Volume Title: NBER Macroeconomics Annual 2012, Volume 27

Volume Author/Editor: Daron Acemoglu, Jonathan Parker, and Michael Woodford, editors

Volume Publisher: University of Chicago Press

ISSN: 0889-3365

Volume ISBN: cloth: 978-0-226-05277-9; paper: 978-0-226-05280-9, 0-226-05280-X

Volume URL: http://www.nber.org/books/acem12-2

Conference Date: April 20-21, 2012

Publication Date: May 2013

Chapter Title: Comment on "Roads to Prosperity or Bridges to Nowhere? Theory and Evidence on the Impact of Public Infrastructure Investment"

Chapter Author(s): Valerie Ramey

Chapter URL: http://www.nber.org/chapters/c12752

Chapter pages in book: (p. 147 - 153)

Comment

Valerie Ramey, *University of California, San Diego, and NBER*

Introduction

This paper fills a large gap in the literature by studying multipliers for an important component of government expenditures: infrastructure investment. The debate over the stimulus package at the start of the Great Recession revealed the fact that there was no consensus on the value of the multipliers for various types of government spending. Although most of the research had focused on either total government purchases or defense purchases, much of the stimulus package took the form of transfers and government infrastructure spending, two categories of expenditures which had received scant attention in the multiplier literature. Thus, the present paper with its focus on estimating multipliers for infrastructure spending is a valuable addition to the literature.

This paper offers numerous contributions. First, the authors have assembled a very rich state panel data set that can be used to study the effects of state-level infrastructure spending. Second, the authors take the lags between grants, obligations, outlays, and the possible impact of anticipations seriously and construct a very plausible measure of shocks. This attention to detail is important because the literature has shown that endogeneity, lags, and anticipations can have important effects on inference. Third, the authors offer insight into numerous empirical aspects and highlight subtle empirical issues that have been overlooked by the literature. Finally, the authors present a useful theoretical model that is qualitatively consistent with their empirical results. This model is very useful for interpreting some of the patterns in the data.

The authors find that their identified shock to the present discounted value of federal infrastructure grants leads to an increase in state GDP

978-0-226-05277-9/2013/2012-0203$10.00

on impact and then again six to eight years out. By ten years out, there is no lasting effect. The pattern is similar for state income, employment and wages, and salaries. The implied multiplier is 3 on impact and as high as 8 six to eight years out. The average over a 10 year horizon, however, is two.

As pointed out in Francesco Giavazzi's discussion, the state-level estimates cannot directly answer the question of interest for aggregate policy because the state fixed effects net out the impact of current or future expected tax increases to finance the spending. Thus, a theoretical model is necessary for interpreting the results. Using an open-economy model to capture their state estimates, the authors find that they can generate a state multiplier of 0.3 on impact and as high as 2 after six to eight years. They are also able to use the model to show that the aggregate multiplier is roughly half of the local multiplier.

The bulk of my discussion offers a few caveats to the conclusions that Leduc and Wilson draw from their empirical analysis. Because they focus mostly on the peak multipliers for GDP, Leduc and Wilson do not discuss some of the estimated negative short-run effects implied by their empirical analysis. Since infrastructure spending has been widely discussed in the debates about short-run stimulus packages, it is important also to highlight the implications of the empirical results for the short run, rather than focusing only on results six or more years out.

The next section discusses a potentially important, but overlooked, possible role of credit markets. The third section highlights the unusual dynamic patterns estimated from the data and discusses the implications for policy. The fourth section discusses whether the impulse response estimates are driven by the Jordà method. The final section concludes.

Mechanics of the Highway Program and the Potential Importance of Credit Markets

The mechanics of the highway program described by Leduc and Wilson suggest to me that credit markets must be part of the transmission mechanism. The highway spending process begins when the federal government distributes *grants* to states. States have four years to incur *obligations* by writing contracts with vendors. As construction work is completed, the vendors send bills to the state, which then forwards them to the federal government. The *outlays* occur when the federal government pays those bills.

Note that no money is being disbursed by governments at the time that construction companies are hiring workers and buying materials. Companies are not paid until the federal government has sent the funds after receiving the bills. According to the second column of Leduc and Wilson's table 1, only 44 percent of the obligations are paid in the first two years. This result implies that contractors are hiring workers and paying suppliers before receiving any government funds. Thus, they must be financing their receivables with credit. Presumably, it is easier to get a loan based on a contract written with a state government that has received federal grants than with a private firm. Thus, one of the ways that government purchases may stimulate the economy is that it allows firms easier access to credit. A key question, though, is whether this credit demand is crowding out loans for other private companies that do not have contracts with the government. It would be interesting to see what is happening to credit flows in the state after these infrastructure grants.

Puzzling Dynamic Effects of Highway Spending

The estimated impulse responses have some very puzzling features that are not discussed by the authors. Consider the graph of GDP in figure 3, panel A, and the graphs of productivity, employment, personal income, and wages and salaries displayed in figure 4. All show the same basic dynamic pattern: the variables rise somewhat on impact, then fall below normal in years two through four, rise above normal in years six through eight, and then return to normal by year ten. Leduc and Wilson's theoretical model is able to produce a small dip in the early years because of time-to-build considerations, but the dip is miniscule relative to the positive effects much later in the theoretical model. In contrast, the employment response in the data shows that the trough is deeper than the expansion.

These declines in years two through four of numerous series are particularly puzzling when compared to the dynamic patterns of government spending. Figure 5 shows that grants, obligations, and outlays are all significantly above normal during these times. In fact, the outlay series peaks at year two. The analysis thus begs the question: Why are employment, wages and salaries, and GDP below normal when infrastructure spending is above normal? Who is working on these projects if employment is depressed?

Leduc and Wilson focus all of their attention on the positive impact effect on GDP and the positive effects on multiple variables at years six through eight. However, the anomalous drops in key variables at years two through four are just as large and statistically significant as the effects on which Leduc and Wilson focus. One could have easily used these results to write a paper entitled "Infrastructure Spending Depresses Short-Run Employment."

These results suggest that infrastructure spending is not a good stimulus tool. While the GDP multiplier peaks at eight years in the future, the average multiplier in the first three to four years appears to be about zero, since the positive impact effect shown in figure 3, panel A, is balanced by the negative effects in years two through four. Moreover, employment falls significantly below zero for most of the first five years after a shock. Since policymakers care more about the short-run employment response than the GDP response, these results suggest that infrastructure spending would actually be counterproductive. Leduc and Wilson cite evidence that labor costs represent only 8 percent of the expenditures on road construction because it is so capital intensive. Thus, infrastructure spending is clearly not an activity that can create many jobs in the short-run.

Econometric Issues

The authors use a nonstandard method for calculating the impulse responses. In this section, I will discuss how their method compares to two other methods that are more frequently used.

The authors use Jordà's (2005) method of direct projections. Jordà has argued that his method is more robust to misspecification because it does not depend on estimating the entire system. I worry, however, that the unusual oscillatory behavior of many of the variables may be a result of their use of this method. A comparison of panel A of figure 3 with panels D and E show that the Jordà method produces pronounced oscillations in the GDP response whereas the other two leading methods do not. In fact, the other two methods produce patterns that are closer to the ones predicted by the theoretical model. The other two methods are (1) standard VAR impulse response functions and (2) dynamic simulations, as used by Romer and Romer (1989, 2010) and Ramey and Shapiro (1998). Both of these solve for the impulse responses recursively based on one estimated equation per variable, so that the response at horizon $t + h$ shares estimated parameters with the responses at other horizons. In contrast, Jorda's method does not impose any link between responses at different horizons; the estimate response for each variable at each horizon comes from a different estimated equation.

To determine whether these three methods yield differences in other contexts, I compare the results of the three methods in studying the dynamic impact of a shock to military news on government spending and private spending, in a simplified version of the system I used in Ramey (2012). Unlike the authors' arguments for infrastructure spending, we have no reason to believe that rises in government spending due mostly to military spending should have echo effects years later. In all three methods, I trace shocks to my military news variable on log per capita values of real government spending and private spending. Data are quarterly from 1939 through 2010. I include 4 lags and quadratic trends.

The following is a brief summary of the implementation of each method in the present context. All systems include constant terms and trend terms, but they are suppressed for ease of exposition.

Standard VAR

$$z_t = B_1 z_{t-1} + B_2 z_{t-2} + B_3 z_{t-3} + B_4 z_{t-4} + e_t$$

Variable z is a vector of news, govt spending (g), private spending (y), and the first element of e_t is the shock to news. Impulse response functions for horizon $t + h$ are calculated by setting e_{1t} to a nonzero value and calculating the responses recursively from the estimated B's, allowing all three variables to respond dynamically. Confidence intervals are constructed using bootstrap standard errors.

Dynamic Simulations ("Romer Method")

$$x_t = \sum_{i=1}^{4} a_i \, news_{t-i} + \sum_{i=1}^{4} c_i x_{t-i} + u_t$$

Variable x is government spending or private spending. To calculate the impulse responses, one sets news$_t$ to a nonzero value and calculates the responses recursively for each x separately. Standard errors are computed using Monte Carlo methods.

Jordà Direct Projections

$$x_{t+h} = \delta^h news_t + \sum_{i=1}^{4} \gamma_i^h g_{t-i} \sum_{i=1}^{4} \lambda_i^h y_{t-i} + u_t$$

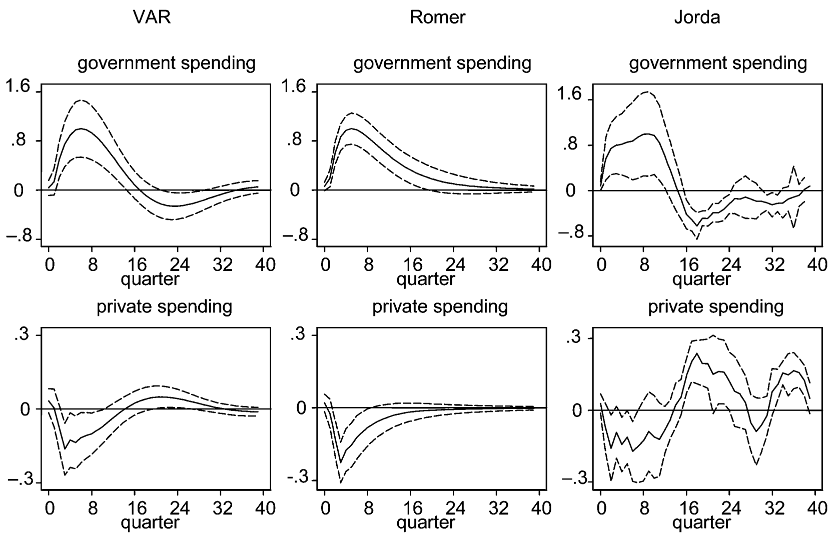

Fig. C1. A comparison of three methods for calculating impulse responses
Note: 90 percent confidence intervals shown.

where again x is g or y. To calculate the impulse response at each horizon h, one regresses x_{t+h} on news at t and lagged values of g and y. The estimated δ^{h}'s are the responses. Standard errors are heteroscedastic and autocorrelation consistent standard errors.

To facilitate comparison, I have set the value of the shocks so that the peak response of government spending is 1 percent in each case. Figure C1 shows the responses for government spending and private spending calculated in three different ways.

The results show that the standard VAR method and the Romer method give similar results. There is a small amount of oscillation in the standard method compared to the Romer method, but it is not pronounced. In contrast, the Jordà method predicts that while private spending falls in the short run, it rises significantly between quarters 16 and 24 and again between quarters 32 to 40. This pattern is qualitatively similar to the one identified for infrastructure spending by Leduc and Wilson. In the present case, which uses shocks to news about military spending as the shock, the economic intuition for the ripple effect after four years is not so obvious. These results make me wonder whether there is something about the Jordà method that produces oscillations in the responses rather than it being a real feature of the economy. More research is needed about the properties of the various methods for calculating impulse responses.

Conclusions

Overall, Leduc and Wilson's paper fills an important gap in the literature by studying the effects of shocks to infrastructure spending. The data constructed are particularly rich and the attention to details regarding exogeneity and lags is commendable. While translating their state-level multipliers to aggregate estimates is difficult, the results are still interesting for answering questions about policies at the state level.

I have raised two caveats about their overall conclusions, however. First, while their results suggest substantial multipliers six years after the initial shock, the short-run effects are often significantly negative. These results suggest that infrastructure spending is not a useful component of stimulus packages. Second, I have raised the issue that the Jordà method that they use to estimate impulse responses seems to induce oscillations in the responses that are not present when other methods are used. Thus, more research is needed to determine which methods are best for estimating impulse responses.

Endnote

For acknowledgments, sources of research support, and disclosure of the author's material financial relationships, if any, please see http://www.nber.org/chapters/c12752.ack.

References

Jordà, Oscar. 2005. "Estimation and Inference of Impulse Responses by Local Projections." *American Economic Review* 95 (1): 161–82.

Ramey, Valerie A. 2012. "Government Spending and Private Activity." Forthcoming in *Fiscal Policy after the Financial Crisis*, edited by Alberto Alesina and Francesco Giavazzi. Chicago: University of Chicago Press.

Ramey, V. A., and M. D. Shapiro. 1998. "Costly Capital Reallocation and the Effects of Government Spending." *Carnegie-Rochester Conference Series on Public Policy* 48 (1): 145–94.

Romer, Christina D., and David H. Romer. 1989. "Does Monetary Policy Matter? A New Test in the Spirit of Friedman and Schwartz." *NBER Macroeconomics Annual 1989*, edited by Olivier Jean Blanchard and Stanley Fischer, 121–84. Cambridge, MA: MIT Press.

———. 2010. "The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks." *American Economic Review* 100 (3): 763–801.