). Even the Gulf of Tonkin incident on August 2, 1964 brought no forecasts of increases in defense spending. However, after the February 7, 1965 attack on the U.S. Army barracks, Johnson ordered air strikes against military targets in North Vietnam. The February 13, 1965, *Business Week* said that this action was "a fateful point of no return" in the war in Vietnam. Fighting escalated in the spring and expenses increased beyond initial estimates. In July 1965, Johnson told the nation "This is really war" and doubled draft quotas. On December 4, 1965, *Business Week* said that the price tag for the Vietnam War was drastically marked up that week and that there was no end in sight.

The Carter-Reagan Buildup

The long decline in defense spending began to turn around slightly when Carter promised NATO that the US would increase defense spending by an inflation-adjusted three percent a year. The Soviet invasion of Afghanistan on December 24, 1979 led to a significant turnaround in U.S. defense policy. The event was particularly worrisome because some believed it was a possible precursor to actions against Persian Gulf oil countries. The January 21, 1980 *Business Week* (p.78) printed an article entitled "A New Cold War Economy" in which it forecasted a significant and prolonged increase in defense spending. Reagan was elected by a landslide in November 1980 and in February 1981 he proposed to increase defense spending substantially over the next five years.

These dates were based on data up through 1998. Owing to recent events, I now add the following date to these war dates:

9/11

On September 11, 2001, terrorists struck the World Trade Center and the Pentagon. On October 1, 2001, *Business Week* forecasted that the balance between private and public sectors would shift, and that spending restraints were going "out the window." In this case, though, it was clear that some of the increased spending they were discussing was not defense, but rather industry bailouts and the like. To recall the timing of key subsequent events, the U.S. invaded Afghanistan soon after 9/11. It invaded Iraq on March 20, 2003.

The military date variable takes a value of unity in 1950:3, 1965:1, 1980:1, and 2001:3, and zeroes elsewhere. Table 2 compares the predictive power of this variable for real per capita defense spending growth relative to two other variables. The first row shows the R-squared from a regression of the growth of real defense spending on eight lags of itself. The R-squared is 0.37. The second row reports the R-squared of the growth of real defense spending on the current value plus eight lags of the military date variable. The R-squared is still quite high, 0.27. The last row investigates the predictive power of a scaled military date variable. Burnside, Eichenbaum and Fisher (2004) and Eichenbaum and Fisher (2005) advocate scaling the variable to allow the different sizes in build-ups. Based on the peak rise in spending, Eichenbaum and Fisher assign a value of 1 to Korea, 0.3 to Vietnam, and 0.1 to Carter-Reagan. I updated this with an 0.1 for 9/11. The third row shows that the scaled military date variable explains 57 percent of the growth of defense spending. Of course, one reason this variable has better predictive power is that it contains information about the future peak of military spending. In any case, even the basic military date variable has substantial explanatory power for defense spending.

To identify government spending shocks, the military date variable is embedded in the standard VAR, but ordered before the other variables.[4] Choleski decomposition shocks to the military date variable rather than to the government spending variable are used to identify government shocks.

[4] The original Ramey-Shapiro (1998) implementation did not use a VAR. They regressed each variable of interest on lags of itself and the current and lagged values of the military date variable. They then simulated the impact of changes in the value of the military date variable. The results were very similar to those obtained from embedding the military variable in a VAR.

C. Comparison of Impulse Response Functions

Consider now a comparison of the effects of government spending increases based on the two identification methods. In both instances, I use a VAR similar to the one used recently by Perotti (2007) with a few modifications of variable definitions. The VAR consists of the log real per capita quantities of total government spending, GDP, the Barro-Sahasakul tax rate, total hours worked, nondurable plus services consumption, and private fixed investment, as well as the log of nominal compensation in private business divided by the deflator in private business. The Barro-Sahasakul tax rate is from Perotti (2007). Chained nondurable and services consumption are aggregated using Whelan's (2000) method. I use total hours worked instead of private hours worked based on Cavallo's (2005) work showing that a significant portion of rises in government spending consists of increases in the government payroll. Total hours worked are based on unpublished BLS data and are available on my web site. Also, note that I use a product wage rather than a consumption wage. Ramey and Shapiro (1998) show both theoretically and empirically why it is the product wage that should be used when trying to distinguish models of government spending. Defense spending tends to be concentrated in a few industries, such as manufactured goods. Ramey and Shapiro show that the relative price of manufactured goods rise significantly during a defense buildup. Thus, product wages in the expanding industries can fall at the same time that the consumption wage is unchanged or rising.[5]

Both VARs are specified in levels, with a time trend and four lags included. Because the tax rate series only extends to 2003:4, I use quarterly data from 1947:1 to 2003:4. In the VAR identification, the government spending shock is identified by a Choleski decomposition in which government spending is ordered first. In the war dates identification, the current value

[5] The main reason that Rotemberg and Woodford (1992) find that real wages increase is that they construct their real wage by dividing the wage in manufacturing by the implicit price deflator. Ramey and Shapiro show that the wage in manufacturing divided by the price index for manufacturing falls during a defense buildup.

and four lags of a dummy variable with the military date are also included. The military date takes a value of unity in 1950:3, 1965:1, 1980:1, and 2001:3.[6] I compare the effects of shocks that are normalized so that the log change of government spending is unity at its peak in both specifications.

Figures 5A and 5B show the impulse response functions. Following the government spending VAR literature (e.g. Galí et al (2007) and Perotti (2007)), the standard error bands are 68% bands, based on bootstrap standard errors. This standard of significance is far below the standards in other literatures, but I conform in order that the graphs be clearer. Also, more parsimonious representations tend to give similar point estimates with smaller standard errors, so the results are often significant at conventional levels (e.g. Ramey-Shapiro (1998) and Edelberg et al (1999).

The first column shows the results from the VAR identification and the second column shows the results from the war dates identification. Figure 5A shows the effects on government spending, GDP, and hours. The results are qualitatively consistent across the two identification schemes for these three variables. By construction, total government spending rises by the same amount, although the peak occurs several quarters earlier in the VAR identification. This is the first indication that a key difference between the two methods is timing. GDP rises in both cases, but its rise is much greater in the case of the war dates identification. Hours rise slightly in the VAR identification, but much more strongly in the war dates identification. A comparison of the output and hours response shows that productivity rises slightly in both specifications.

Figure 5B shows the cases in which the two identification schemes differ in their implications. The VAR identification scheme implies that government spending shocks raise

[6] Burnside, Eichenbaum and Fisher (2004) allow the value of the dummy variable to differ across episodes according to the amount that government spending increase. They obtain very similar results.

consumption, lower investment for two years, and raise the real wage. In contrast, the war dates identification scheme implies that government spending shocks lower consumption, raise investment for a few quarters before lowering it, and lower the real wage.

Overall, these two approaches give diametrically opposed answers with regard to some key variables. The next section presents empirical evidence and a theoretical argument that can explain the differences.

IV. The Importance of Timing

A concern with the VAR identification scheme is that some of what it classifies as "shocks" to government spending may well be anticipated. Indeed, my reading of the narrative record uncovered repeated examples of long delays between the decision to increase military spending and the actual increase. At the beginning of a big buildup of strategic weapons, the Pentagon first spends at least several months deciding what sorts of weapons it needs. The task of choosing prime contractors requires additional time. Once the prime contracts are awarded, the spending occurs slowly over time. Quarter-to-quarter variations are mostly due to production scheduling variations among prime contractors.

From the standpoint of the neoclassical model, what matters for the wealth effect are changes in the present discounted value of government purchases, not the particular timing of the purchases. Thus, it is essential to identify when news becomes available about a major change in the present discounted value of government spending.

Blanchard and Perotti (2002) worried about the timing issue, and devoted Section VIII of their paper to analyzing it. To test for the problem of anticipated policy, they included future values of the estimated shocks to determine whether they affected the results. They found that the response of output was greater once they allowed for anticipation effects (see Figure VII). Unfortunately, they did not show how the responses of consumption or real wages were affected. Perotti (2005) approached the anticipation problem by testing whether OECD forecasts of government spending predicted his estimated government spending shocks. For the most part, he found that they did not predict the shocks.

In the next subsection, I show that the war dates as well as professional forecasts predict the VAR government spending shocks. I also show how in each war episode, the VAR shocks are positive several quarters after *Business Week* or the Office of Management and Budget started forecasting increases in defense spending. In the second subsection, I present simulations from a simple theoretical model that shows how the difference in timing can explain most of the differences in results. In the final subsection, I show that delaying the timing of the Ramey-Shapiro dates produces the Keynesian results.

A. Empirical Evidence on Timing Lags

To compare the timing of war dates versus VAR-identified shocks, I estimate shocks using the VAR discussed above except with defense spending rather than total government spending as the key variable. I then plot those shocks around the war dates.

Figures 6A and 6B shows the path of log per capita real defense spending, the series of identified shocks, and some long-term forecasts. Consider first the Korean War in Figure 6A. The first vertical line shows the date when the Korea War started. The second vertical line indicates when the armistice was signed in July 1953. According to the VAR estimates, shown in the middle graph, there was a series of three large positive shocks to defense in 1950:4, 1951:1, and 1951:2. However, as *Business Week* made clear, the path of defense spending during these three quarters was anticipated as of August 1950. The bottom graph shows *Business Week's* forecasts of defense spending. The June 1950 forecast, made before the Korean War started, predicted that defense spending would rise slowly from $13 billion per year. Two months later in August 1950, *Business Week* correctly predicted the rise in defense spending through fiscal year 1954. Thus, it is clear that the positive VAR shocks are several quarters too late. It is also interesting to note that while *Business Week* was predicting a future decline in defense spending as early as April 1953 when a truce seemed imminent, the VAR records a negative defense spending shock late in the first quarter of 1954. Thus, the VAR shocks are not accurately reflecting news about defense spending.

Forecasts were not so accurate for Vietnam. The Office of Management and Budget (OMB) forecasts shown in the bottom panel suggests that there were continuing positive defense surprises during the Vietnam War. Each January, the forecasts jumped significantly. However, the press at the time often cited leading senators who predicted much higher expenditures months in advance. In any case, it is unclear whether the VAR has the timing right on its shocks since it is clear from the press that spending was known at least a few quarters in advance.

The VARs show many positive shocks during the Carter-Reagan build-up through 1985. The bottom panel shows, however, that as of January 1981, the OMB was very accurately predicting spending in fiscal years 1981-1984. On the other hand, the October 1981 forecast over-predicted defense spending in fiscal years 1985 and 1986. However, all of the forecast error for 1985 and 1986 can be attributed to the fact that inflation fell much more quickly than expected. In real terms, the October 1981 predictions for the 1985 and 1986 fiscal years were very accurate. Yet the VARs produce large positive shocks for those years.

After 9/11 the VAR implies virtually no shocks until the second quarter of 2003. Yet the February 2002 forecast for the next several years was raised significantly relative to the pre-9/11 April 2001 forecast. The February 2003 OMB forecast under-predicted spending, primarily because it assumed no invasion of Iraq, although many believed that it would happen.

As additional evidence of the ability of the private sector to forecast, Figure 7 shows the government spending growth forecasts from the Survey of Professional Forecasters, available from the Federal Reserve Bank of Philadelphia. Before the third quarter of 1981, forecasters were asked to predict nominal defense spending. I convert the forecasts to real defense spending using the forecasts of the GDP deflator. Starting in the third quarter of 1981, forecasters were asked to predict real federal spending. The forecasts shown in the graph for quarter t are the forecast made in t for the growth rate of spending between t - 1 and t + 4. It is clear that forecasters predicted significantly higher defense spending growth for the year ahead starting in the first quarter of 1980, which was just after the Soviet invasion of Afghanistan in December 1979. Similarly, forecasters predicted higher federal spending growth beginning in the fourth quarter of 2001, just after 9/11.[7] Note also that the invasion of Iraq in March 2003 did not lead to a jump up in forecasts in the second quarter of 2003. In fact, the initial invasion went so well that forecasters reduced their forecasts in the third quarter of 2003.

Overall, it appears that much of what the VAR might be labeling as "shocks" to defense spending may have been forecasted. To test this hypothesis formally, I perform Granger causality tests between various variables and the VAR-based defense and total government spending shocks. In addition to the military dates variable, I also use estimates from the Survey of Professional Forecasters for real federal government spending forecasts starting in the third

---

[7] The higher predictions do not show up in the third quarter of 2001 because the forecasters had already returned their surveys when 9/11 hit.

quarter of 1981.[8] I use the implied forecast dating from quarter t-4 of the log change in real spending from quarter t-5 to quarter t.

Table 3 shows the results. The evidence is very clear: the war dates Granger-cause the VAR shocks but the VAR shocks do not Granger-cause the war dates. Moreover, the VAR shocks, which are based on information up through the previous quarter, are Granger-caused by professional forecasts *made four quarters earlier.* Thus, the VAR shocks (which use information through the previous quarter) are forecastable even with information that was available four quarters in advance.

One should be clear that timing is not an issue only with defense spending. Consider the interstate highway program. In early 1956, *Business Week* was predicting that the "fight over highway building will be drawn out." By May 5, 1956, *Business Week* thought that the highway construction bill was a sure bet. It fact it passed in June 1956. However, the multi-billion dollar program was intended to stretch out over 13 years. It is difficult to see how a VAR could accurately reflect this program. Another example is schools for the Baby Boom children. Obviously, the demand for schools is known several years in advance. Between 1949 and 1969, real per capita spending on public elementary and secondary education increased 300%.[9] Thus, a significant portion of non-defense spending is known months, if not years, in advance.

B. The Importance of Timing in a Theoretical Model

Macroeconomists have known for a long time that anticipated policy changes can have very different effects from an unanticipated change. For example, Taylor (1993, Chapter 5)

---

[8] Forecasts of nominal defense spending are available starting in 1969 to 1981. However, the series is relatively short and contains some missing values, so I did not use it.

[9] The nominal figures on expenditures are from the *Digest of Education Statistics*. I used the GDP deflator to convert to real.

shows the effects of a change in government spending, anticipated two years in advance, on such variables as GDP, prices, interest rates and exchange rates. He does not consider the effects on consumption or real wages, however. More recently, Yang (2005) shows that foresight about tax rate changes significantly changes the responses of key variables in theoretical simulations.

To see how important anticipation effects can be for government spending, consider a simple neoclassical growth model with government spending and nondistortionary taxes. The model is deliberately stylized in order to show how important these effects can be. The key equations of the social planner problem are as follows, where the parameters are calibrated to a quarterly frequency:

$Y_t = (Z_t N_t)^{0.67} K_t^{0.33}$	Production Function
$U = \log(C_t) + \varphi_t \cdot \log(1 - N_t)$	Utility of the representative household
$Y_t = C_t + I_t + G_t$	Resource constraint
$K_{t+1} = I_t + (1 - 0.023) K_t$	Capital Accumulation

Y is output, *N* is labor, *K* is capital, *C* is consumption, *I* is investment, and *G* is government purchases. Government purchases are financed with nondistortionary taxes. Households maximize the present discounted value of utility *U* with discount factor $\beta = 0.99$.

The driving processes are calibrated as follows:

$$\ln Z_t = .95 \cdot \ln Z_{t-1} + ez_t, \quad \sigma_{ez} = 0.01$$

$$\ln \varphi_t = .95 \cdot \ln \varphi_{t-1} + e\varphi_t, \quad \sigma_{e\varphi} = 0.008$$

$$\ln GF_t = \text{constant} + 1.4 \ln GF_{t-1} - 0.18 \ln GF_{t-2} - 0.25 \ln GF_{t-3} + eg_t, \quad \sigma_{eg} = 0.028$$

$$\ln G_t = GF_{t-2}$$

The calibration for the technology shock is standard. The calibrations for the marginal rate of substitution shock and the government spending shock are based on my estimates from data. *GF* is the forecast of government spending whereas *G* is actual government spending. This specification allows agents to know the shock to actual government spending two quarters in advance. The model is solved with Dynare. All variables reported below are in logs.

Figure 7 shows the theoretical impulse response functions from this model. News becomes available in quarter 0, but government spending does not start to increase until quarter two. In contrast, output, hours, consumption, investment and real wages all jump in quarter 0 when the news arrives. Output and hours rise immediately, while consumption and real wages fall immediately. Interestingly, investment rises for several quarters before it falls. Investment rises because the increase in government spending is expected to be prolonged, so firms want to build up their capital stocks.

Looking at these graphs, one wonders what happens in a VAR if one identifies the government spending shock from when government spending actually changes. To study this effect, I simulate data from this stylized model. In order to increase the number of shocks to four so that I can include at least three variables in the VARs, I also assume that there is measurement error in the logarithm of output, and that it follows an AR(1) with autocorrelation coefficient of

0.95 and standard errors of 0.005. I then run two types of trivariate VARs on the simulated data. In the "faulty timing" VARs, I use actual government spending, output, and the variable of interest, and identify the shock as the shock to government spending, which is ordered first. In the "true timing" VARs, I use the news ("GF"), output, and the variable of interest, and identify the shock as the shock to the news variable which is ordered first. In all cases, I use four lags of the variables.

Figures 8A and 8B show the results.[10] The faulty timing VARs, which identify shocks from actual government spending, are shown in the left column and the true timing VARs, which identify shocks from the news about government spending, are shown in the right column. The patterns across the two columns are strikingly similar to the patterns across the VARs on real data using the two methods shown in Figures 5A and 5B. In particular, the faulty timing finds a much smaller response of GDP than the true timing, just as the standard VAR identification method finds a smaller response than the Ramey-Shapiro military dates. The same is true of hours. In Figure 8B, we see that the faulty timing VAR leads to a rise in consumption, whereas the true timing VAR leads to a fall in consumption. In both cases, real wages fall, although the pattern is different. The problem with the faulty timing VAR is that it often catches the variables *after* they have already had an initial response to the news. The contemporaneous correlation of the estimated shocks using actual government spending in the VAR with the true shocks is -0.01. The correlation of the identified shock and the true shock *two quarters ago* is 0.4, but the timing is off. Thus, the faulty timing VAR picks up nothing of the true shocks. In contrast, the correlation between the estimated shocks using the news and the true shocks is 0.97.

[10] The government spending and output responses shown are from the trivariate VAR that also contains consumption.

As stated above, this model is very stylized in order to make the point in the simplest possible model. One could modify the theoretical model to incorporate elements such as habit persistence in consumption and/or adjustment cost in investment. In this case, the true responses would be more dragged out and missing the timing by two quarters would have a somewhat smaller effect. However, introducing more realistic lags would likely have changed the impulse responses just as they did in the previous example.

C. Would Delaying the Ramey-Shapiro Dates Lead to Keynesian Results?

If the theoretical argument of the last section applies to the current situation, then delaying the timing of the Ramey-Shapiro dates should result in VAR-type Keynesian results.[11] To investigate this possibility, I shifted the four military dates to correspond with the first big positive shock from the VAR analysis. Thus, instead of using the original dates of 1950:3, 1965:1, 1980:1, and 2001:3, I used 1951:1, 1965:3, 1980:4, and 2003:2.

Figure 9 shows the results using the baseline VAR of the previous sections. As predicted by the theory, the delayed Ramey-Shapiro dates applied to actual data now lead to rises in consumption and the real wage, similarly to the shocks from the standard VARs. Thus, the heart of the difference between the two results appears to be the VAR's delay in identifying the shocks.

Alternatively, one could try to estimate the VAR and allow future identified shocks to have an effect. Blanchard and Perotti (2002) did this for output, but never looked at the effects on consumption or wages. Based on an earlier draft of my paper, Tenhofen and Wolff (2007)

[11] This idea was suggested to me by Susanto Basu.

analyze such a VAR for consumption and find that when the VAR timing changes, positive shocks to defense spending lead consumption to fall.

Thus, all of the empirical and theoretical evidence points to timing as being key to the difference between the standard VAR approach and the Ramey-Shapiro approach. The fact that the Ramey-Shapiro dates Granger-cause the VAR shocks suggests that the VARs are not capturing the timing of the news. The theoretical analysis shows that timing is crucial in determining the response of the economy to news about increases in government spending.

V. The Importance of the Composition of Government Spending

The second section of this paper made the argument that nondefense components of government spending did not necessarily have the required properties for implementing a clean test of Keynesian versus neoclassical predictions. In particular, I argued that most nondefense spending is done by state and local governments, and much of it is arguably productive. Moreover, aggregate VARs may not adequately capture the drivers of state and local spending.

To determine the importance of composition, I re-estimated the baseline VAR with defense spending substituted for all government spending. The shocks are identified by a standard Choleski decomposition, and no adjustments are made to the timing.

Figure 10 shows the impulse response functions. The figure shows that when defense spending is substituted for all government spending, even the standard VAR identification fails to produce rises in consumption or real wages. Thus, the composition of government spending appears to matter significantly of the effects of shocks.[12]

The issue of composition has implications for studies of the U.S. versus other OECD countries. Since other OECD countries spend much less on defense, most of their spending may

[12] This result is not necessarily robust to other possible VAR specifications.

potentially complement private consumption and private capital. The U.S. is unique in both the level and variance of its defense spending relative to other types of spending. Moreover, the U.S. is unusual in that it has not fought a war on its own territory since the Civil War. For these reasons, defense spending in the U.S. is probably the best test of pure fiscal theories.

VI. Are Consumers Really so Rational?

The results of the previous sections support the neoclassical model in its predictions about the effects of news of pure government spending shocks on consumption and real wages. A key part of the explanation is that consumers react quickly to news. One might be skeptical, however, that consumers could be so rational.

In fact, the results presented do not require consumers to be "too" Ricardian. After most of the military dates, *Business Week* talked of either planned tax increases or a delay in a proposed tax cut. The narrative made it clear that most of the public believed that at least part of the increase in spending would be financed by tax increases in the near future.

The top panel of Figure 11 shows the effect of government spending shocks on tax rates in the baseline VARs. The tax rate used is Perotti's (2007) update of the Barro-Sahasakul tax rate. Tax rates were included in all previous specifications, but were not included with the earlier graphs. The graphs show that for both identification schemes, taxes rise after a government spending shock. However, they rise by significantly more after a war date than after a VAR shock. Thus, it is possible that the neoclassical response of consumption could be due to the immediate rise of taxes.

The behavior of interest rates is also of interest when considering consumer responses. The bottom panel of Figure 11 shows the impulse response functions for interest rates. These are based on specifications in which the real interest rate, defined as the Baa bond rate less the CPI inflation rate, is substituted for wages. The VAR shocks imply that interest rates fall, whereas the war dates shocks imply that real interest rates rise. It is likely that the timing issues discussed above explain the difference.

As final support for the hypothesis that consumers respond quickly to news, Figure 12 shows the behavior of durable consumer expenditures just after North Korea invaded South Korea at the end of June 1950. Both the Keynesian and Neoclassical theories predict that durable consumption expenditures should have fallen. In fact, they rose dramatically. According to press reports, consumers feared a return to WWII rationing and immediately went out and bought goods that they thought would be rationed. Thus, this evidence supports the notion that consumers can respond very quickly to news.

VII. Conclusions

This paper has explored possible explanations for the dramatically different results between standard VAR methods and the narrative approach for identifying shocks to government spending. I have shown that the main difference is that the narrative approach shocks appear to capture the timing of the news about future increases in government spending much better. In fact, these shocks Granger-cause the VAR shocks. My theoretical results show how timing can account for all of the difference in the results across the two methods. Because the VAR approach captures the shocks too late, it misses the initial decline in consumption and real wages that occurs as soon as the news is learned. I show that delaying the timing on the Ramey-Shapiro dates replicates the VAR results.

Moreover, I have argued that for testing between competing theories of the effects of pure government spending shocks, U.S. defense expenditures are the best measure to use. I have shown that most nondefense spending occurs at the state and local level, and that much of it is productive spending. When I substitute defense spending for government spending in the baseline VAR, I show that even standard VAR identification implies that consumption and real wages fall in response to a positive spending shock.

References

Aiyagari, Rao, Laurence Christiano and Martin Eichenbaum, "The Output, Employment and Interest Rate Effects of Government Consumption." *Journal of Monetary Economics* 30 (1992), pp. 73–86.

Baxter, Marianne and Robert G. King, "Fiscal Policy in General Equilibrium," *American Economic Review* 83 (1993), pp. 315–334.

Blanchard, Olivier and Roberto Perotti, "An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output," *Quarterly Journal of Economics* (November 2002): 1329-1368.

Burnside, Craig, Martin Eichenbaum, and Jonas Fisher, "Fiscal Shocks and their Consequences," *Journal of Economic Theory* 115 (2004): 89-117.

Caldara, Dario and Christophe Kamps, "What are the Effects of Fiscal Policy Shocks? A VAR-based Comparative Analysis," September 2006 working paper.

Cavallo, Michele, "Government Employment Expenditure and the Effects of Fiscal Policy Shocks," Federal Reserve Bank of San Francisco Working Paper 2005-16, 2005.

Cullen, Joseph and Price V. Fishback, "Did Big Government Largesse Help the Locals? The Implications of WWII Spending for Local Economic Activity, 1939-1958." NBER Working Paper 12801, December 2006.

Devereux, Michael, A. C. Head and M. Lapham, "Monopolistic Competition, Increasing Returns, and the Effects of Government Spending," *Journal of Money, Credit, and Banking* 28 (1996), pp. 233–254.

Edelberg, Wendy, Martin Eichenbaum and Jonas D. M. Fisher**, "**Understanding the Effects of a Shock to Government Purchases," *Review of Economic Dynamics* 2 (January 1999): 166-206.

Eichenbaum, Martin and Jonas D. M. Fisher, "Fiscal Policy in the Aftermath of 9/11," *Journal of Money, Credit, and Banking* 37 (February 2005): 1-22.

Fatas, Antonio and Ilian Mihov, "The Effects of Fiscal Policy on Consumption and Employment: Theory and Evidence," CEPR Discussion Paper No. 2760, April 2001.

Gali, Jordi, J. David López-Salido, and Javier Vallés, "Understanding the Effects of Government Spending on Consumption," *Journal of the European Economic Association,* 5 (March 2007): 227-270.

Giavazzi, Francesco and Marco Pagano, "Can Severe Fiscal Contractions be Expansionary?" *NBER Macroeconomics Annual,* 1990.

Mountford, Andrew and Harald Uhlig, "What are the Effects of Fiscal Policy Shocks?" CEPR Discussion Paper 3338, 2002.

Pappa, Evi, "New Keynesian or RBC Transmission? The Effects of Fiscal Shocks in Labor Markets," LSE discussion paper.

Perotti, Roberto, "Estimating the Effects of Fiscal Policy in OECD Countries," CEPR Discussion Paper 4842, January 2005.

-------------------, "In Search of the Transmission Mechanism of Fiscal Policy," forthcoming *NBER Macroeconomics Annual* 2007.

Ramey, Valerie A. and Matthew Shapiro, "Costly Capital Reallocation and the Effects of Government Spending," *Carnegie Rochester Conference on Public Policy* (1997).

Rotemberg, Julio and Woodford, Michael, "Oligopolistic Pricing and the Effects of Aggregate Demand on Economic Activity," *Journal of Political Economy* 100 (1992), pp. 1153–1297.

Taylor, John B., *Macroeconomic Policy in a World Economy,* New York: W.W. Norton, 1993. (Available online at: http://www.stanford.edu/~johntayl/MacroPolicyWorld.htm)

Tenhofen, Jorn and Guntram Wolff, "Does Anticipation of Government Spending Matter? Evidence from an Expectations Augmented VAR," Deutsche Bundesbank working paper, 2007.

Whelan, Karl, "A Guide to the Use of Chain Aggregated NIPA Data," June 200 working paper.

Yang, Shu-Chun Susan, "Quantifying Tax Effects under Policy Foresight," *Journal of Monetary Economics* 52 (November 2005): 1557-1568.

Table 1. Variance Decomposition of Government Spending
Percent Explained by Shocks to Defense Spending

	VAR ordering		
Horizon (in quarters)	**Defense, state & local, total**	**State & local, defense, total**	**Federal nondefense, defense, total**
1	57	57	66
4	85	86	83
10	92	92	91
20	91	91	92

Based on estimated VARs with four lags and a time trend. All variables are log real per capita.

Table 2. Predictive Power of Ramey-Shapiro Dates

Dependent Variable: Log difference in real per capita defense spending

Variables included	**R-squared**
8 lags of log difference in real per capita defense spending	0.37
Current plus 8 lags of military date variable	0.27
Current plus 8 lags of scaled military date variable	0.57

The military date variable takes the value of 1 in 1950:3, 1965:1, 1980:1, and 2001:3. The scaled military date variable takes the values 1 in 1950:3, 0.3 in 1965:1, 0.1 in 1980:1, and 0.1 in 2001:3.

Table 3. Granger Causality Tests
1948:1 – 2003:4

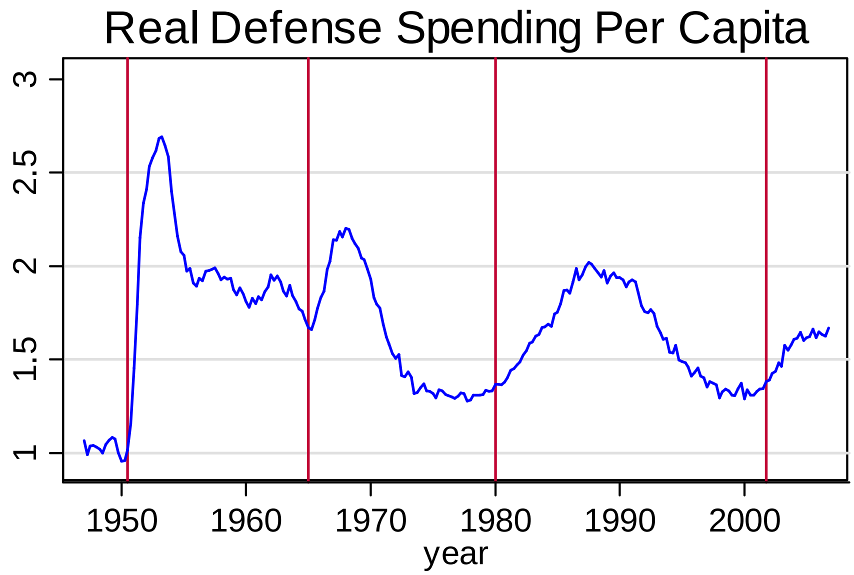
Hypothesis Tests	p-value in parenthesis
VARs with Total Government Spending	
Do War dates Granger-cause VAR shocks?	Yes (0.026)
Do 4-quarter ago Professional Forecasts Granger-cause VAR shocks? **1981:3 – 2003:4**	Yes (0.004)
Do VAR shocks Granger-cause War dates?	No (0.154)
VARs with Defense Spending	
Do War dates Granger-cause VAR shocks?	Yes (0.006)
Do VAR shocks Granger-cause War dates?	No (0.325)

VAR shocks were estimated by regressing the log of either real per capita defense or total government spending on 4 lags of itself, the Barro-Sahasakul tax rate, log real per capita GDP, log real per capita nondurable plus services consumption, log real per capita private fixed investment, log real per capita total hours worked, and log compensation in private business divided by the deflator for private business. Except for the professional forecasts, 4 lags were also used in the Granger-causality tests.

For the professional forecaster test, the VAR shock in period t is regressed on the forecast of the growth rate of real federal spending from t-5 to t, where the forecast was made at quarter t-4. The professional forecast regressions were estimated from 1981 to 2003 because of data constraints.

The war dates are a variable that takes a value of unity at 1950:3, 1965:1, 1980:1, 2001:3.

Figure 1: Real Government Spending Per Capita
(in thousands of chained (2000) dollars)



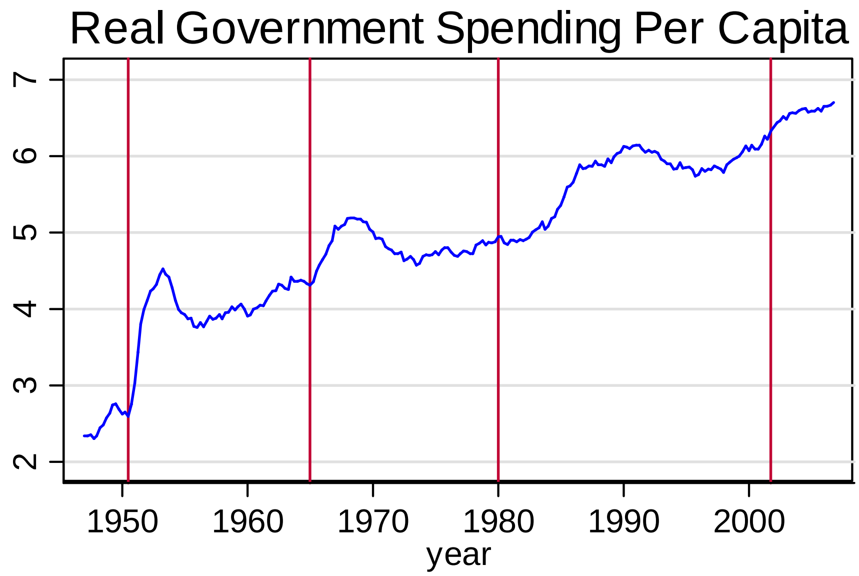


Figure 2: Historical Trends in Real Government Spending Per Capita

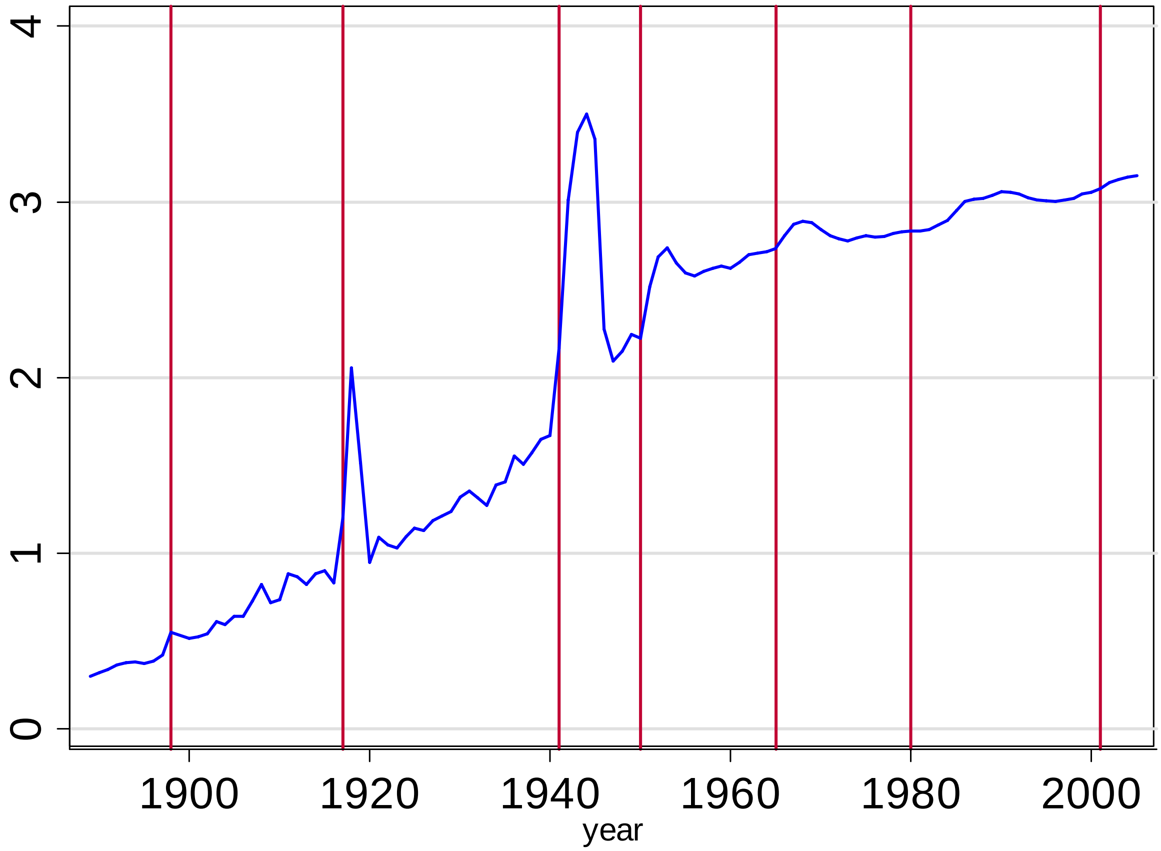


Figure 3: Components of Government Spending
Fraction of nominal GDP

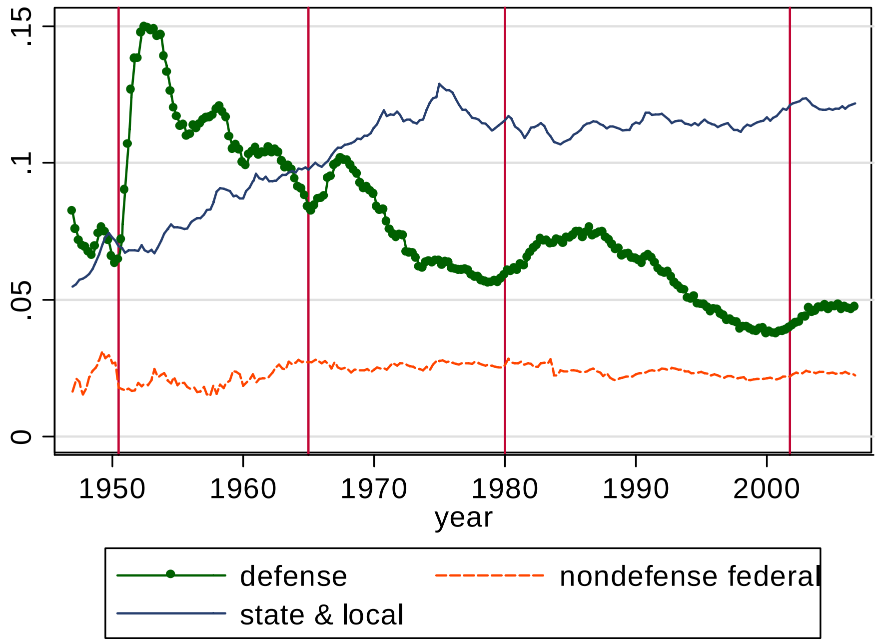


Figure 4. Government Expenditures by Function
Fraction of nominal government spending

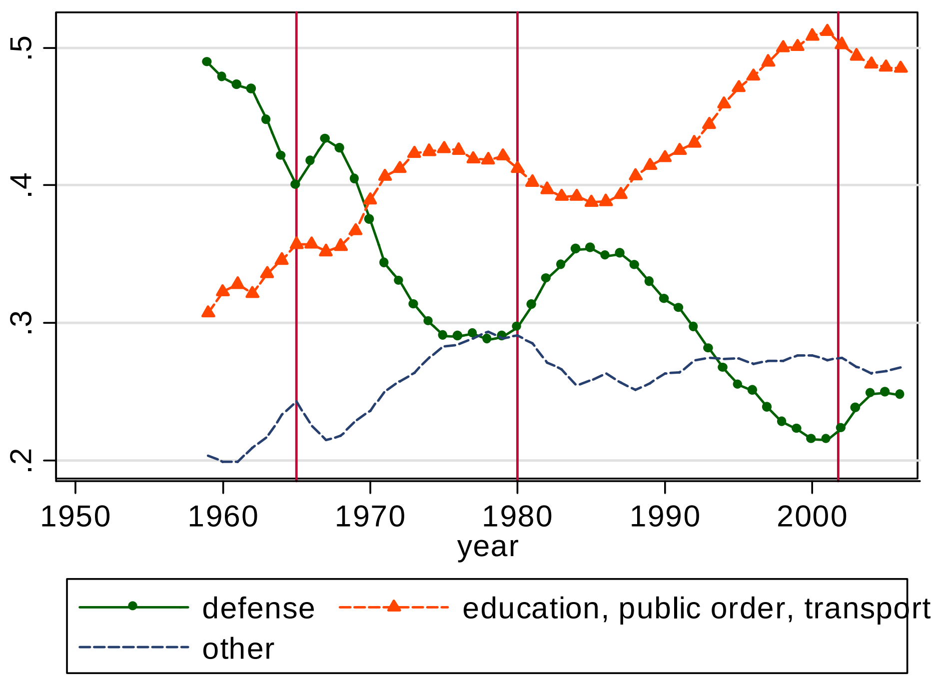


Figure 5A. Comparison of Identification Methods
Response to a government spending shock

(Standard error bands are 68% confidence intervals)

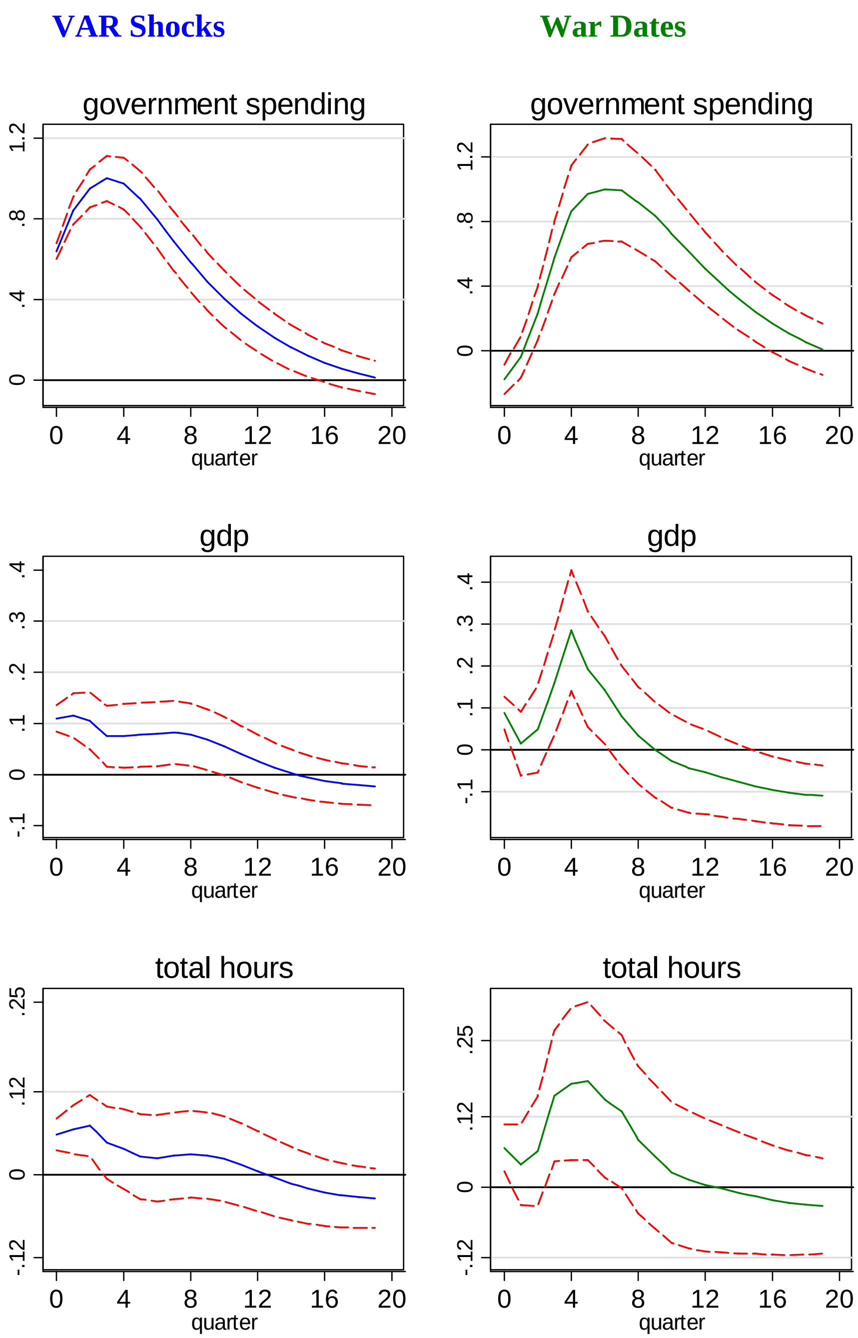


Figure 5B. Comparison of Identification Methods (continued)
Response to a government spending shock

(Standard error bands are 68% confidence intervals)

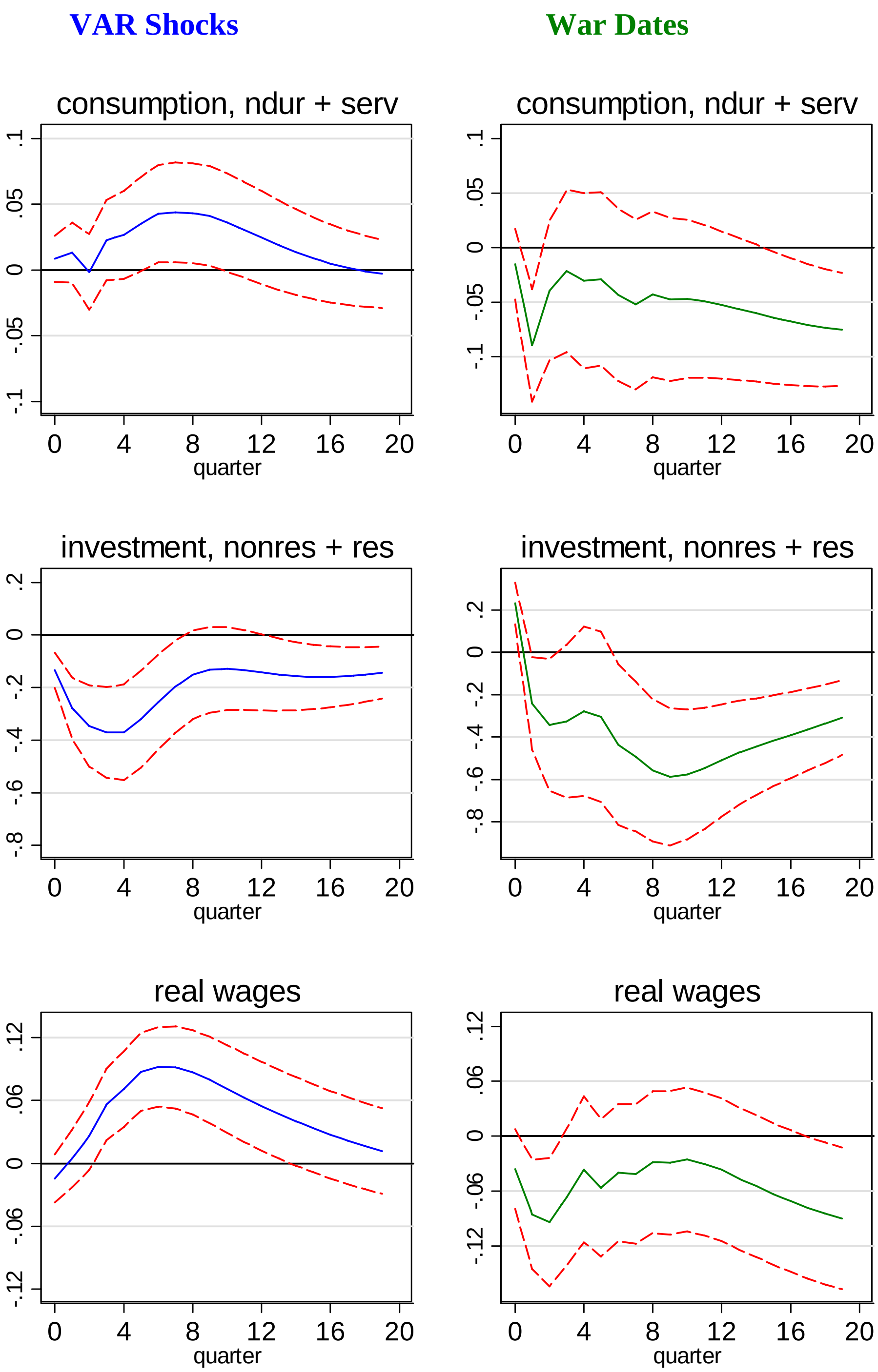
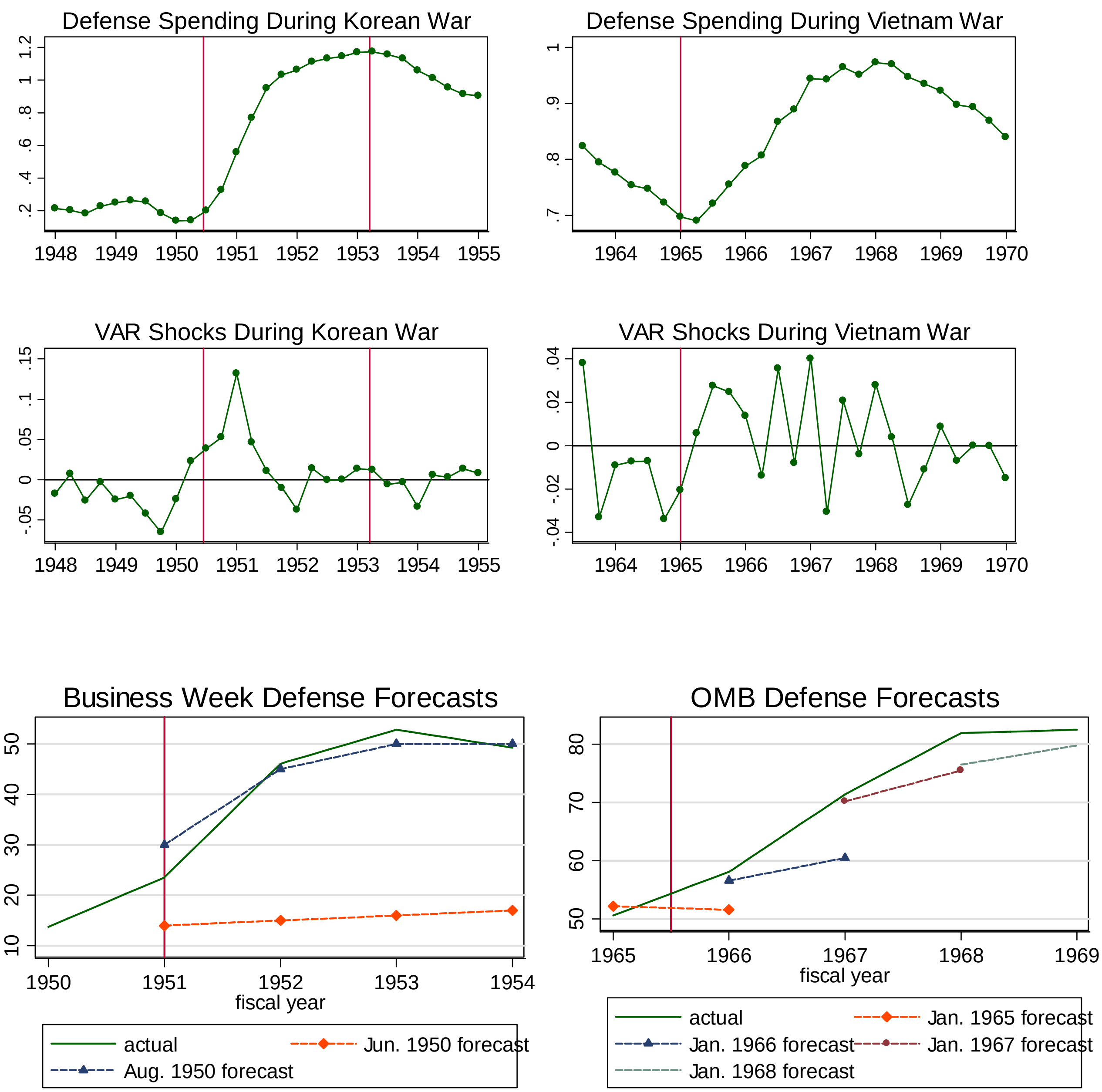
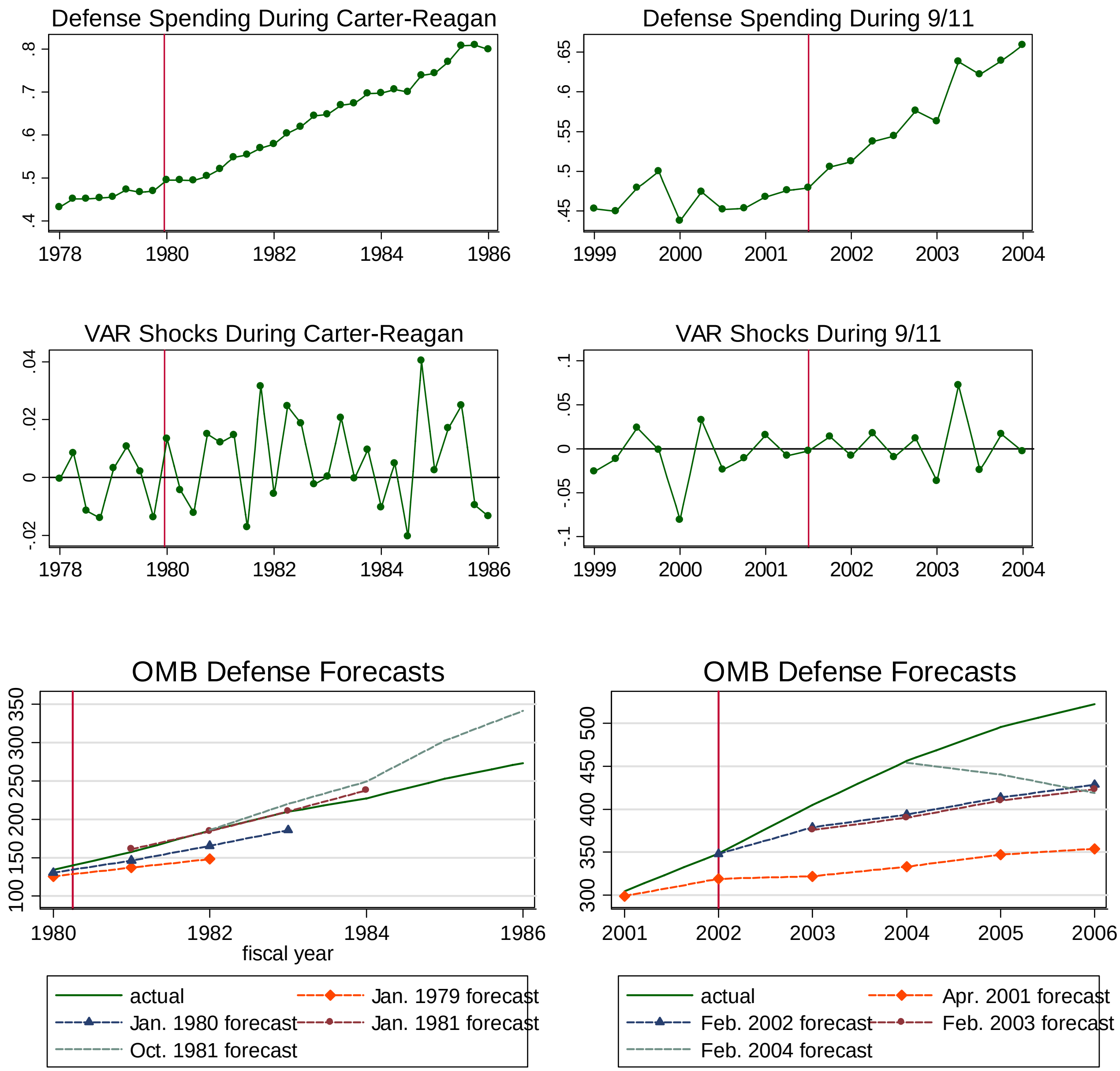


Figure 6A: Comparison of VAR Defense Shocks to Forecasts
Korea and Vietnam



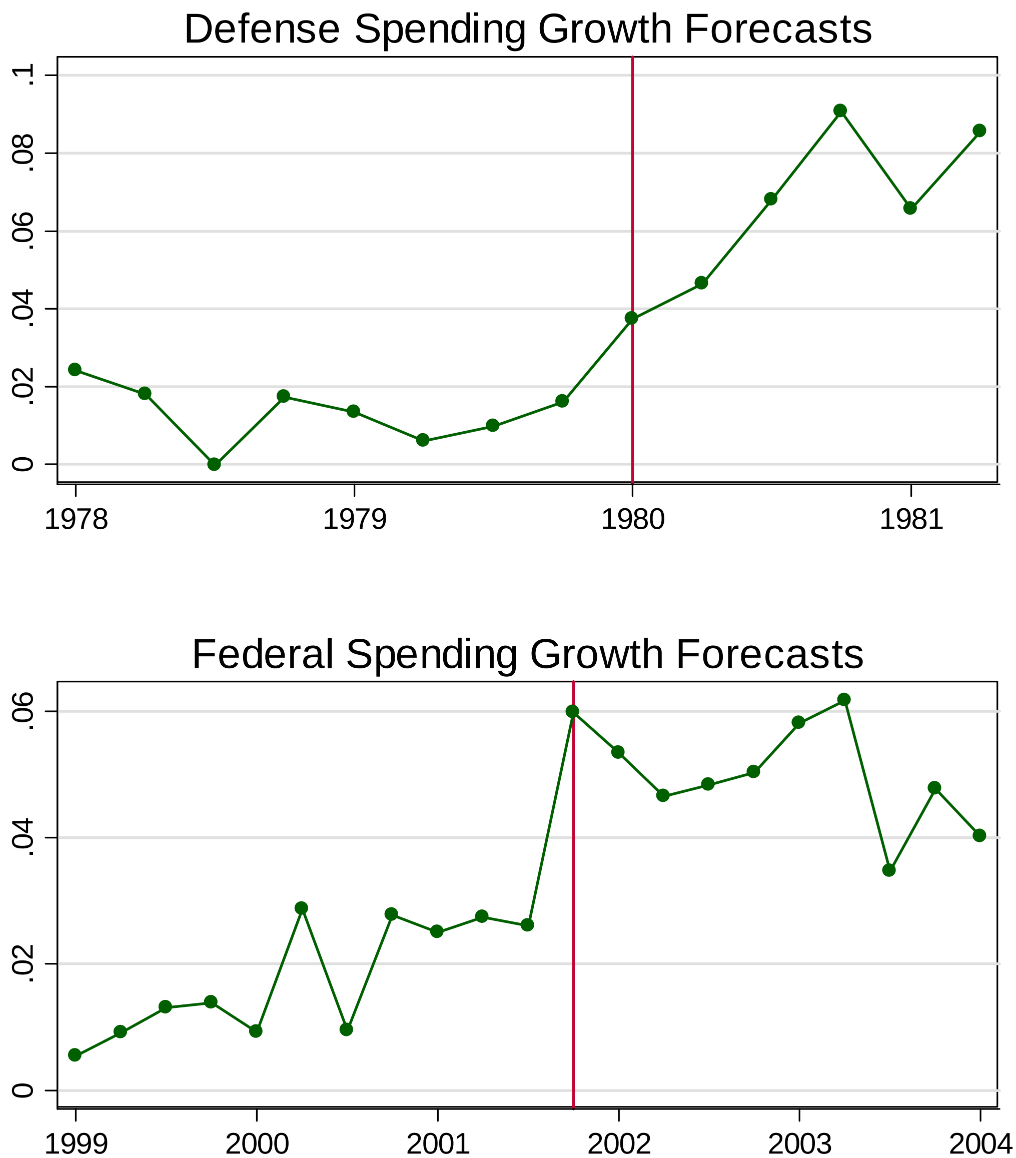
The top and middle panels are based on log per capita real defense spending on a quarterly calendar year basis. The bottom panels are nominal, annual data on a fiscal year basis.

Figure 6B: Comparison of VAR Defense Shocks to Forecasts
Carter-Reagan and 9/11



The top and middle panels are based on log per capita real defense spending on a quarterly calendar year basis. The bottom panels are nominal, annual data on a fiscal year basis

Figure 6C: Survey of Professional Forecasters Predictions



The variable shown at time t is the forecast of the growth rate of real spending from quarter t - 1 to quarter t + 4.

Figure 7. The Theoretical Effect of an Increase in Government Spending Announced Two Quarters in Advance (variables in logarithms)

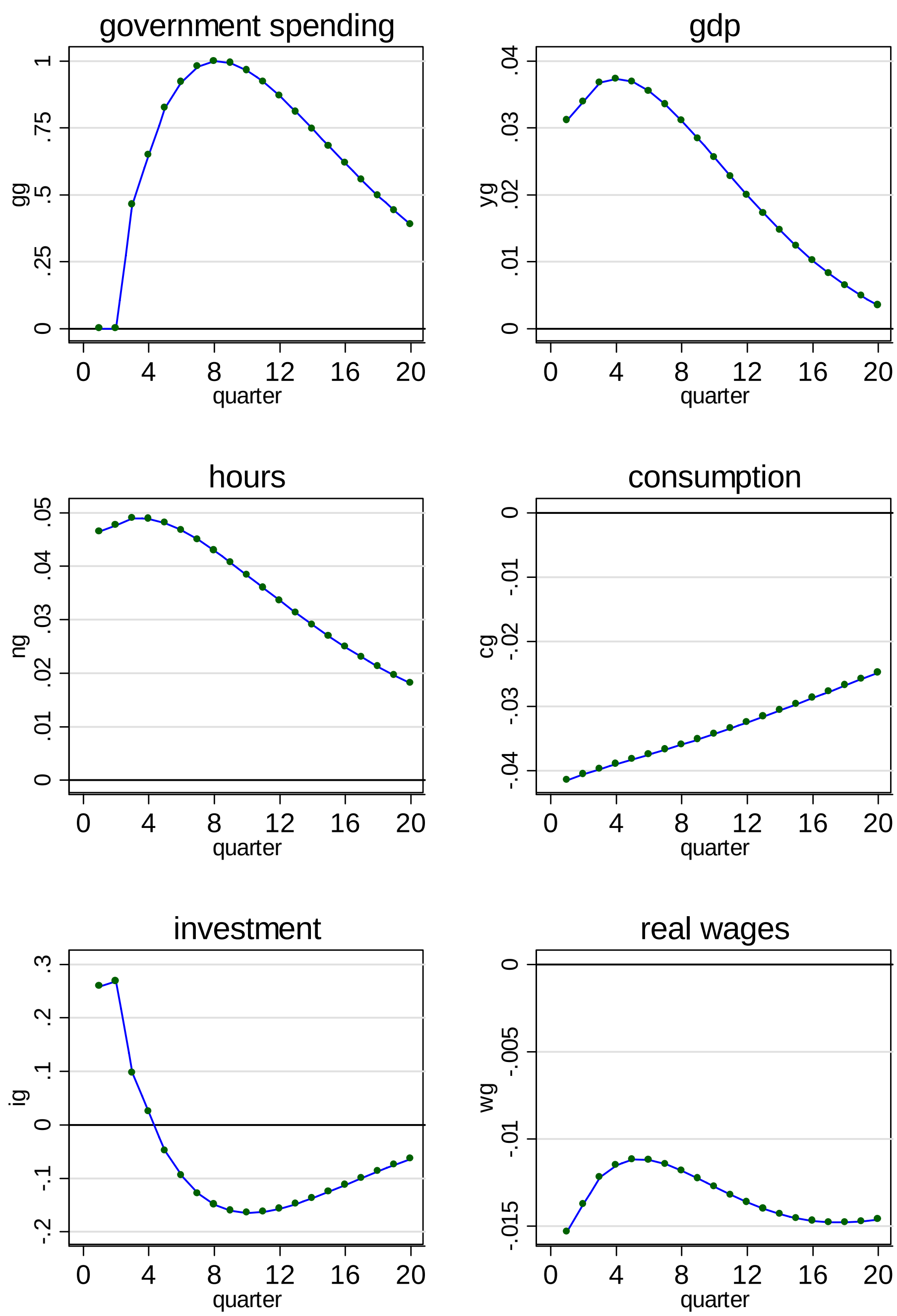


Figure 8A. The Effect of Missing the Timing

Based on trivariate VARs on model simulations in which government spending changes are anticipated two quarter in advance

Faulty Timing **True Timing**

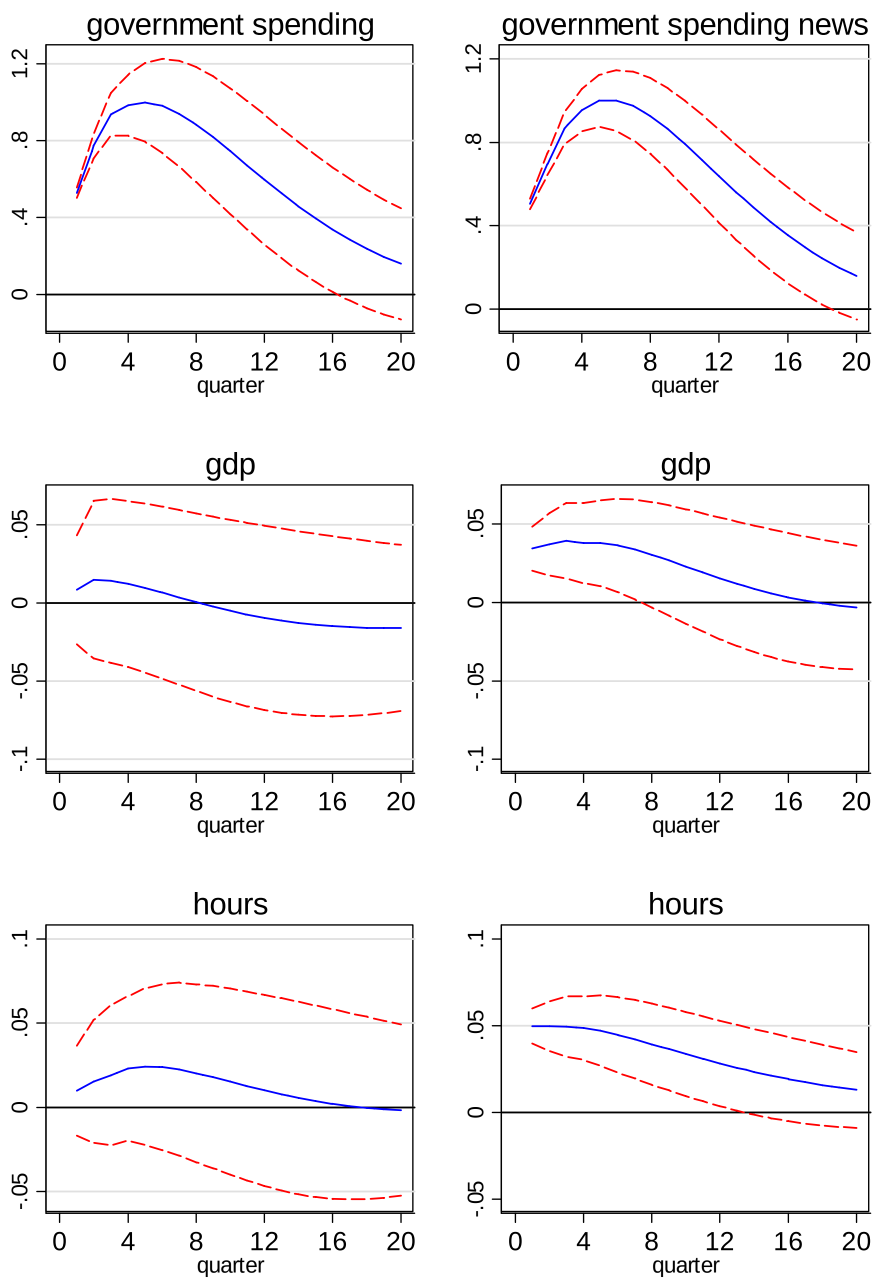


Figure 8B. The Effect of Missing the Timing

Based on trivariate VARs on model simulations in which government spending changes are anticipated two quarter in advance

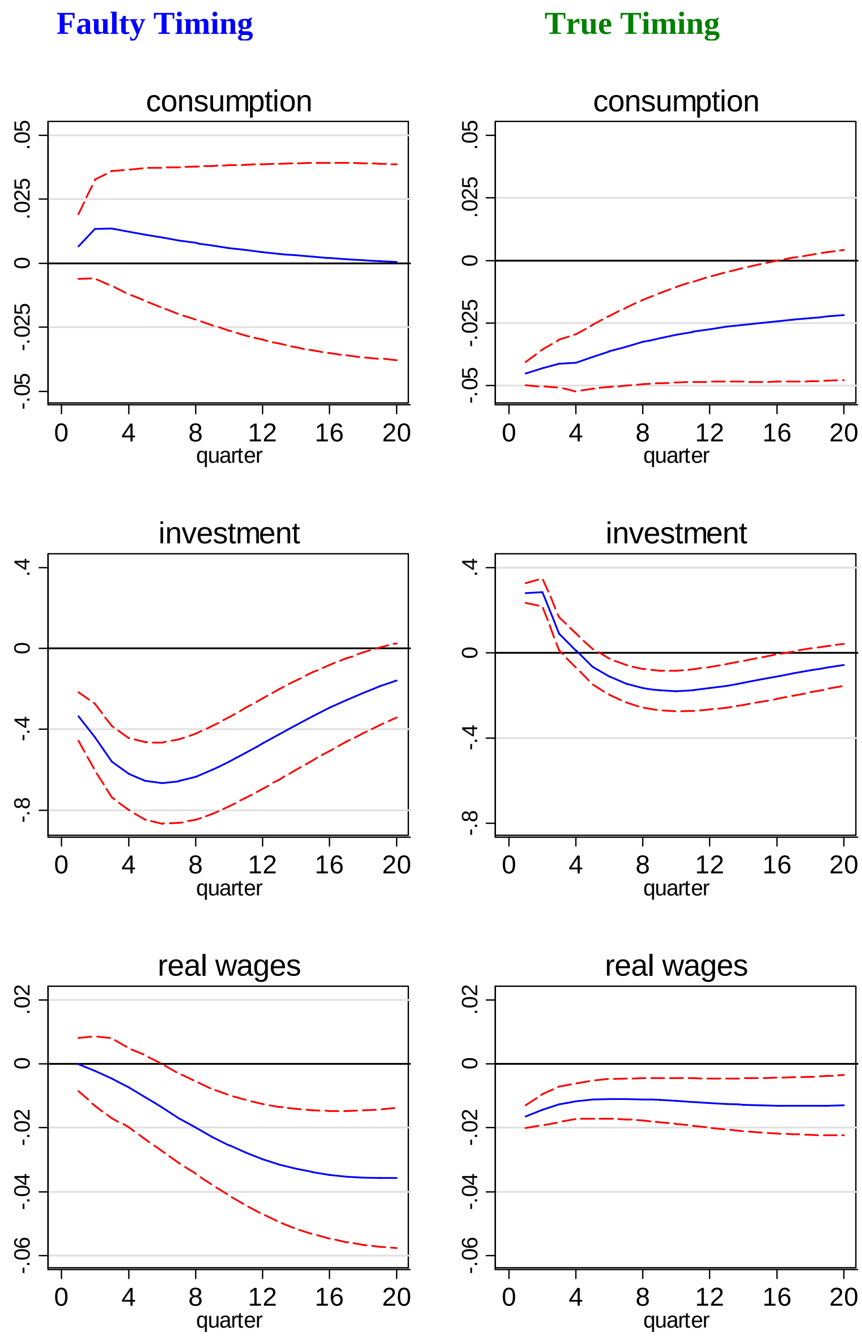


Figure 9. The Effect of Mistiming the Ramey-Shapiro Dates
(The Ramey-Shapiro dates are shifted later to coincide with
the large positive VAR shocks in 1951:1, 1965:3 1980:4 2003:2)

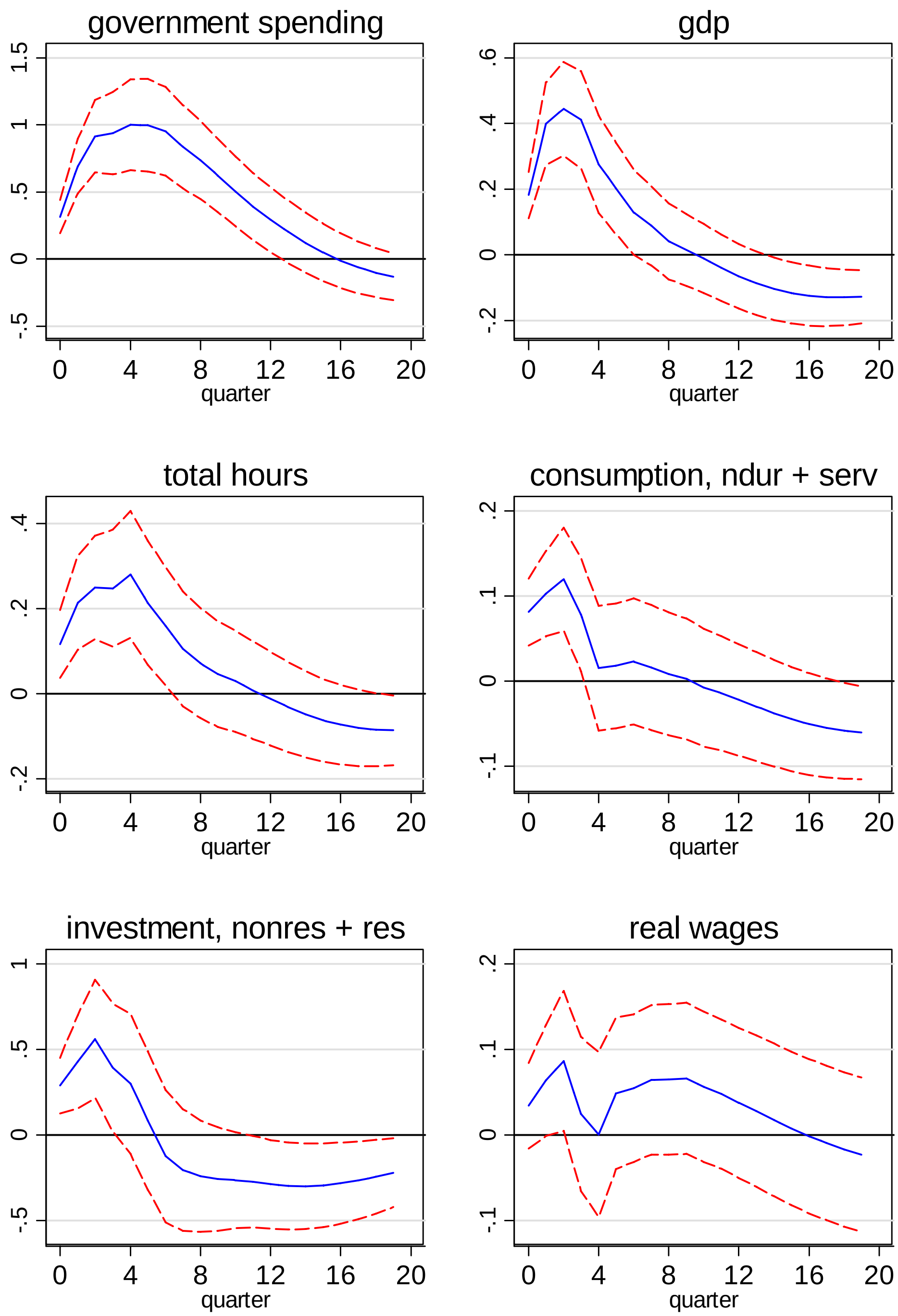


Figure 10. The Effect of Using Defense Spending Shocks in a Standard VAR

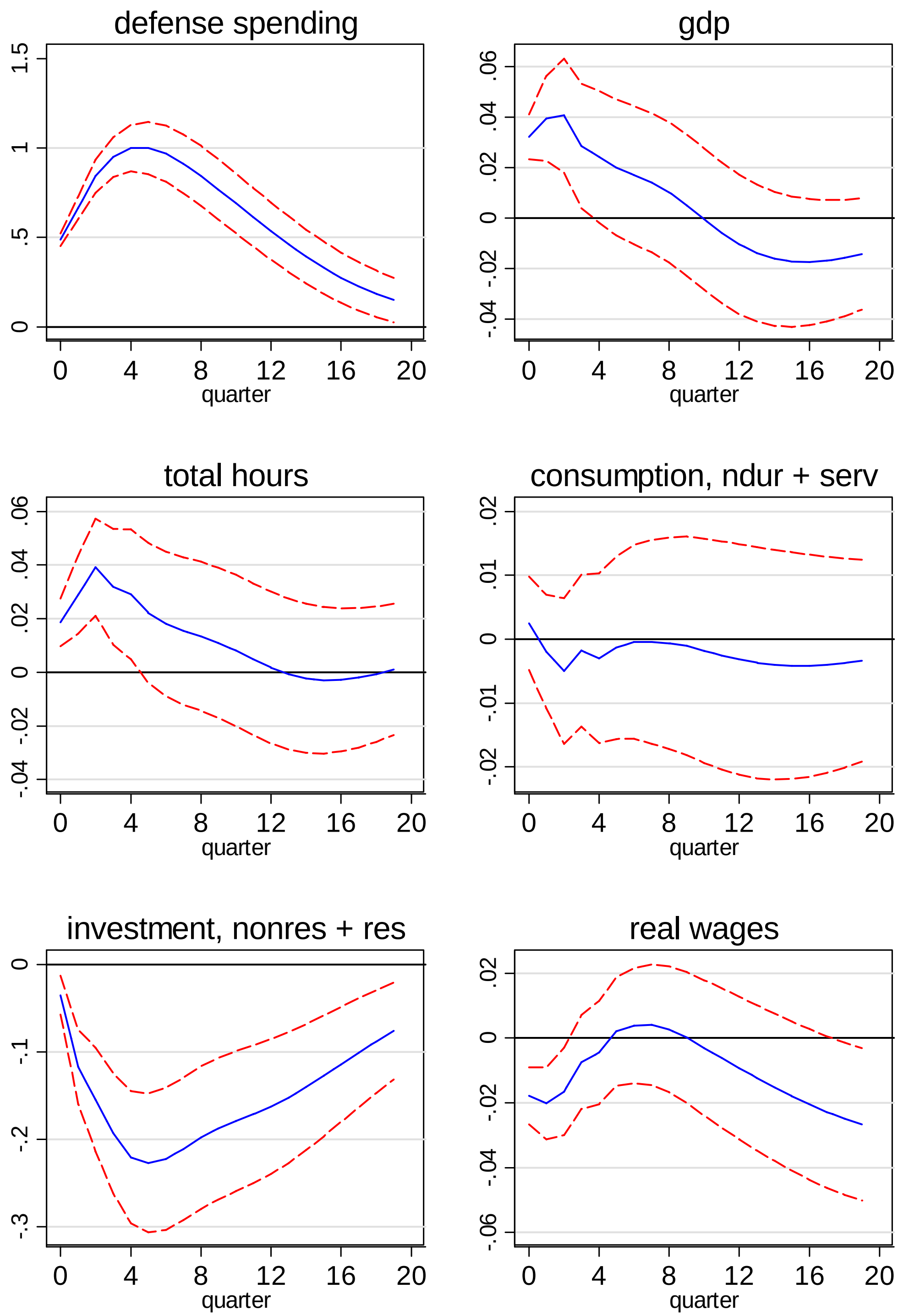


Figure 11. The Effects of Government Spending Shocks on Taxes and Interest Rates

(Standard error bands are 68% confidence intervals)

VAR Shocks **War Dates**

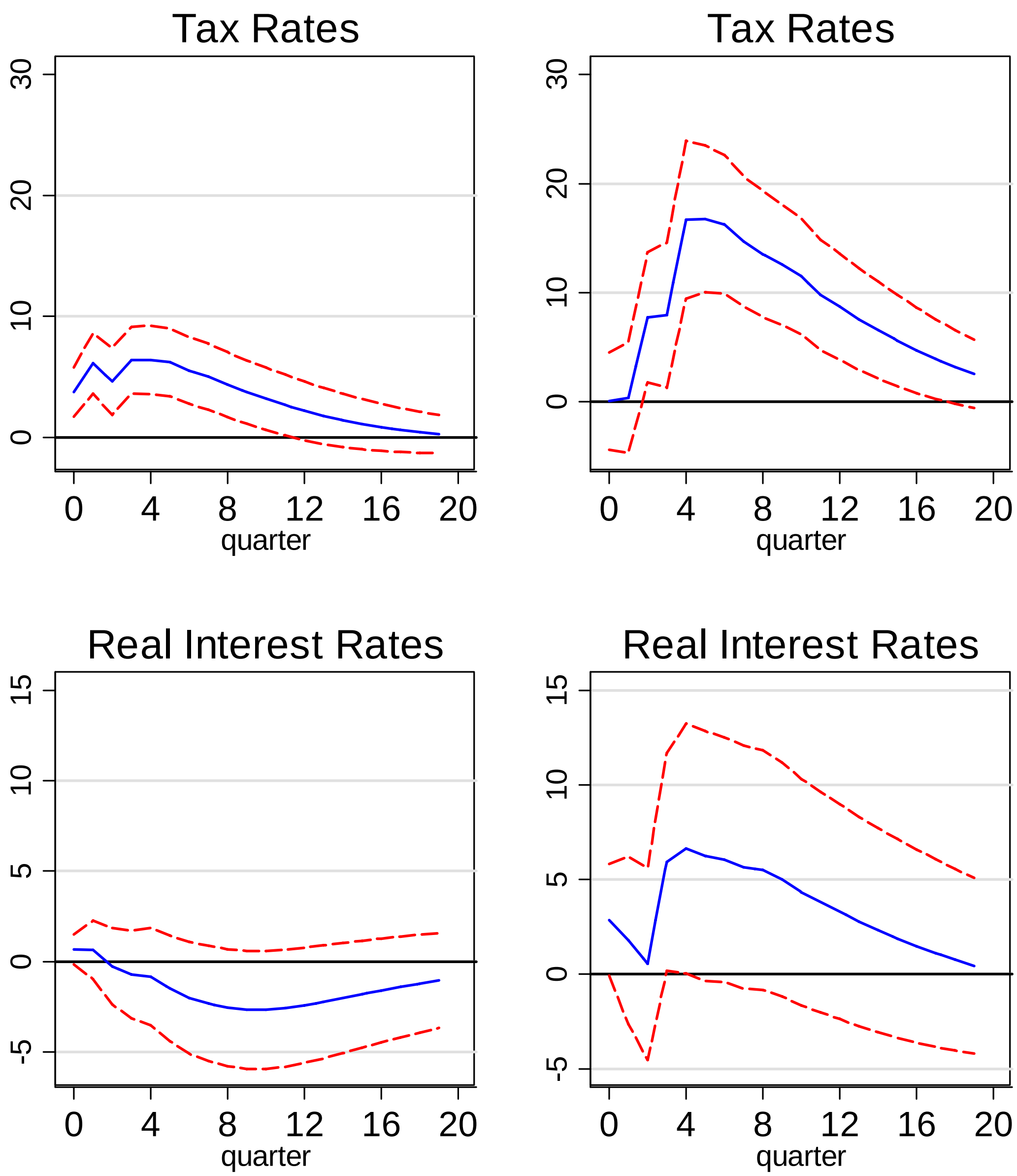


Figure 12. Evidence that Consumers React Quickly to Expectations

The Behavior of Per Capita Consumer Durable Expenditures at the onset of the Korean War When Consumers Feared Rationing