Volume Title: NBER Macroeconomics Annual 2007, Volume 22

Volume Author/Editor: Daron Acemoglu, Kenneth Rogoff and Michael Woodford, editors

Volume Publisher: University of Chicago Press

Volume ISBN: 978-0-226-00202-6

Volume URL: http://www.nber.org/books/acem07-1

Conference Date: March 30-31, 2007

Publication Date: June 2008

Chapter Title: Comment on "In Search of the Transmission Mechanism of Fiscal Policy"

Chapter Author: Valerie A. Ramey

Chapter URL: http://www.nber.org/chapters/c4086

Chapter pages in book: (p. 237 - 246)

Comment

Valerie Ramey, University of California, San Diego and NBER

1 Introduction

Roberto Perotti's paper is an ambitious effort to shed light on why different empirical methods lead to varying conclusions about the effects of government spending on key economic variables. Perotti's previous work used vector autoregressions (VARs) and found old-fashioned Keynesian effects of government spending. In contrast, my previous work used dates of exogenous military buildups based on narratives and found neoclassical effects of government spending. Numerous others have used these two types of methods, with similar results. In this current paper, Perotti critiques the military date method, answers critiques I have made of the VAR method, and uses the VAR identification method to estimate multiple variations over different time periods and countries. He also analyzes industry-level input-output data.

I wholeheartedly agree with Perotti that the question asked in this paper is crucial to our basic understanding of how the economy works. Ex ante, it seems that a consensus could easily be reached, since unlike the case of technology shocks we actually have data on government spending. Unfortunately, finding a consensus on this issue is proving to be just as elusive as in the case of technology shocks. I do not think this paper will settle the issue, for while I applaud his efforts to try to find robust results, I will ultimately disagree with his conclusion.

In my discussion, I will begin by providing some background to help readers who have entered this debate midstream. I will then argue that it makes sense to focus on defense spending rather than other components of government spending. I will then assess Perotti's critique of my method as well as assess his defense of his method. I will argue that both methods have weaknesses, but that the weaknesses of the VAR approach are potentially more problematic. In particular, I highlight results from a new paper showing that allowing for differences in timing in the VAR reverses the consumption results. Finally, I will show that his input-output results are due entirely to the classification of two computer industries in defense. Thus, while impressed with the volume of new results, I do not find them to be convincing evidence for the presence of old-fashioned Keynesian-style effects of government spending.

2 Background

The old-fashion Keynesian model predicts that an increase in government spending will lead to a rise in output, hours, consumption, and real wages. The neoclassical model offers the same predictions for output and hours, but disagrees on the predictions for consumption and real wages. In particular, the neoclassical model predicts that the same negative wealth effect that raises labor supply should also lower consumption. Moreover, the increase in labor supply should be accompanied by a lower product wage in the short run (Baxter and King [1993]). Interestingly, the new Keynesian model also predicts that consumption should fall, since in that model agents are assumed to choose consumption optimally. It is only when one returns to an old Keynesian assumption of "rule-of-thumb" consumers that one can obtain an increase in consumption in the model (Galí, López-Salido, and Vallés [2007]).

Ramey and Shapiro (1998) estimated the effects of government spending by identifying exogenous and unanticipated increases in military spending. To do this, they used a narrative approach and chose dates when *Business Week* suddenly predicted a large increase in government spending due to exogenous political events. The dates chosen were the Korean War in 1950, the Vietnam War in 1963, and the Soviet invasion of Afghanistan at the end of 1979. Using a dummy variable that took the value of unity for each of these dates, Ramey and Shapiro found that the increase in government spending raised output, hours, nonresidential investment, interest rates, and the relative price of manufactured goods, while it lowered consumer durable purchases, nondurable consumption, residential investment, and various measures of real wages. Subsequently, Edelberg, Eichenbaum, and Fisher (1999), Burnside, Eichenbaum, and Fisher (2004), Cavallo (2005), and Eichenbaum and Fisher (2004) have used these military dates in a variety of specifications and robustness tests and have found similar results. More recent work has also added 9/11 as a fourth date (e.g., Eichenbaum and Fisher [2004] and Ramey [2006]).

Blanchard and Perotti (2002) used a VAR approach instead to identify government spending shocks. While they used a structural VAR method for identifying tax shocks, their government spending shocks were identified using a standard Cholesky decomposition in which it was assumed that government spending did not react within the quarter to GDP, and so on. Perotti uses the same technique in the current paper and finds results similar to the earlier ones. In this framework, an increase in government spending raises both nondurable consumption and real wages.

What accounts for the different results? The military dates method differs from the VAR shocks in at least four ways: (1) the military dates are few in number; (2)they capture only one type of government spending—military spending—which is widely and nationally perceived; (3) they apply only to increases in government spending, not decreases; and (4) their timing is based on the historical record of news rather than on when the spending actually occurs. Most of Perotti's critiques of the military date method center on (1). My critique of the VAR method has previously focused on point (4) (Ramey [2006]). Thus, a good part of Perotti's current paper addresses these two points. I will also discuss point (2) as another potential crucial difference between the two methods.

3 Why We Should Focus on Defense Buildups

As discussed before, the military date methods focuses only on defense spending, whereas the recent VAR methods have focused on total government spending or government consumption spending. I will present two arguments for why it makes sense to focus on military spending.

First, to distinguish theories, it is very important that we consider types of government spending that do not affect the aggregate production function. If government spending is socially productive, then the net effect may be a positive wealth effect, which can increase consumption and real wages even in the neoclassical model. Military spending, particularly in the U.S. context, is ideal, since it has negligible direct effects on the aggregate production function. What other types of spending does the government undertake? Table 3C1.1 shows the composition of government purchases by function in 1959 and 2005 (earlier data are not available). Defense spending was almost 50 percent of total government purchases in 1959 and 25 percent in 2005. Perotti uses government consumption expenditures plus defense. As the table makes clear, fed-

Table 3C2.1
Consumption of Government Purchases (Percent, Nominal Terms from NIPA Table 3.17)

By Function:	1959	2005
Defense	49	25
Nondefense Investment	14	13
Federal Nondefense Consumption	9	11
State and Local Nondefense Consumption	28	51
S&L Consumption Spending on Education	13	24
S&L Consumption Spending on Public Safety, Health, and Economic Affairs	11	16

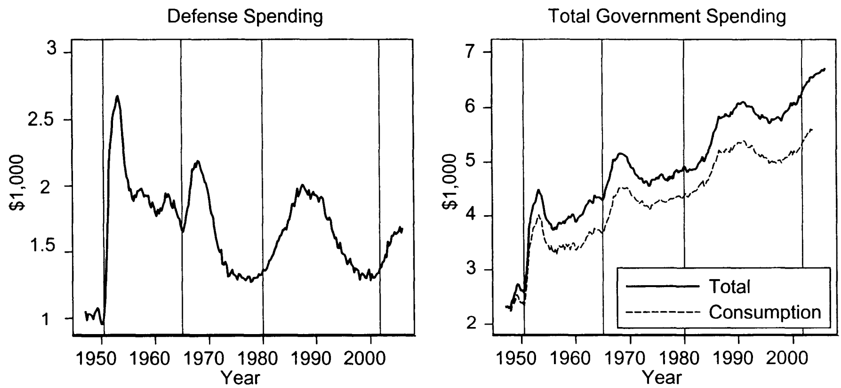

Figure 3C2.1
Real Per Capita Government Spending

eral nondefense consumption is a trivial part of government spending. On the other hand, state and local nondefense consumption is more important, equal to 51 percent in 2005. Moreover, most of the spending is on public education, health, and public safety. It is very possible that these other categories of government spending have direct productive effects on the economy, and hence are not the types of shocks we seek in order to distinguish models.

Military spending also accounts for the bulk of fluctuations in government spending. Figure 3C2.1 plots per capita defense spending as well as per capita total government spending and consumption government spending. Except for an upward trend, almost all of the variation in government spending is due to military buildups. The vertical lines denote the Ramey-Shapiro dates, augmented with 9/11. Defense spending goes up dramatically after each date, with Korea being by far the biggest shock in the post-WWII period. The defense buildups continue to be evident in total government spending, shown in the graph on the right. The other fluctuations are minor blips in comparison.

Based on this evidence, I am led to two conclusions. First, the military buildups are the major source of shocks to total government spending. Korea is particularly important and should not be omitted. Eliminating the Korean War period from a study of the effects of government spending shocks makes as much sense as eliminating the 1990s from a study of the effects of information technology. Second, most of the other fluctuations represent state and local spending on possibly productive activities such as education, and thus are not the sorts of shocks we should be studying for the exercise at hand.

4 Perotti's Critique of the Military Date Method

Perotti's principle critique of the military date method, which uses current and lagged values of one set of dummy variables for the four major military buildups, is twofold. First, he argues that the methodology imposes the constraint that the dynamic responses of the endogenous variables must be the same across all four periods. He argues that each episode may have been associated with different tax policies. Indeed, his critique is potentially valid but *it applies equally to his preferred VAR method.* Both methods impose the constraint that the dynamic response of the endogenous variables to shocks is the same across the sample. Both methods study *the average response* to a shock. Thus, this critique cannot be used as a basis for preferring the VAR methodology over the military date methodology.

Perotti then goes on to estimate a new specification of the military date method (DV3) that allows the dynamics to change across the four buildups. He does this by defining four separate dummy variables. He then criticizes this new proposed method, appropriately so, by arguing that it assumes that the shock to each variable during the start of a military buildup consists entirely of the military buildup and allows no shocks from other sources. Also, as he points out, this method is very imprecise. I agree with his critique of his own new DV3 method. The standard method is preferable because it does not suffer from this problem.

To see an example of the problems associated with Perotti's DV3 procedure, consider his analysis of the behavior of taxes, which he claims supports his conclusions. Using his DV3 method, he finds that tax rates declined during the Vietnam War (figure 3.2). One has only to look at the average marginal tax rate series from figure 8 in Edelberg, Eichenbaum, and Fisher (1999) or the capital and labor income tax rates in figure 1 of Burnside, Eichenbaum, and Fisher (2004) to see that Perotti's DV3 procedure gives counterfactual results.

To summarize, Perotti's critique of the military date method also applies to the VAR method and thus does not cause us to prefer one method over the other. His proposed alternative DV3 method is even more problematic.

5 Potential Timing Problems with the VAR Method

In Ramey (2006), I argued that the differences in results from the VAR method and the military date method could be due to timing. To see this, suppose that the government announces an increase in government spending a few periods before it actually occurs. According to the neoclassical model, consumption and the real wage should fall immediately, and then recover partially. If one uses a VAR to identify the shock from the actual path of spending, then one only sees consumption and the real wage when they are rising after the initial fall. I supported these arguments both with specific examples of the lags in government spending after the decision to build up the military and with Granger causality tests.

In this paper, Perotti tries to respond to this critique. First, he shows that when the Korean War period is eliminated the military dates no longer Granger-cause the VAR shocks. My response is two-fold. First, eliminating the Korean War period simply reduces the power of the tests. Second, even without the Korean War period, the military dates still Granger-cause the defense spending component of the VAR shocks. Perotti's second response is to use the Congressional Budget Office (CBO) forecasts. This idea is nice in theory, but has little power in practice because these forecasts (1) are available only from 1984 on; (2) are only available semiannually; and (3) only apply to federal spending. Thus, the finding that they do not Granger-cause the VAR residuals is very weak evidence in favor of the VAR.

I was surprised that Perotti did not try to defend the VARs by using the Blanchard and Perotti (2002) robustness check for the timing issue. In their earlier paper, Blanchard and Perotti allowed for the possibility that the shocks identified by the VAR were known one period in advance. However, they only applied this check to the effect on output. What happens if one applies this check to consumption? A recent paper by Tenhofen and Wolff (2007, abstract) does just this and finds that "consumption falls in reaction to an expenditure shock once the model allows for one-period ahead anticipation of this shock." Thus, it appears that the inability of the standard VAR to get the timing right is the source of the finding of the increase in consumption. This result leads us to doubt all of the VAR results he shows that claim that a government spending shock leads to an increase in consumption. These apply to the post-WWII quarterly results, the annual historical results, as well as the international results.

To summarize, Perotti's defense against the timing issue is rather weak, and Tenhofen and Wolff's application of Blanchard and Perotti's own method shows that changing the timing in the VAR by just one quarter reverses the consumption results.

6 Industry Evidence from Input-Output Tables

In section 8, Perotti turns to industry evidence to distinguish whether government effects are more neoclassical or Keynesian. To see why industry-level evidence can shed light on this issue, consider the demand for labor curve for industry i:

$$\frac{W_{it}}{P_{it}} = A_{it} F_L(L_{it}, K_{it})$$

where W/P is the real product wage, A is total factor productivity (TFP), L is labor input, K is the capital stock, and F_L is the marginal product of labor. According to the neoclassical demand curve, if A and K are held constant, then an increase in L must be accompanied by a decrease in the real product wage.

Perotti seeks to cast doubt on the neoclassical production function by showing that both hours and real wages increased in the industries that received the greatest increase in defense spending during the Vietnam War and the Carter-Reagan buildup. He shows that most of the top ten receiving industries experience an increase in both hours and the real product wage during the buildups. These results stand in contrast to some of my previous findings on the effects of defense spending on real wages. For example, figure 3C2.2 is a modification of a graph shown in Barth and Ramey (2001). It shows that real product wages fell in the aircraft industry just as defense spending was increasing dramatically.

Why are Perotti's results so different? There are several problems with his analysis. First, his timing for the Carter-Reagan buildup is wrong. He

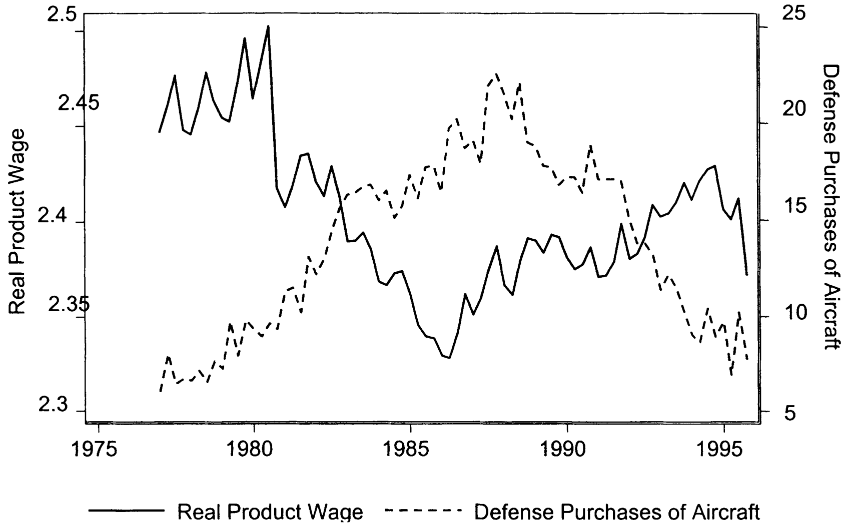

Figure 3C2.2
Defense Purchases and Real Product Wages in Aircraft

compares 1977 to 1982. In fact, most of the increase in military prime contracts occurred after 1982, which can also be seen in the Barth-Ramey graph. Second, it is not surprising that real wages in defense industries would increase, even with an increase in hours. As the previous equation shows, an increase in both hours and the real wage contradicts neoclassical theory only if both A and K are constant. However, from 1963 to 1967 real wages grew 12 percent in all private business. As long as A and K were increasing at a similar rate to the rest of the economy, then an increase in real wages is not inconsistent with the neoclassical view.

The third, and most important reason, for the difference appears to be due to the classification of two key industries. In particular, Perotti classifies semiconductors and electronic computing equipment as industries in which the change in government spending was a big share of total shipments. These two industries, however, are huge outliers among all of the industries. In particular, while the average annual growth rate of labor productivity from 1958 to 1996 was 2.5 percent for all industries in the NBER database, the annual growth rate of labor productivity was 16 percent for semiconductors and almost 19 percent for electronic computing equipment! It is clear that the unusually high rate of technological progress in these industries would cause real wages to increase even if the government suddenly raised its demand for their products.

The extraordinary behavior of these two industries turns out to account for *all* of Perotti's results. To demonstrate this, I performed the following experiment. Using the same NBER productivity database used by Perotti, I defined a defense dummy variable equal to unity if the industry was listed in Perotti's table 3.4. I then ran regressions of annual growth rates of hours, capital, and real product wages on the defense dummy variable, as well as year fixed effects, for the buildup periods 1965–1967 and 1980–1987. I found that the defense industries had significantly higher growth of hours, capital, and real wages, consistent with Perotti's finding. However, when I redefined the defense dummy variable to omit semiconductors and electronic computing equipment, I found that while hours and capital growth rates were still significantly higher in defense industries, real product wage growth was not. Thus, classifying the two extremely high productivity industries as defense accounts for Perotti's findings.

7 Conclusions

Neither of the leading methods for identifying government spending shocks is without its flaws. The military date method suffers from few episodes, whereas the VAR method suffers from potential timing problems. Despite a large set of empirical results, the present paper does not resolve the issue. Unfortunately, the critique of the military date method offered by Perotti applies equally to the VAR method. I have argued that his proposed new DV3 method is even more problematic than the standard method. The defense of the timing of the VARs is weak, and the new results by Tenhofen and Woolf (2007) cast doubt on every VAR estimated in this paper. The input-output analysis is promising, but my investigation suggests that the results are due to the extraordinary productivity growth experienced by the two computer industries he included in the defense category.

References

Barth, Marvin J., and Valerie A. Ramey. The cost channel of monetary transmission. *NBER Macroeconomics Annual 2001*, ed. B. Bernanke and K. Rogoff, 199–239. Cambridge: MIT Press.

Baxter, Marianne, and Robert King. 1993. Fiscal Policy in General Equilibrium. *American Economic Review* 83 (3): 315–34.

Blanchard, Olivier, and Roberto Perotti. 2002. An empirical characterization of the dynamic effects of changes in government spending and taxes on output. *Quarterly Journal of Economics* 117 (4): 1329–68.

Burnside, Craig, Martin Eichenbaum, and Jonas D. M. Fisher. 2004. Fiscal shocks and their consequences. *Journal of Economic Theory* 115:89–117.

Cavallo, Michele. 2005. Government employment expenditure and the effects of fiscal policy shocks. Unpublished working paper, Federal Reserve Bank of San Francisco.

Edelberg, Wendy, Martin Eichenbaum, and Jonas D. M. Fisher. 1999. Understanding the effects of a shock to government purchases. *Review of Economics Dynamics* 2 (1): 166–206.

Eichenbaum, Martin, and Jonas D. M. Fisher. 2004. Fiscal policy in the aftermath of 9/11. Forbearance of Chicago Working Paper no. 2004-06. Chicago: Federal Reserve Bank.

Galí, Jordi, David López-Salido, and Javier Vallés. 2007. Understanding the effects of government spending on consumption. *Journal of the European Economic Association* 5 (1): 227–70.

Ramey, Valerie A. 2006. Identifying government spending shocks: It's all in the timing. Mimeograph, University of California at San Diego.

Ramey, Valerie A., and Matthew D. Shapiro. 1998. Costly capital reallocation and the effects of government spending. *Carnegie Rochester Conference on Public Policy* 48:145–94.

Tenhofen, Jörn, and Guntram B. Wolff. 2007. Does anticipation of government spending matter? Evidence from an expectation-augmented VAR. Deutsche Bundesbank Working Paper no. 14/2507.