Macroeconomic Consequences of Infrastructure Investment

Valerie A. Ramey

University of California, San Diego and NBER

November 13, 2019

Abstract

In this paper, I examine macroeconomic theory and empirical evidence on the short-run and long-run effects of government investment in developed economies such as the U.S. I begin by presenting a stylized dynamic general equilibrium model in order to elucidate the economic intuition behind the effects. I use the model to explain the recent findings from the macroeconomic quantitative literature that short-run multipliers on government investment are likely to be *lower* than those on government consumption in most instances. I next analyze the leading empirical estimates of the long-run effects of public investment. Using insights and artificial data from the stylized model as a guide, I demonstrate the econometric biases that may be present in some estimates from the literature. I then review the empirical estimates on the short-run effects, with particular attention to the ARRA. I build on some of the existing literature to search for direct effects of highway infrastructure grants on construction employment. In common with several leading papers from the literature, I do not find positive effects. I conclude that most of the research suggests that while government investment is likely to increase output in the long run, its short-run effects are near zero in most situations.

Prepared for the November 15-16, 2019 NBER Conference *Economics of Infrastructure Investment*. I am grateful for helpful discussions with Hafedh Bouakez, Gabriel Chodorow-Reich, John Fernald, Per Krusell, Daniel Leff Yaffe, Johannes Wieland and Sarah Zubairy and helpful comments from seminar participants at Duke. Dan Wilson kindly provided supplemental data from his 2017 paper with Sylvain Leduc.

I. **Introduction**

Public capital can play an important role in increasing long-run output and standards of living. Because of nonrivalry in consumption and/or non-excludability in use, the private sector will tend to underprovide key types of productive capital. Hence, there is a role for government to raise social welfare by providing productive public capital, even when it must tax private resources to finance it. Economic history is replete with examples of public capital, and infrastructure in particular, that had significant impacts on long-run GDP and/or welfare. For example, Gordon (2016) highlights the contributions of publicly provided sanitation, clean water, and electrical infrastructure to both the rise in life expectancy and increase in productivity in the U.S. during the first part of the 20th Century. In the post-WWII period, the U.S. interstate highway program has been linked to significant increases in productivity and output (e.g. Aschauer (1989), Fernald (1999), Leff Yaffe (2019)).

More recently, government infrastructure spending has also figured prominently in policy discussions regarding short-run stimulus. Government infrastructure spending is viewed by many policymakers as having advantages over government consumption spending for stimulating the economy during a recession. In a traditional Keynesian model, both productive and wasteful government spending stimulate the economy in the short run through standard income and multiplier effects and help push output back to potential output. Government investment spending such as infrastructure spending, however, has the additional advantage that it can change the path of potential output. In particular, if a short-run increase in government spending also raises the stock of productive public capital or long-run total factor productivity (TFP), then government spending provides two benefits: Keynesian demand stimulus in the short run and neoclassical supply stimulus in the long run. These lasting effects are particularly welcome since typically stimulus packages must be financed with an increase in distortionary taxes after the recession is over. If output remains higher because of the long-run effects of more public capital, then the tax base expands and the necessary increases in tax rates are less.

In this paper, I examine the macroeconomic theory and empirical evidence on the benefits of infrastructure spending, both in the long run and the short run. Much of the theory and the empirical work suggests that even when there are substantial long-run benefits of infrastructure investment, the short-run benefits are probably lower than for non-productive government spending. In the last few years, the macroeconomic theory literature has discovered that realistic features of infrastructure investment, such as the importance of time to build and sector-specific demand effects, can work to reduce the short-run aggregate stimulus effects, even when the long-run supply-side benefits are present. Moreover, much of the existing macroeconomic empirical evidence is consistent with the predictions of these theories. I conclude that infrastructure investment may not be the most powerful short-run stimulus.

On the other hand, based on the theory and empirical estimates, I conclude that there is probably a significant long-run benefit to infrastructure spending. Based on some rough calculations using the benchmark neoclassical model and leading empirical estimates, I find that the optimal steady-state level of government investment spending may be above the current U.S. rate of 3.5 percent.

The paper proceeds as follows. Section II works through the effects of government investment and consumption in a benchmark neoclassical model. It develops the intuition for the mechanism at work and performs some experiments. It then extends the model in several important ways and shows how the implications change. It derives and compares multipliers on government consumption and government investment in both the short run and long run. Section

III considers the addition of New Keynesian features and studies how the predictions change. It reviews the quantitative New Keynesian model results from the literature and discusses how predictions can change when monetary policy is constrained by the zero lower bound. Section IV adds a brief note on the geography of trade models of benefits of infrastructure.

Section V then moves on to the empirical evidence on the long-run effects of public investment in the U.S. I show that it is important to distinguish exactly which elasticity of output to public capital is being measured, with an illustration using the theory from Section II. I then discuss the empirical challenges and how various studies attempt to overcome them. Section VI studies the shorter-run effects of government investment spending. Much of the focus is on the ARRA studies, and in particular on the infrastructure part of the ARRA. I offer new estimates of the effects of the ARRA on employment in highway construction. Section VII summarizes and discusses the results that emerge from the previous sections and concludes.

II. **Government Investment in a Neoclassical Model**

This section analyzes the short-run and long-run effects of government investment and public capital in a stylized neoclassical model. The New Keynesian model adds features, such as sticky prices, to an underlying neoclassical base, so neoclassical mechanisms continue to be key drivers of results even in New Keynesian models. Hence, it is useful to begin by highlighting the mechanisms by which government investment has its effects in the benchmark neoclassical model. In later empirical sections, I use insights and artificial data generated from this model to explain why some empirical methods estimate higher returns than others.

A. Neoclassical Model Structure and Transmission Mechanisms

Most of the macroeconomics analysis of government investment builds on the pioneering work of Baxter and King (1993), who were the first to analyze both the short-run and long-run effects of government investment in a fully dynamic general equilibrium neoclassical macroeconomic model.[1] In the typical neoclassical model, government purchases have direct impacts on the economy in several ways. Let G_t^C denote government consumption goods purchases in period t and let G_t^I denote government investment goods purchases. The sum of government purchases has a direct impact through the economywide resource constraint:

(1) $$C_t + I_t + G_t^C + G_t^I \leq Y_t.$$

C_t is private consumption, I_t is private investment, and Y_t is output. This resource constraint is key to the wealth effects that drive the labor and output response in both neoclassical and New Keynesian models. A government that purchases goods and services extracts resources from the economy. Financing through current or future lump sum taxes adds no additional effects, so the resource constraint captures the key impacts. If there is no direct effect of government spending on the production possibilities of the economy, a rise in government purchases leaves the private sector with fewer resources. Households respond by lowering their own consumption and leisure and raising their labor supply. Employment rises not because the demand for labor has risen (since government spending does not directly affect the aggregate marginal product of labor) but because

[1] Baxter and King build on the earlier work by Aschauer (1988) and Barro (1989), which introduced the neoclassical approach to analyzing the effects of government spending. Their analyses were conducted within simpler analytical models that necessarily constrained the dynamic interactions of capital and labor. As Baxter and King show, the use of quantitative models allows for the relaxation of those constraints and produces different predictions. Other strains of the literature have studied the growth consequences of public capital. See for example Glomm and Ravikumar (1994, 1997),

labor supply has risen. The rise in labor supply induced by the wealth effect is the key mechanism by which an increase in government purchases raises output in virtually all modern macroeconomic models.

While government consumption and government investment enter symmetrically in the resource constraint in equation (1), they play different roles in the rest of the economic structure. Most modelers assume that government consumption enters household utility, but in a separable way, so that it has no impact on the marginal utility of consumption. In this case, there is no additional impact of government consumption on the economy, other than raising household welfare. Allowing instead for government consumption to be a complement or substitute for private consumption in the utility function can lead to a wide variety of possible effects, which are not considered here. To be concrete, suppose that a representative household maximizes lifetime utility U:

$$(2) \qquad U = E_0 \sum_{t=0}^{\infty} \beta^t \left[\ln(C_t) + \varphi \ln(1 - N_t) + \Gamma(G_t^C)\right]$$

β is the discount factor. The middle term is the natural log of leisure, where the time endowment has been normalized to 1 and N_t is hours worked. Note that both C_t and N_t are normal goods.

Government investment, on the other hand, can have direct effects on the production function. Baxter and King (1993) specify the following stylized Cobb-Douglas aggregate production function:

$$(3) \qquad Y_t = A_t K_{t-1}^{\theta_k} N_t^{\theta_n} (K_{t-1}^G)^{\theta_G} .$$

A_t is the level of total factor productivity (TFP), K_t is the private capital stock at the end of period t, K_t^G is the public capital stock at the end of period t, and N_t is the quantity of labor. Typical analyses assume constant returns to private inputs, so that $\theta_k + \theta_n = 1$. The size of θ_G, the exponent on public capital, plays an important role in the long-run impact of government investment, which can have consequences for its short-run impact. If θ_G is greater than zero, then in this calibration there are increasing returns to scale.

Note that virtually all of the short-run effect of government spending on output must operate through labor input for the following reason. Both private and public capital are relatively fixed in the short run, so if government spending does not affect TFP (A_t) in the short run, government spending can raise GDP in the short run only to the extent that it raises labor input.

Finally, government investment and public capital are linked since government investment this period adds to the public capital stock available at the beginning of next period:

(4) $$K_t^G = G_t^I + (1 - \delta)K_{t-1}^G ,$$

δ is the depreciation rate on public capital. Since government investment is typically a small fraction of the steady state stock of public capital, it takes numerous periods of elevated government investment to raise the public capital stock a noticeable amount. The capital accumulation equation for private capital is similar:

(5) $$K_t = I_t + (1 - \delta)K_t - 1$$

Equations (3) and (4) capture the distinguishing characteristics of government investment relative to government consumption. A dollar increase in government investment raises the stock of public capital through equation (4), which has multiple effects on the production function in equation (3). First, for fixed TFP, private capital, and labor, the higher public capital stock leads to higher output. Second, because the higher public capital stock raises the marginal products of both private capital and labor, it incentivizes firms to invest in more capital and to hire more workers. In the neoclassical model, the only type of government spending that raises the demand for labor is government spending that directly raises TFP or public capital.

How the government spending is financed has first-order effects on the response of output and labor. The simplest case, which I will use for my benchmark case, is that the government uses lump sum taxes. The government budget constraint is given by:

$$(6) \qquad G_t^C + G_t^I = T_t$$

where T_t is lump sum taxes. In the representative household, perfect financial markets, and rational expectations case, the timing of the lump sum taxes has no effect: deficit spending with later increases in lump sum taxes is equivalent to balanced budget lump sum taxes. In this case, the social planner solution is equivalent to the decentralized competitive equilibrium. In the more realistic case that the government must raise distortionary taxes, the timing of those taxes matter and the positive effects of government spending on output can be severely muted.

In this benchmark economy, the social planner chooses sequences $\{C_t\}$, $\{N_t\}$, $\{I_t\}$, $\{Y_t\}$, and $\{K_t\}$ to maximize the lifetime utility of the representative household given in equation (2), subject to the economywide resource constraint in equation (1), the production function in (3), the capital accumulation equations in (4) and (5), as well as exogenous processes for the two types of government spending. Of course, it would make perfect sense to allow the social planner to choose the optimal level of public capital as well. However, since we want to do experiments on the effects of more government investment, we take the government spending as exogenous for now. In the empirical section, I will discuss the implications of optimal choices of public capital.

The first order conditions and steady-state conditions for this model are presented in the appendix.

B. Quantitative Predictions from the Neoclassical Model

Even the simple model presented above cannot be solved analytically unless the depreciation rate on capital is set at 100 percent, so I analyze the model quantitatively. The calibration of the parameters is for a quarterly model and is similar to the calibration/estimation from Leeper et al. (2010). In equation (2), the discount factor β is 0.99, which implies an annual real interest rate of 4 percent and φ is set to 4.5 in order to produce a steady state in which the representative household spends 20 percent of its time endowment on work in the baseline model. In equation (3), the capital share $\theta_k = 0.64$ and the labor share $\theta_n = 0.36$. I will consider two values for θ_G of 0.05 and 0.1. 0.05 was the baseline used by Baxter and King (1993). A meta analysis by Bom and Ligthart (2014) finds a mean estimate of 0.08 in the short run and 0.12 in the long run. The quarterly depreciation rate on both types of capital, δ, in equation (5) is set at 0.025.

The experiments involve shocks to either government consumption or government investment. I assume that each follows a first-order autoregressive (AR(1)) process:

(7) $\quad G_t^J = constant + \rho\, G_{t-1}^J + \epsilon_t^J \quad$ for J = C, I

The constant terms are chosen to yield steady-state fractions of government spending relative to GDP that match their values for 2019 in the U.S., which are approximately 14 percent for government consumption and 3.5 percent for government investment. This calibration sets the government investment-to-GDP ratio below the optimal value of 4.4 percent that a fully-maximizing social planner would choose. Similar to Leeper et al. (2010), I assume an AR(1) process for government spending with a serial correlation parameter 0.9, which involves a fairly persistent increase.

1. Baseline Experiments

I consider three baseline experiments. The first is an unanticipated increase in government consumption G_t^C, the second is an unanticipated increase in government investment G_t^I when the exponent on public capital in the production function in (3) is θ_G = 0.05, and the third is an unanticipated increase in government investment when is θ_G = 0.10.

Figure 1 shows the impulse responses for three experiments for the baseline model. These show the endogenous response of key variables to an unanticipated increase in government consumption or government investment that is autocorrelated. All are shown in percentage terms. Government spending, output, consumption, private investment, and public capital are expressed in deviations from their own steady state values as a percent of steady-state output. Labor input and wages are percent deviations from their own steady state values. The real interest rate is in percentage point deviations from its own steady state.

Consider first an increase in government consumption, whose effects are depicted by the black solid line. As discussed above, the direct effect is a negative wealth effect on consumption and leisure. The government is extracting resources from the economy, so consumption falls and labor supply rises. Because of diminishing marginal product of labor, real wages fall. This rise in the labor supply boosts output; there is no demand channel. Real interest rates rise and as a result, investment falls. There is no change to public capital. All values eventually return to their original steady-state levels since the government spending increase is not permanent.

The effect of an increase in government investment when the exponent on public capital θ_G = 0.05 is shown by the blue short dashed line in Figure 1. In this case, the impact effect on labor, consumption, and output is less than for a government consumption increase. A muted negative wealth effect is key to this difference: the government is still extracting the same amount from current output, but it is using it to contribute to future wealth in the form of productive capital. Recall that we are assuming the economy starts from a steady state in which public capital is below the optimal level.

However, private investment falls more during the first six quarters than in the government consumption case. The weaker wealth effect on labor means that output rises less in the short-run, so more private spending must be crowded out by the government spending. The weaker wealth effect means that households do not reduce their consumption as much so the brunt of the crowd-out falls on private investment. The differential short-run response of consumption and investment is a key theme in Boehm's (forthcoming) analysis of the short-run multipliers on government consumption versus government investment. Building on insights from Barsky, House, and Kimball (2007) and others, Boehm notes that the long service life of private capital leads to a very high intertemporal elasticity of substitution in investment demand. Because investment rates are typically small relative to the capital stock, agents are very willing to intertemporally substitute investment, much more so than for consumption. As I will discuss below, the additional features of Boehm's model magnify these effects.

The real interest rate rises about the same amount on impact, but then continues to rise. As the public capital stock is built up, output continues to grow. Labor input remains high and private investment becomes elevated since the higher level of public capital raises the marginal products of both labor and private capital. Wages also rise above their initial steady state.

The green long dashed line in Figure 1 shows the effect of the government investment change for even more productive public capital, with capital $\theta_G = 0.10$. All of the mechanisms discussed in the last case are even stronger in this case, so output and labor rise little in the short run and private investment falls even more. However, as the public capital stock is built up, output rises significantly for a prolonged period of time. The effects are even more pronounced for higher values of θ_G.

The most important insight offered by this experiment is that the short-run effects of government spending on output and labor are *lower* for government investment than for government consumption. The positive wealth effects of more public capital in the future have a dampening effect on the stimulus effects of government spending in the short run.

2. Experiments with Time to Build

Leeper, Walker, Yang (2010) highlight two important limitations to the stimulus effects of government investment: implementation delays and future fiscal financing adjustments. They estimate a more elaborate neoclassical model and consider the effects of these two realistic additions. Implementation delays are very realistic for infrastructure spending. As Leeper et al. (2010) point out, typically there are delays in appropriations and the subsequent outlays occur slowly over time. While routine maintenance of roads may involve delays of a year between appropriations and completion, new highways, roads and bridges can involve delays of four years. Leeper et al. modeled both the slow outlay process as well as a time-to-build feature.

The American Recovery and Reinvestment Act (ARRA) illustrates how difficult it is to fast track infrastructure project investment. The ARRA stimulus package specifically targeted "shovel-ready" projects because of the urgency for immediate government spending. Even then, there were significant delays between the appropriations, the obligations, the outlays and the actual use of the new infrastructure.

Figure 2 shows the cumulative spending as a percent of Federal Highway Administration Appropriations in the ARRA. Although the ARRA was passed in February 2009, by the end of 2009 only 11 percent had been spent. A year later, just over half had been spent. The cumulative percent spent did not approach 100 percent until the end of 2012.

I now illustrate Leeper et al.'s (2010) insight about implementation delays in the context of my simplified model. I add only time to build, since my baseline experiment already builds in the persistent spending path. I assume that there is an 8-quarter delay between the initial government investment and the addition to the useable public capital. To be specific, I replace equation (4) with:

(4′) $$K_t^G = G_{t-8}^I + (1-\delta)K_{t-1}^G ,$$

Everything else is the same.

Figure 3 shows the results of these experiments. The black line repeats the results for the baseline case for government consumption, which is not affected by time to build. The blue short dashed line and the green long dashed line show the results for government investment with time to build for the two values of θ_G. Time to build effects further mute the short-run stimulus effects of government investment. The negative wealth effects continue to be muted, so labor and output rise less and consumption falls less. Private investment continues to fall more. However, the positive effect of rising public capital in the baseline experiments is delayed eight quarters. This delay results in lower stimulus to output for almost three years relative to the case of government consumption increases. Eventually the strong positive effects on output dominate, but this would typically be long after a recession is over. As Leeper et al. explain, implementation delays can lead to similar effects to those for announced but slowly phased in tax cuts: because everyone knows that the (after-tax) returns to labor and private investment will be higher in the future than now, there is an incentive to delay productive activity.

3. Neoclassical Multipliers

I now consider the output multipliers associated with each of these experiments. It should be noted that government spending multipliers are typically low, around 0.4, in neoclassical models when the changes in government spending are temporary. Only permanent changes in government spending can lead to short-run multipliers that are unity in the typical neoclassical model. New Keynesian features can raise multipliers, but most would raise the government consumption and investment multipliers similarly, so the relative ordering remains similar, as I will show in the next section. Thus, it is useful to compare the multipliers across the experiments without necessarily accepting the actual level of the multiplier.

The multipliers are calculated as recommended by Mountford and Uhlig (2009), as the present discounted value of the integral of the output response up to quarter h divided by the present discounted value of the integral of the government spending response up to quarter h. The interest rate used for discounting is the equilibrium real interest rate generated by the simulated model.

Figure 4 shows the multipliers for each horizon for the first 20 quarters. With no delays due to time to build, the government investment multipliers are lower than the government consumption multipliers for the first six quarters, but then exceed them as time goes on. With 8-quarter time-to-build delays in government infrastructure investment, the output multiplier for government investment is less than the multipliers for the government consumption for the first five years. Thus, evaluated only by the short-run multiplier, government infrastructure investment is inferior to government consumption investment in its potential to stimulate the economy.

Table 1 shows the long-run multipliers for each of the cases. Here is where government investment spending has its great advantages. While the present value long-run multiplier for government consumption is a measly 0.3, the present value long-run multiplier for government investment is ranges from 1.5 to 1.7 when θ_G = 0.05 and 2.6 to 3.1 when θ_G = 0.1. The range depends on whether there are time-to-build delays. The higher real interest rate in the short-run has noticeable effects, as illustrated in the final column which shows undiscounted integral multipliers. In those cases, the government investment multiplier is higher but there is little difference between the no delay experiment and the time-to-build experiment. Thus, the message from Table 1 is that government investment is unambiguously superior to government consumption in generating long-run increases in output, as long as public capital is productive.

The actual levels of multipliers, however, can depend on the details of the model and the experiment. Table 2, Panel A, shows the multipliers from Baxter and King (1993) and Leeper et al. (2010). Baxter and King's government investment experiments consider only permanent increases in the ratio of government investment to GDP. The long-run multiplier depends crucially on the assumed value of the elasticity of output to public capital, θ_G. Their long-run multipliers range from 1.2 for government consumption (i.e. θ_G=0) to 13 for θ_G= 0.4. In contrast, Leeper et al. (2010) report long-run multipliers that are smaller for both values of θ_G because they also include the response in distortionary taxes that they estimate from the data. Nevertheless, the result that the long-run multiplier for government investment is greater than for government consumption is robust to these details.

III. Government Spending in New Keynesian Models

A. Overview of New Keynesian Mechanisms

New Keynesian (NK) models typically use the basic structure of the neoclassical model, but add elements intended to capture traditional Keynesian intuition. The benchmark NK model relies on mechanisms that are not closely related to the traditional Keynesian intuition, though.

Consider the effects of government consumption spending in a benchmark NK model, which features monopolistic competition in product markets and sticky prices. In this model, there is a steady-state markup of prices over marginal cost. The stickiness of prices makes the markup countercyclical in response to monetary and government shocks. When those shocks raise output, real marginal cost rises because of the diminishing returns to labor. Sticky prices, however, prevent prices from rising in the short run, which reduces the markup distortion. As Broer et al. (2019a) have recently pointed out, the countercyclical profits associated with the countercyclical markups lead to additional negative wealth effects on household, increasing labor supply more than the neoclassical wealth effect alone. They show that this is an important mechanism for the transmission of monetary policy. In answer to my recent query about the importance of this mechanism for government spending multipliers, Broer et al. (2019b) demonstrate that the negative wealth effect of countercyclical profits is the *entire* reason that multipliers in the NK model are greater than those in the neoclassical model during times of normal monetary accommodation. Woodford (2011) shows that these NK features can raise the government spending multiplier above the neoclassical model multiplier, but the multiplier only reaches unity if monetary policy can hold real interest rates steady.

An exception to the limit of one on the multiplier is the case of the zero lower bound (ZLB). When interest rates are at their zero lower bound, the monetary authority wants to reduce nominal interest rates more but cannot. Thus, the monetary authority cannot lower real interest rates. The only way that real interest rates can fall is if a fiscal stimulus can generate higher expected inflation. Carefully timed fiscal stimulus that lasts during the zero lower bound period but not after can generate higher expected future inflation. These expectations lower the *ex ante* real interest rate and spur economic activity now. It is this mechanism, identified by Eggertsson (2009), Woodford (2011), and others, that can lead to high government spending multipliers at the ZLB.

There are several reasons to be skeptical of some of the NK predictions at the ZLB, though. First, Wieland (2018) highlights the result that previous theoretical work finding large multipliers at the ZLB relied on multipliers changing discontinuously for small changes in parameters. Wieland discovers that this discontinuity is due to their changing the equilibrium selection mechanism. Once a stable equilibrium selection mechanism is used, multipliers vary continuously with the parameters and are almost always equal to unity.

Second, the results depend crucially on two links: the increase in government spending generates higher expected inflation and higher expected inflation raises private spending. There is mixed evidence on whether government spending increases during ZLB periods actually generate the required increase in inflationary expectations. Dupor and Li (2015) study the response of inflation to fiscal expansions in the post-WWII U.S. and particularly during the Great Recession. They study times when monetary policy is accommodative and find that the inflation response is either nonexistent or far too small to generate the large multipliers. Miyamoto, Nguyen, and Sergeyev (2018) find some evidence of higher inflationary expectations during the Japanese ZLB period. Bachman, Berg, and Sims (2015) test the second link by studying the impact of individual consumer inflation expectations on their spending propensities in the Michigan Survey of Consumers. They find that higher inflationary expectations have no impact on the readiness to spend during normal times and in fact have a *negative* effect on the readiness to spend during zero lower bound periods.

A third reason to be skeptical of the theoretical results for the NK model at the ZLB are the predictions regarding the effects of negative supply shocks. As first highlighted by Eggertsson (2011), a negative supply shock, which in normal times would result in a fall in output, is predicted to *stimulate* output during a ZLB period. The mechanism is the same as the one that generates higher spending multipliers during the ZLB: higher expected inflation, which lowers the real interest rate. In this case, a negative supply shock leads to higher expected inflation, which lowers the ex ante real interest rate and spurs demand. Wieland (2019) tests this prediction by studying the impacts of the earthquake and tsunami Japan as well as the effect of oil price shocks. The NK model predicts that these shocks should have been expansionary since Japan has been at the ZLB for decades. He finds that they were contractionary, contrary to the prediction of NK theory.

The expansionary effects of negative supply shocks at the ZLB are not just a side show with respect to implications for optimal fiscal policy. If one believes the NK mechanism that predicts higher multipliers on government spending at the ZLB, then one must also accept the prediction that *raising distortionary income taxes* at the ZLB is expansionary. Eggertsson (2011), Woodford (2011), and Drautzburg and Uhlig (2015) demonstrate this prediction in both simple calibrated NK models and estimated medium scale NK models. Thus, anyone recommending greater government spending at the ZLB because of higher multipliers should also recommend that the spending be financed with increases in distortionary taxes rather than deficits.

That said, while there are reasons to be skeptical of the NK mechanisms that lead to higher spending multipliers at the ZLB, there is some empirical evidence that indeed multipliers can be higher at the ZLB. For example, in Ramey and Zubairy (2018), we estimate multipliers around 1.4 at the ZLB in historical data if we exclude periods of WWII rationing. Miyamoto, Nguyen and Sergeyev (2018) apply Ramey and Zubairy's methods to Japan and find higher multipliers at the ZLB, around 1.5 on impact. Further, as discussed later Boehm (forthcoming) finds higher multipliers for government investment spending at the ZLB. Thus, whatever the mechanism, multipliers may be higher at the ZLB.

As just highlighted, the mechanisms in the benchmark NK model are not closely related to the intuition of traditional Keynesian models. In an effort to bring New Keynesian models closer to old Keynesian intuition, researchers have introduced additional elements. For example, Galí, Lopez-Salido, Vallés (2007) explore extensions of the benchmark model designed to recapture traditional Keynesian intuition about the effects of government spending. They do not consider

ZLB effects. They first demonstrate that a benchmark NK model makes the same prediction about the response of private consumption as the neoclassical model: an increase in government consumption spending leads consumption to decline because of the negative wealth effect. The NK model shares this feature because households are assumed to be rational and forward looking and labor markets are assumed to be competitive. Thus, the same negative wealth effect that generates higher labor supply and thus output necessarily generates lower consumption. Galí et al. add two additional features to the benchmark NK model to try to reverse the negative effect on consumption: they assume a fraction of consumers are rule-of-thumb (also known as "hand to mouth") and a noncompetitive labor market in which all wages are set by unions and households are off their labor supply curves. They find that if labor markets are competitive, the fraction of consumers required to be rule of thumb is implausibly high. However, the combination of noncompetitive labor markets and a fraction of rule of thumb consumers above 0.25 can lead to rises in private consumption and multipliers above unity, at least on impact.

To summarize, in the benchmark NK model, output rises in response to government spending entirely because of negative wealth effects operating through two channels. The first channel is the neoclassical channel whereby government use of resources leads households to work harder. The second channel is the countercyclicality of markups leading to countercyclical profits, which create additional negative wealth effects after a rise in government spending. When the economy is not constrained by the zero lower bound on nominal interest rates, the benchmark NK model can produce multipliers somewhat above the neoclassical model but typically not above unity. The joint addition of rule of thumb consumers and noncompetitive labor markets can overcome the negative wealth effect on consumption. Multipliers can be significantly higher at the zero lower bound. I have offered several reasons to be skeptical of those mechanisms. I have also highlighted the fact that those mechanisms would also suggest that policy makers should raise income taxes during recessions. I will now review the NK literature that has specifically investigated the effects of government investment.

B. New Keynesian Analyses of Government Investment

One of the first explorations of government investment specifically in a NK model is by Linnemann and Schabert (2006). They were also seeking mechanisms that could overturn the negative response of consumption to government spending increases. They provided analytical results from a model without private capital. They found that if the government spending contributed to aggregate production, and the elasticity of output to public capital was sufficiently high, then positive wealth effects of the supply-side effects of government spending outweighed the negative wealth effects. In general, high elasticities of labor supply and monetary policy that responded to the positive supply shock effect by lowering nominal interest rates contributed to this result. The paper analyzes the effects of various features, such as tax policy and monetary policy, on generating this effect.

Many of the subsequent NK analyses of the relative stimulus effects of government investment spending were conducted in response to the financial crisis and the stimulus programs adopted in response. Some of these are summarized in Panel B of Table 2. Coenen plus 17 co-authors (2012) analyze the effects of various fiscal policies in the leading large scale New Keynesian dynamic stochastic general equilibrium (NK DSGE) models used by the Federal Reserve, the European Central Bank, the IMF and other leading policy institutions. These are very rich models that incorporate a host of additional NK elements, such as rule-of-thumb consumers and noncompetitive labor markets. They report the average first year multipliers for a 2-year stimulus, financed with deficits. As is typical in NK models, the results depend crucially on the responses to monetary policy. The multipliers on both government consumption and investment are 0.9 if monetary policy follows the usual Taylor rule rather than being accommodative. The multiplier rises as high as 1.6 in both cases if monetary policy is accommodative. Coenen, Straub, and Trabandt (2013) conduct an analysis in the ECB model with a richer fiscal sector and the range for their multipliers in the short run and long run are similar to those of Coenen et al. (2012). These are shown in the second row of Panel B. Note that the short-run government consumption multiplier tends to lie above the government investment multiplier, just as we saw in the neoclassical model.

The results by Albertini, Poirier, and Roulleau-Pasdeloup (2014), shown in the third row of Panel B of Table 2, illustrate the importance of the accommodative monetary policy assumption. Their impact multipliers are below one for both government consumption and investment during normal times but one or above at the ZLB. The Drautzburg-Uhlig (2015) results, shown in the fourth row of Table 2B show how including a realistic delayed tax response significantly lowers the multipliers for both government consumption and investment.

Boehm (forthcoming) highlights a potentially important limitation of the short-run stimulus effects of government investment spending. As I discussed briefly in my analysis using the neoclassical model, Boehm notes that the long service life of private capital leads to a very high intertemporal elasticity of substitution in investment demand. Because investment rates are typically small relative to the capital stock, agents are very willing to intertemporally substitute investment, much more so than for consumption. These effects are magnified in Boehm's NK model which has two sectors, a consumption goods sector and an investment goods sector, and where labor is not mobile in the short run between these two sectors. He considers temporary increases in government consumption or investment spending, financed by lump-sum taxes. Because of the sectoral immobility of labor, government consumption competes with the private sector for consumption goods whereas government investment competes with the private sector for investment goods. Consumers are less willing to intertemporally substitute their purchases of consumption goods, so there is less crowding out of private consumption by government consumption. In contrast, because investment is small relative to the capital stock, firms are much more willing to intertemporally substitute their investment spending. As a result, a temporary increase in government investment spending has a large crowd out effect on private investment. As Table 2B shows, his model implies that short-run multipliers are lower for government investment than for government consumption. Both are below unity in his model in the short run. In the long run, the beneficial production effects of public capital lead to a multiplier of 1.6.

Bouakez, Guillard, and Roulleau-Pasdeloup (2017, forthcoming) demonstrate a further reversal of both neoclassical and NK results during normal times when ZLB mechanisms are in force. Recall that Leeper et al. (2010) had found that introducing time-to-build delays in public capital lowered the short-run multiplier on government investment spending in a neoclassical model. Bouakez et al. (2017) show that Leeper et al.'s qualitative results continue to hold in a NK model when the economy is not constrained by a ZLB and when monetary policy behaves normally. However, when the economy is thrown into a liquidity trap by certain types of shocks, longer time-to-build delays lead to *higher* short-run multipliers. As explained above, the amplification of government spending multipliers and reversal of results about supply shocks at the ZLB all come about through expected inflation effects. Time-to-build delays prevent increases in the public capital stock from occurring in the ZLB period, which helps counter any deflationary pressures. As the final row of Table 2B shows, their impact multipliers for both government consumption and investment are below unity during normal times but are 2.3 in ZLB periods when there is no extra time-to-build delay and reach four for government investment when there is a 4-year time-to-build delay.

Bouakez et al. (2017) assume that government spending is financed with lump-sum taxes in all of their experiments. However, we know from the work of Woodford (2011) and Eggertsson (2011), that at the ZLB even larger multipliers can be generated by using income taxation rather than deficit financing or lump sum taxation. Thus, if one accepts the mechanisms that lead to Bouakez et al.'s (2017) flipping of the effects of time-to-build delays, one must also believe that higher income tax rates during the ZLB raise multipliers even higher. This uncomfortable policy implication is probably not understood by many who believe that spending multipliers are higher at the ZLB.

IV. **A Brief Note on Trade Models of Infrastructure Spending**

This section offers a brief summary of the important work in the trade literature that has much to say about the returns to transportation infrastructure. These models focus on the longer-term benefits of transportation infrastructure.

The geography of trade literature takes transportation costs and spatial features seriously in modeling the potential benefits of transportation infrastructure. Much of the technical work of this literature builds on pioneering work of Eaton and Kortum (2002). The quantitative analyses in these models directly model and measure the extent to which transportation infrastructure reduces trade costs between two points, opens access to markets, and allows for a variety of spillovers, agglomeration effects, and congestion effects. This literature, which is also known as

"Quantitative Spatial Economics," has been surveyed recently by Redding and Turner (2014) and Redding and Rossi-Hansberg (2017). Recent contributions include those by Donaldson and Hornbeck (2016), who revisit Fogel's (1962, 1964) classic analyses of the contributions of railroads to U.S. economic growth; Donaldson (2018), who studies the impact of railroads in India during the Raj, and Allen and Arkolakis (2019), who develop a new geographic framework and use it to study the welfare effects of improving each segment of the U.S. highway system. The results of the Allen and Arkolakis (2019) paper are particularly pertinent to current policy debates. Though they find heterogeneity in the welfare effects across segments, their quantitative analysis indicates that for all highway links the welfare benefits of additional lane-miles substantially exceed the construction costs.

V. Empirical Evidence on the Long-Run Effects of Government Investment in Public Capital and Infrastructure

This section reviews and analyzes some of the leading estimates of the long-run effects of government investment. I discuss some key methodology used and estimates obtained that reveal the challenges of estimating causal effects of public capital. I illustrate the econometric problems by estimating the effects of public capital on artificial data generated by a simple extension of the model in Section II. Finally, I discuss a promising way to address the challenges and the estimates that emerge.

A. Brief summary of estimates

There is a long literature that seeks to measure the returns to infrastructure investment. An early example is Fogel's (1964) pioneering analysis of the contributions of railroads to U.S. economic development. Several decades later, Aschauer's (1989, 1990) famous hypothesis that the productivity slowdown in industrialized countries was caused by reductions in infrastructure investment led to renewed research in this area. He estimated an aggregate production function and found an elasticity of output to public capital of 0.39 in U.S. data. Munnell's (1990) extension of his work found similar results, with elasticities between 0.31 and 0.39.

Much of the recent macroeconomics literature has focused on short-run effects of general government spending, but several papers also provide estimates for long-run multipliers on government investment spending. For example, Iltzetzki, Mendoza, Vegh (2013) use structural vector autoregressions on a panel of OECD countries to study the effects of government spending in a wide range of circumstances. They use Choleski decompositions to identify shocks. When they focus on government investment they find multipliers for public investment that ranged between 0.4 in the short-run to 1.6 in the long run. Boehm (forthcoming) specifically compares multipliers for government investment and consumption spending in a panel of OECD countries. He also uses Cholesky decompositions, but does control for forecasts. He also finds a long-run multiplier of 1.6 for government investment spending.

Some of the most convincing evidence of the productivity of public capital has used regional or industry variation in the U.S. to estimate the output effects of road construction in the U.S. It is important to note that these estimates give only *relative* effects because aggregate effects are typically taken out by constant terms or time-fixed effects. Fernald (1999) exploits the differences in benefits of the U.S. interstate highway system across industries. He specifically models transportation services as an input into the production function, taking into account the complementarity between vehicles owned by the industries and roads and the difference uses across industries. He finds that industries that rely more heavily on transportation experienced greater increases in productivity than other industries as a result of the building of the U.S. interstate highway system. Using additional identifying assumptions, he translates his relative estimates into a production function elasticity of output to roads of 0.35, an estimate even higher than Aschauer (1989). However, he argues that the effects are not large enough to be the principal explanation of the productivity slowdown.

Leff Yaffe (2019) uses state panel data and narrative evidence to estimate the output effects of the building of the U.S. interstate highway system, accounting for anticipation effects and crowding-in of state and local spending on roads. His multiplier estimates are significantly affected by the estimated "crowd-in" of state highway spending. In particular, an infusion of funds to a state (instrumented using Bartik-style instruments) typically led to additional road building to connect to the interstate highway system. When he includes the additional state and local spending in the government spending measure, Leff Yaffe's long-run relative multiplier estimate is 1.8.

Leduc and Wilson (2013) estimate the effects of Federal highway grants to states during more recent times using annual state-level data starting in the 1990s. They report various long-run (i.e. 10 year) multipliers. Their favored ones are just under 2.

Bom and Ligthart's (2014) excellent literature review discusses the variety of estimates for the role of public capital. Their meta analysis settles on a mean production function elasticity of output to public capital of 0.08 in the short run and 0.12 in the long-run. They find that the elasticity is higher for public capital installed by local or regional governments and for core infrastructure. The mean estimate of the output elasticity for these latter types of public capital is 0.19 in the long-run.

The estimates are less optimistic for emerging economies. Perhaps because of less efficient governments, many of the estimated returns are surprisingly low. Henry and Gardner (2019) survey the evidence in numerous countries and conclude that in only a minority do infrastructure projects, such as paved roads and electricity, clear the required hurdles.

In the next two sections, I highlight two major challenges associated with estimating the production function elasticity of output. The first is associated with the difference between the production function elasticity and the steady-state general equilibrium elasticity. The second is the problem of the endogeneity of public capital spending. I illustrate the challenges by comparing the approaches used in three leading sets of papers: (1) Aschauer (1989) and Munnell's (1990) static production function estimates; (2) Flores de Frutos and Pereira (1999) and Pereira's (2000) structural vector autoregression (SVAR) estimates; and (3) Bouakez, Guilliard, and Roulleau-Pasdeloup's (2017) TFP and cointegrating relation estimates.

B. Production Function vs. General Equilibrium Output Elasticities

Aschauer (1999) and Munnell (1999) estimated their production elasticities using log levels of contemporaneous variables. Essentially they regressed the logarithm of aggregate output on the logarithms of labor, private capital, and public capital. Thus, temporarily leaving aside the endogeneity issues that I will discuss below, they were both estimating the production function elasticity, θ_G from the production function from Section II, repeated here for reference:

(3) $$Y_t = A_t K_{t-1}^{\theta_k} N_t^{\theta_n} \left(K_{t-1}^G\right)^{\theta_G} .$$

Let us now compare their method and results to the analysis by Pereira and Flores de Frutos (1999), denoted "PF" in the following exposition. PF noted several possible problems with the estimation method of Aschauer and Munnell, including issues of possible spurious regression (e.g. because the macroeconomics variables are nonstationary, omission of dynamic feedbacks, and possible simultaneous equation bias. They addressed all three of these issues by using a structural vector autoregression (SVAR) to estimate the elasticity of output to public capital. First, they tested and found unit roots in the logs of output, labor, and the two capital stocks. They could find no evidence of cointegration, so they estimated their system in first differences to avoid spurious regression. Second, their use of the SVAR allowed complete dynamics. Third, they allowed for reverse causality from output and the other variables to public capital and identified exogenous movements in public capital as the innovation to public capital not explained by lagged values of the other endogenous variables, i.e., they used a Cholesky decomposition to identify the exogenous shock.

PF fully recognized that they were estimating a different elasticity from the one estimated by Aschauer and Munnell. PF's headline number is a long-run elasticity of private output to public capital of 0.63. To obtain this number, they first estimate the impulse responses of all the endogenous variables, including public capital, to their identified exogenous shock to public capital. They then calculate the long-run elasticity (shown in their Table 6) as the ratio of the impulse response of log output at 5 to 10 years to the impulse response of log public capital at 5 to 10 years, since both impulse responses have stabilized at their new levels by that time.[2]

[2] Those impulse responses are shown in their Figure 1.

This elasticity of output to public capital estimated by PF is not, however, the production function elasticity θ_G. The production function elasticity of output to public capital, θ_G, is the elasticity of output to an increase in public capital, *holding TFP, labor, capital constant.* There is another elasticity of output to public capital, however, that includes the general equilibrium-induced changes in private inputs. The increase in public capital raises the marginal products of private inputs, which leads to incentives to accumulate more private capital. This is in fact what PF estimate. PF's impulse response function estimates show that private capital also rises permanently. (Employment bounces around in the short run, but then returns to a level slightly above its former value.) Because private capital is allowed to respond, PF's elasticity is not the production function elasticity.

The dynamic general equilibrium neoclassical model presented in Section II allows us to map the relationship between the production function elasticity and the general equilibrium steady-state elasticity for our particular calibration.[3] I use the model to simulate how the elasticity of steady-state output to public capital, which allows for general equilibrium effects on private inputs, is related to the production function elasticity, θ_G. I use the same calibration as Section II, setting the ratio of government investment to GDP equal to 0.035 to match the value for 2019. I then calculate elasticities based on increasing the public capital stock by one unit.

Figure 5 shows the results. The blue line is the simulated relationship and the dotted line is the 45º line. The relationship between the two elasticities is affine, and is given by:

General Equilibrium Steady-State Output Elasticity = $0.047 + 1.49 \cdot \theta_G$

[3] Pereira and Flores de Frutos (1999) instead conduct the comparison by manipulating their estimates to find the steady state implied by their time series model.

The positive constant term means that even when public capital is not directly productive, output increases by 0.05 percent when public capital increases by one percent in steady state. This effect stems from the negative wealth effect on labor supply: if the government raises the level of unproductive public capital, it must do so by siphoning resources from the private sector. Households respond by lowering their consumption and raising their labor. The rise in labor also induces a rise in private capital. Thus, the steady-state elasticity of output to steady-state public capital is always greater than the elasticity of output to public capital in the production function. Part of this difference is due to the negative wealth effect raising labor supply and part is due to the induced investment in private capital, which grows as θ_G rises.

We can use this relationship to calculate what PF's estimated elasticity would imply for the value of θ_G.[4] Their long-run elasticity allowing private inputs to change of 0.63 is the general equilibrium steady-state elasticity. Using the equation above, this implies that an estimate of θ_G of 0.39, exactly equal to Aschauer's estimate!

C. The Econometric Problem of Endogenous Public Capital

As Flores de Frutos and Pereira (1999) recognize, the long-run elasticity they estimate also includes dynamic feedback into the government's public capital decision. Their headline estimates are based on the assumption that the government chooses public capital in part based on developments in the economy, but only with a lag. Their estimated regressions show significant effects of those lags. Thus, part of the overall response they estimate is due to the feedback effect of a growing economy on the endogenous part of public capital.

[4] PF in fact report the elasticity of *private* output. Since I am not sure how they define private output, I abstract from this issue and just consider total output.

The endogeneity of public capital is a potentially serious problem, recognized by many of researchers. Aschauer (1989) used OLS for his main estimates, but attempted to deal with possible reverse causality by using lagged endogenous variables as instruments. Using lagged endogenous variables as instruments was a common practice in the late 1980s, but is now known to require implausible exclusion restrictions in most macroeconomic applications. PF recognized the problem, but I could not find an estimate of the extent of the bias implied by their SVAR estimates.

The simultaneity problem occurs because larger and more wealthy economies invest in more public capital. In fact, since a benevolent social planner should choose a level of public capital that maximizes the discounted utility of the representative household, it should respond to technological progress by increasing the amount of public capital.

We can make this point concrete by exploring the effects of endogeneity on estimates by simulating artificial data from an extension of the simple model presented in Section II and determining whether standard methods can estimate the correct value of θ_G. This exercise is what I have called a "DSGE Monte Carlo" (Ramey (2016)).

To be specific, I augment the calibrated neoclassical model to allow the social planner to *choose* the optimal level of public capital, based on maximizing the discounted utility of the representative household.[5] I use the baseline calibration with θ_G = 0.05. I then allow true TFP, *A* in the production function above, to vary. Because I am interested in long-run effects, I calculate how steady-state values of the key variables change with changes in TFP.

I estimate a regression similar to the one used by Bouakez et al. (2017). In particular, rather than regressing output itself on the inputs, they use Fernald's (2014) measure of TFP as the

[5] Note that the social planner problem is not concave, since I assume constant returns in the private inputs, so existence and uniqueness are not guaranteed. See Glomm and Ravikumar (1994, 1997) for a thorough analysis of model in which the government chooses the public capital optimally. My explorations with the simple model suggest that there exists a unique maximum of the social planner problem, as long as θ_G is not too large.

dependent variable. Fernald makes very general assumptions and carefully measures TFP at the industry level using factor shares and then aggregates them to get aggregate TFP. He also adjusts it for cyclical utilization. In the context of the simple aggregate production function in my model, Fernald's measure is defined as follows:

(8) $$Fernald\ dTFP \equiv dln(Y_t) - \theta_K \cdot dln(K_t) - \theta_N \cdot dln(N_t)$$

The growth rate of TFP is defined as the growth rate of output less the share-weighted growth rates of private capital and labor. Fernald also assumes constant returns to scale in the private inputs so he sets $\theta_K + \theta_N = 1$ and uses NIPA tables to calculate the shares. This definition and the production function from equation (3) above implies the following relationship between Fernald's measure of TFP and public capital:

(9) $$Fernald\ dTFP = dln(A_t) + \theta_G \cdot dln(K_t^G)$$

Thus, Fernald's (2014) dTFP measure consists of both true technological change, *dln(A),* and the effects of public capital.

Suppose we regress Fernald's dTFP measure on the growth rate of public capital. Since true technological change is not observed, it shows up in the error term of the regression, i.e., the ε_t in

(10) $$Fernald\ dTFP = \theta_G \cdot dln(K_t^G) + \varepsilon_t$$

Bouakez et al. (2017) cumulate these measures and estimate the regression as a cointegrating equation. I will describe more details of their procedure below.

In the artificial data I generate from my model, I calculate a measure of TFP as the log of output minus the share-weighted logs of private capital and labor, just as Fernald does. I set the weights equal to the actual shares from the model. I then regress the log level of TFP measure on the log of public capital, or the growth rates on each other, on the artificial data generated by the model. Recall that I am focusing only on steady-state equilibrium values.

Whether I estimate in log levels or in growth rates, I obtain an estimate of θ_G equal to 0.64. Hence, the estimate is severely biased upward relative to the true value of 0.05. The reason for the upward bias is intuitive. When there is an increase in technology, A, the marginal product of all inputs increases. As a result, private agents increase private capital and the social planner increases public capital. Thus, the error term ε_t in equation (9) is correlated with public capital.

One could in principle solve the problem by using instrumental variables, but it is difficult to find instruments for public capital in aggregate data. That is why most modern analyses still use simple Cholesky decompositions. Bouakez, Guillard, and Roulleau-Pasdeloup (2017), however, use a method that turns out to reduce the bias significantly. Although they do not discuss endogeneity issues, their method goes far to deal with this type of bias.

In a short discussion section at the end of their mostly-quantitative New Keynesian model effects at the zero lower bound paper, Bouakez et al. (2017) review the literature on the productivity of public capital and then present some independent evidence. Their motivation is as follows. They use Fernald's (2014) carefully constructed TFP measure to avoid estimating a complete production function. They then add "it is still important to account for the additional factors that may affect TFP in the long run" (Bouakez et al. (2017), p. 75), but do not explain *why* it is important. The DSGE Monte Carlo analysis I developed above provides the perfect motivation: any changes in measured TFP (apart from public capital) are likely to lead the government to change public capital endogenously. Thus, in order to reduce the bias in the regression in equation (10), one should control for as many sources of TFP as possible in order to remove them from the error term, ε. Bouakez et al. (2017) construct measures of the stock of research and development spending and the stock of human capital. Their finding of cointegration between the log level of Fernald's TFP, log public capital, log R&D stock and log human capital is strong evidence that they have identified the key drivers of TFP. Pereira and Flores de Frutos (1999) estimated their model in first-differences because they could not find cointegration. Bouakez et al.'s (2017) analysis shows that more key variables needed to be included. By estimating the cointegration equation, Bouakez et al. (2017) are picking up the long-run, presumably steady-state, relationships because the estimates are driven by the stochastic trends.[6]

Bouakez et al.'s main estimates, shown in their Table 2, imply a production function elasticity of output to public capital of 0.065. When I use their data and omit the other determinants of TFP (i.e. the R&D stock and human capital stock) I estimate a coefficient on the log of public capital of 0.33. The difference between these two estimates is explained by the type of bias I demonstrated in my DSGE Monte Carlo. Thus, their controls for other factors affecting TFP go far to reduce the bias.

VI. Empirical Evidence on the Short-Run Effects of Government Investment in Public Capital

[6] See King, Plosser, Stock and Watson (1987, 1991) for a discussion of the role of stochastic trends in long-run growth. The 1987 NBER working paper version is much more complete than the 1991 AER version.

During the Great Recession, government infrastructure spending received much attention because of its possible role in stimulating the economy. The American Recovery and Reinvestment Act (ARRA), enacted in early 2009 in the depths of the Great Recession, used both transfers and government purchases to try to stimulate the economy. Infrastructure spending was an important component of the purchases. The stimulus package specifically targeted "shovel-ready" projects because of the urgency for immediate government spending. As I showed earlier in Figure 2, the delays in spending were nevertheless substantial.

As I discussed in Section II, the theoretical evidence suggests that, dollar for dollar spent, government investment spending has lower short-run stimulus effects than government consumption. What does the empirical evidence say?

A. Aggregate Evidence

Pereira and Flores de Frutos (1999), reviewed in detail in the context of long-run estimates, also studied the short-run effects. They found negative short-run effects of infrastructure spending on employment in all of their specifications. This fact, coupled with the recognition of the delays in investment, led them to recommend against using public investment for short-run stimulus. They argued that it could actually be counterproductive.

As discussed earlier, Iltzetzki, Mendoza, Vegh (2013) used structural vector autoregressions on a panel of OECD countries to study the effects of government spending in a wide range of circumstances. When they focused on government investment they found multipliers for public investment around 0.4 in the short-run.

The work of Boehm (forthcoming), which I discussed earlier for its quantitative model predictions, also tested those predictions using a panel of OECD countries. Recall that his key economic insight was that government investment should have a lower short-run multiplier than government consumption because the elasticity of intertemporal substitution for investment is much higher than for consumption. This feature means that government investment spending will crowd out much more private investment spending than government consumption spending will crowd out private consumption. He tests this prediction of his model using a panel of OECD countries from 2003 to 2016. He identifies exogenous shocks to government consumption and investment using a Choleski identification, controlling for forecasts to avoid anticipation effects. He estimates of multipliers near zero for government investment and around 0.8 for government consumption. He also finds evidence supporting the mechanisms he highlights in his theory. In particular, he finds that a government consumption shock does not crowd out private consumption, but a government investment shocks significantly crowds out private investment. Consistent with this evidence, he also finds little change in the real interest rate in the consumption goods sector after a consumption shock, but a significant increase in the real interest rate in the investment goods sector.

He also offers some final evidence that provides some support to the models predicting higher multipliers at the zero lower bound. When he estimates his model separately over zero lower bound periods and normal periods, he finds evidence of a multiplier around 1 for government consumption and around 1.2 for government investment during zero lower bound periods. Recall that Bouakez et al. (2017, forthcoming) showed that at the ZLB, the NK model predicted a flipping of the ranking of multipliers, with government investment multipliers higher at the ZLB. Boehm's point estimates qualitatively support this prediction. The standard errors of the estimates are higher, though, so the estimates are not statistically different from each other.

B. Cross-State Evidence

Many of the recent studies have estimated the effects of infrastructure by exploiting variation across states. This is especially true of the studies of the effects of the ARRA. These studies can estimate only *relative* effects because they exploit subnational data; that is, they answer the question "how much more employment or output occurs in State A receives $1 more in spending than State B?" Thus, the estimates do not provide direct evidence on aggregate effects because, by construction, they net out financing effects and they do not measure the net effects of positive spillovers versus business-stealing effects. Moreover, most do not account for induced state and local spending, so the multiplier estimate may undercount the total government spending required to produce the result. Nevertheless, they provide valuable insight into the mechanisms.

As an aside, the state employment data is typically much better than gross state product data. As a result, most studies focus on employment effects rather than gross state product effects. This focus is reasonable for short-run studies that are interested in the stimulus effects of government investment.

Leduc and Wilson (2013) estimate the effects of Federal highway grants to states during using annual state-level panel data from 1993 to 2010. Their long-run multipliers were discussed in a previous section. As noted by Ramey (2013, 2018), however, their short-run estimated effects do not suggest much stimulus effect. Consider one of the graphs from Figure 4 of their paper, reproduced here from Ramey (2018):

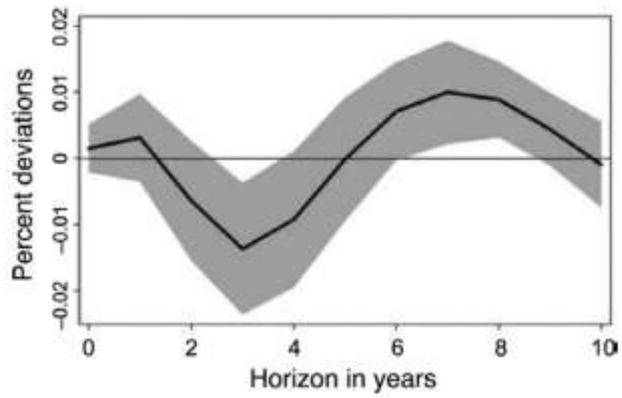


Figure 9.3

The Effects of State Highway Spending on State Employment. Shaded area is 90 percent confidence interval.
Source: Reproduced from one graph of figure 4 of Leduc and Wilson (2012).

This graph shows the effects of state highway spending on state total employment. The impulse response shows little effect on impact or at year 1, but then a significantly negative effect on state employment at years 2 through 5. Thus, these results suggest that highway spending is *counterproductive* as a stimulus. These results echo those found by Flores de Frutos and Pereira (1999) in aggregate data.

Studies that focused all or in part on the infrastructure elements of the ARRA include Wilson (2012), Chodorow-Reich et al. (2012), Leduc and Wilson (2017), Dupor (2017), and Garin (forthcoming). Chodorow-Reich (2019) synthesizes and standardizes the various studies of the ARRA for all types of spending and finds very similar employment multiplier estimates once they are standardized to calculate multipliers the same way. He finds that all of the leading instruments, whether they be Medicaid formulae, Department of Transportation factors, or a mixture of many factors, produce similar results. In particular, he estimates that two job-years were created for each $100,000 spent. As I point out in Ramey (2019), however, these estimates are based on unweighted data and do not take into account crowd-in of state and local spending. Once I make those adjustments, I find that each $100,000 spent led to 0.8 job-years created. These estimates are based on weak instruments, though, since the literature's instruments that are so strong for the ARRA grants are unfortunately weak for spending including additional state and local spending.

Leduc and Wilson (2017) used cross-state variation in ARRA appropriations for highways to study flypaper effects, i.e., whether federal grants for highway construction crowd in or crowd out state and local spending on highways and roads. They found significant crowd in, with each dollar in federal aid resulting in a total of $2.30 in state highway spending. The focus of their paper was the response of state and local spending and how that interacted with rent seeking, but in the appendix they showed regressions of the change in employment in the highway, street and bridge construction industry on the instrumented appropriations. They were able to find a significant positive results in only one case of several. The failure to find positive results echoes my point that the earlier Leduc and Wilson (2013) analysis of highway spending before the ARRA did not find positive effects on total employment.

As Garin (2019) argues, a positive effect of highway spending on construction employment is a necessary condition for any further effects, such as local spillovers and Keynesian multipliers. Therefore, I examine in more detail the impacts of the ARRA highway grants on employment in highway, street and bridge construction, which I will call "highway construction" for short. I use Leduc and Wilson's (2017) data and a similar specification, which they describe in the text associated with Table B1. In particular, the regressions, which use cross-state variation for identification, estimate the effect of ARRA highway apportionments per capita in 2009 on the variables of interest in the succeeding years. I use the baseline sample of 48 states of Leduc and Wilson, and instrument for apportionments with their two road factors. I include their political variables as controls, though I lag them in my local projection specification so that all right-hand side variables are dated 2009 or earlier. I include the change in per capital employment in highway construction between 2007 and 2008 as an additional control for pretrends. I estimate the impulse response in each year using a series of local projection regressions in which the left-hand side variable is the change in the variable of interest from 2008 and year h, where h ranges from 2009 to 2013.

Figure 6, Panel A shows the impulse responses for the specification just described. The upper left graph accurately estimates that all of the ARRA obligations occurred in 2009. The upper right graph shows that the outlays occurred mostly in 2009 and 2010. The lower left graph of Panel A supports the main result of Leduc and Wilson (2017), which is that total highway spending rose by more than the outlays. My new result is the impulse response for highway construction employment, shown in the lower right graph of Panel A. According to the estimated impulse response function, highway employment barely responds in 2009 and 2010, but then falls significantly after that. These effects are clearly contrary to the intended effects of the ARRA.

Dupor (2017) in "So, Why Didn't the 2009 Recovery Act Improve the Nation's Highways and Bridges" argues that the ARRA did not improve the highways and bridges because the federal grants completely crowded out state and local spending. Thus, Dupor argues for the opposite result of Leduc and Wilson (2017), who find significant crowding in. Dupor notes that the difference might be due to his controlling for the logarithm of state population, which Leduc and Wilson do not include. He does not make clear the econometric motivation for adding this control.

To determine how the results change when log population is included as a control, I re-estimate impulse responses by including Dupor's log population control in the model I used to estimate the impulse responses shown in Panel A. The results when the population control is included are shown in Panel B. The top two graphs are similar to those from the previous specification, but the bottom left graph showing the impact on total highway spending is very different. In contrast to the analogous graph in Panel A, there is no change in total highway spending in Panel B. This result suggests complete crowd out. The highway construction employment effects, however, is similar, with virtually no change in 2009 and 2010 but a significant negative effect in 2011 through 2013.

The results obtained adding Dupor's control variable no longer imply that increases in highway spending lower highway construction employment, but they imply that no change in highway spending lowers highway construction employment. One might suspect a problem with the instruments. Chodorow-Reich (2019) tested the overidentifying assumptions using those instruments along with other leading ones from the literature and could not reject the overidentifying assumptions. Thus, this explanation seems less likely.

Neither of the implied stories by Leduc and Wilson (2017) or Dupor (2017) is encouraging for highway grants as a stimulus. In the Leduc and Wilson results, total highway spending rises significantly as a result of the federal grants, but it results in a decrease in employment in highway construction. In the Dupor results, federal grants are ineffective in raising total highway spending, and still highway construction employment falls.

Garin (2019) finds slightly more positive results. He uses a database on almost 3,000 counties and ARRA spending on highways to estimate the direct effects on overall construction (not just highways) employment, as well as total employment. The biggest effect he finds is in total construction employment in 2010, with six jobs created per $1 million. He finds that each dollar of stimulus spent in a county led construction payrolls to increase by 30 cents over the next five years, an increase that is consistent with the labor share in the construction industry. However, when he tests for general equilibrium effects on local employment and payroll, he estimates effects that are close to zero. He finds no evidence of a local multiplier effect.

In sum, there is scant empirical evidence that infrastructure investment, or public investment in general, has a short-run stimulus effect. There are more papers that find negative effects on employment than positive effects on employment. The ARRA results are particularly negative, since the ARRA spending occurred at a time when interest rates were at the zero lower bound and the unemployment rate was 9 to 10 percent. Despite the slack in the economy and the accommodative monetary policy, the effects on construction employment were either small positive or negative.

VII. Summary and Implications

This section begins with a summary of the key findings and then discusses their implications.

A. Summary

- Even when government investment has significant long-run effects, the short-run stimulus multipliers are less than those from government consumption in most situations. The two key reasons are (i) the effects of time-to-build delays and (ii) the propensity of government investment to crowd out private spending more than government consumption does. These results are supported by quantitative models, empirical panel studies across OECD countries, and time series analysis in the U.S.

- There is both theoretical support and some empirical support for the short-run multiplier on government investment being higher when interest rates are constrained by the zero lower bound (ZLB). The theoretical mechanisms that lead to this effect, however, also imply that at the zero lower financing government spending with distortionary income taxation leads to higher multipliers than financing it with deficit spending, a result contrary to most economists' priors.

- Cross-section and panel evidence on U.S. states or counties that focuses on bridge, highway, and road infrastructure spending suggests that the spending leads to either no change or a decline in employment in the first several years, even during ZLB periods. There is no clear explanation for these puzzling results.

- My review and small extension of the empirical literature on the long-run estimates suggests that the aggregate production function elasticity of output to public capital is probably around 0.065, but could be as high as the 0.12 found by Bom and Ligthart's (2014) meta analysis. Some studies find higher estimates for core infrastructure, while others do not.

B. Implications

Let us consider the implications of the long-run estimates first. Suppose that the elasticity of output to public capital in the production function is the Bouakez et al. (2017) estimate of 0.065. What does this imply for the optimal steady-state ratio of government investment to private investment? Returning to the extension of the neoclassical model that allows the social planner to choose the optimal steady-state public capital, it is easy to derive the implied social capital. In particular, the expressions for the optimal steady-state ratios of government to private investment and capital are given by:

(11a) $$\frac{K^G}{K} = \frac{\theta_G}{\theta_k} \cdot \frac{\beta^{-1} - 1 + \delta_P}{\beta^{-1} - 1 + \delta_G}$$

(11b) $$\frac{G^I}{I} = \frac{\delta_G}{\delta_P} \cdot \frac{K^G}{K}$$

Recall that K^G is the stock of public capital, K is the stock of private capital, θ_G is the exponent on public capital in the aggregate production function, θ_k is the exponent on private capital, β is the discount rate of the representative household, δ is the depreciation rate of public capital, G^I is government investment, and I is government investment. Equations (11a) and (11b) generalize the earlier model to allow for different rates of depreciation on government versus private capital.

First suppose that the annual rate of depreciation on both types of capital is 0.1, so that the last factor on the right-hand side of equation (11a) becomes unity. Then if θ_k= 0.36, which is a standard calibration, a value of θ_G= 0.065 implies that the optimal ratio of government-to-private investment is 0.18. Equation (11b) implies the optimal ratio of government investment to private investment is also 0.18 if the depreciation rates are the same.

How does this compare to the U.S. data? According to BEA data for 2018, the ratio of gross government investment to gross private domestic investment was 0.19, very close to the optimal rate just calculated.[7] However, the ratio of current cost net stocks of government to private capital was 0.32 in 2018.[8] In steady state, these ratios can be different only if the depreciation rates are different.

Thus, let us use equations (11a) and (11b) allowing for different depreciation rates. The annualized discount factor from the stylized model is β = 0.96. If we add realistic population and

[7] These calculations are based on BEA GDP and income table 1.1.5 and 3.9.5. Some papers have argued that only nondefense government capital enters the production function. If we exclude defense gross investment, the ratio is 0.15, which is a little lower than the optimal rate.

[8] This calculation is based on BEA Fixed Asset Table 1.1. The ratio is equal to 0.28 if defense capital is excluded.

technology growth rates to the model, however, the transformed discount factor would be very close to unity. In this case, the right-hand side of equation (11a) is approximately equal to $\frac{\theta_G}{\theta_k} \cdot \frac{\delta_P}{\delta_G}$. Since the depreciation factors cancel when equation (11a) is substituted into equation (11b), there is no effect on the ratio of public to private investment. Thus, even with different depreciation rates, the ratio of government investment to private investment is still equal to $\frac{\theta_G}{\theta_k}$. Thus, the ratio of investment in the data is approximately equal to the optimal rate calculated from the model with θ_G= 0.065.

The different ratio of government to private capital is explained by the differences in depreciation rates. Recall that the 2018 ratio of public to private investment is 0.19 and the ratio of public to private capital is 0.32. Therefore, according to equation (11b), the ratio of government capital depreciation to private capital depreciation is 0.59. If the annual depreciation rate on private capital is the typical calibrated value of δ_P = 0.1, then δ_G = 0.06, which is within the range of standard estimates.

This stylized comparison suggests that the actual public investment rate is close to the optimal rate. However, the simple model assumes no distortionary taxation. The need to finance government spending with distortionary taxes would reduce the implied optimal government investment rate because a unit of government capital would cost more than a unit of output because of the depressing effect of distortionary taxes on output.

The results on the short-run stimulus effects are less positive. Even with the positive long-run effects on output, and the accompanying benefits of having a larger tax base to help pay off the debt from a stimulus package, it would be difficult to recommend infrastructure spending as a short-run stimulus, at least during normal times. Most empirical studies, including ones that focus on the ARRA during the zero lower bound, find either no effects or counterproductive effects on employment.

Does this mean that stimulus packages should not include productive government spending? Not necessarily. The essence of the arguments in favor of government infrastructure spending as a stimulus might apply more readily to other types of government spending, some of which are classified as government consumption rather than investment in the national accounts. A prime example is federal transfers to states to prevent teacher layoffs. During recessions, states facing balanced budget constraints often lay off public workers when tax revenues fall during a recession. Stimulus packages that prevent teacher layoffs in K-12 and instructor layoffs from junior colleges during a recession are likely to have positive short-run effects because, unlike infrastructure spending, there are no time-to-build delays and the employment effects are immediate. Government expenditures on teacher salaries are classified as government consumption expenditures, not investment expenditures, because they do not contribute to an increase in government-owned fixed assets. However, estimates of the returns to education would suggest significant positive effects on long-run human capital. Thus, it is possible that this type of government spending could have both a positive short-run stimulus effect and a (very long term) positive long-run supply-side stimulus because of the effects on the nation's human capital. More research is required, however, to determine whether these theoretical considerations play out in practice.

References

Albertini, Julien, Arthur Poirier, and Jordan Roulleau-Pasdeloup. "The composition of government spending and the multiplier at the zero lower bound." *Economics Letters* 122.1 (2014): 31-35.

Aschauer, David Alan. "The equilibrium approach to fiscal policy." *Journal of Money, Credit and Banking* 20.1 (1988): 41-62.

Aschauer, David Alan. 1989. "Is Public Expenditure Productive?" *Journal of Monetary Economics* 23(2): 177-200.

Aschauer, David Alan. "Does public capital crowd out private capital?." *Journal of monetary economics* 24.2 (1989): 171-188.

Bachmann, Rüdiger, Tim O. Berg, and Eric R. Sims. "Inflation expectations and readiness to spend: Cross-sectional evidence." *American Economic Journal: Economic Policy* 7.1 (2015): 1-35.

Barro, Robert J. 1981. "Output Effects of Government Purchases." *The Journal of Political Economy* 89 (6): 1086-1121.

Barro, Robert. "The neoclassical approach to fiscal policy." *Modern business cycle theory* (1989).

Barsky, Robert B., Christopher L. House, and Miles S. Kimball. "Sticky-price models and durable goods." *American Economic Review* 97.3 (2007): 984-998.

Baxter, Marianne, and Robert G. King. 1993. "Fiscal Policy in General Equilibrium." *American Economic Review* 83(3): 315–334.

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): 1329-1368.

Boehm, Christoph E. "Government consumption and investment: Does the composition of purchases affect the multiplier?." *Journal of Monetary Economics* (2019).

Bom, Pedro RD, and Jenny E. Ligthart. "What have we learned from three decades of research on the productivity of public capital?." *Journal of economic surveys* 28.5 (2014): 889-916.

Bouakez, Hafedh, Michel Guillard, and Jordan Roulleau-Pasdeloup. "Public investment, time to build, and the zero lower bound." *Review of Economic Dynamics* 23 (2017): 60-79.

Bouakez, Hafedh, Michel Guillard, and Jordan Roulleau-Pasdeloup. "The optimal composition of public spending in a deep recession." *Journal of Monetary Economics* (2019).

Broer, Tobias, Niels-Jacob Harbo Hansen, Per Krusell, and Erik Oberg. “The new Keynesian transmission mechanism: A heterogenous-agent perspective.” *Review of Economic Studies*. Forthcoming.

Broer, Tobias, Niels-Jacob Harbo Hansen, Per Krusell, and Erik Oberg. “A Note on the Transmission Mechanism of Government Spending Shocks in the Basic Neoclassical Model. October 2019 manuscript.

Chodorow-Reich, Gabriel, et al. "Does state fiscal relief during recessions increase employment? Evidence from the American Recovery and Reinvestment Act." *American Economic Journal: Economic Policy* 4.3 (2012): 118-45.

Chodorow-Reich, Gabriel. "Geographic Cross-Sectional Fiscal Spending Multipliers: What Have We Learned?." *American Economic Journal: Economic Policy* 11.2 (2019): 1-34..

Christiano, Lawrence, Martin Eichenbaum, and Sergio Rebelo. 2011. “When is the Government Spending Multiplier Large?” *Journal of Political Economy* 119 (1): 78-121.

Coenen, Günter, Christopher J. Erceg, Charles Freedman, Davide Furceri, Michael Kumhof, René Lalonde, Douglas Laxton et al. 2012. "Effects of Fiscal Stimulus in Structural Models." *American Economic Journal: Macroeconomics* 4 (1): 22-68.

Coenen, Günter, Roland Straub, and Mathias Trabandt. "Fiscal policy and the great recession in the euro area." *American Economic Review* 102.3 (2012): 71-76.

Coenen, Günter, Roland Straub, and Mathias Trabandt. "Gauging the effects of fiscal stimulus packages in the euro area." *Journal of Economic Dynamics and Control* 37.2 (2013): 367-386.

Cogan, John F., Tobias Cwik, John B. Taylor, and Volker Wieland. 2010. "New Keynesian versus Old Keynesian Government Spending Multipliers." *Journal of Economic Dynamics and Control* 34 (3): 281-295.

Drautzburg, Thorsten, and Harald Uhlig. "Fiscal stimulus and distortionary taxation." *Review of Economic Dynamics* 18.4 (2015): 894-920.

Dupor, Bill. 2017. “So, Why Didn’t the 2009 Recovery Act Improve the Nation’s Highways and Bridges? Federal Reserve Bank of St. Louis Review, Second Quarter 2017.

Dupor, Bill and Rodrigo Guerrero. 2017. “Local and Aggregate Fiscal Multipliers.” *Journal of Monetary Economics* 92 (December): 16-30.

Dupor, Bill, and Rong Li. "The expected inflation channel of government spending in the postwar US." *European Economic Review* 74 (2015): 36-56.

Dupor, Bill and M. Saif Mehkari. 2016. "The 2009 Recovery Act: Stimulus at the Extensive and Intensive Labor Margins." *European Economic Review* 85: 208-228.

Eggertsson, Gauti B. "What fiscal policy is effective at zero interest rates?." *NBER Macroeconomics Annual* 25.1 (2011): 59-112.

Farhi, Emmanuel, and Iván Werning. 2016. "Fiscal Multipliers: Liquidity Traps and Currency Unions." In *Handbook of Macroeconomics*, eds. John B. Taylor and Harald Uhlig. Vol. 2, 2417-2492. Amsterdam: Elsevier.

Fernald, John G. "Roads to prosperity? Assessing the link between public capital and productivity." *American economic review* 89.3 (1999): 619-638.

Fernald, John. "A quarterly, utilization-adjusted series on total factor productivity." Federal Reserve Bank of San Francisco, 2014.

Fogel, Robert William. *Railroads and American economic growth*. Baltimore: Johns Hopkins Press, 1964.

Garin, Andrew. "Putting America to work, where? Evidence on the effectiveness of infrastructure construction as a locally targeted employment policy." *Journal of Urban Economics* 111 (2019): 108-131.

Glomm, Gerhard, and Balasubrahmanian Ravikumar. "Public investment in infrastructure in a simple growth model." *Journal of Economic Dynamics and Control* 18.6 (1994): 1173-1187.

Glomm, Gerhard, and Balasubramanian Ravikumar. "Productive government expenditures and long-run growth." *Journal of Economic Dynamics and Control* 21.1 (1997): 183-204.

Gordon, Robert J. *The rise and fall of American growth: The US standard of living since the civil war*. Vol. 70. Princeton University Press, 2017.

House, Christopher L., and Matthew D. Shapiro. 2006. "Phased in Tax Cuts and Economic Activity." *American Economic Review* 96 (5): 1835-1849.

Ilzetzki, Ethan, Enrique G. Mendoza, and Carlos A. Végh. 2013. "How Big (Small?) are Fiscal Multipliers?" *Journal of Monetary Economics* 60 (2): 239-254.

Jordà, Òscar. 2005. “Estimation and Inference of Impulse Responses by Local Projections.” *American Economic Review* 95 (1): 161-82.

Leduc, Sylvain and Daniel Wilson. 2013. "Roads to Prosperity or Bridges to Nowhere? Theory and Evidence on the Impact of Public Infrastructure Investment.” *NBER Macroeconomics Annual* 27 (1): 89-142.

Leduc, Sylvain, and Daniel Wilson. "Are State Governments Roadblocks to Federal Stimulus? Evidence on the Flypaper Effect of Highway Grants in the 2009 Recovery Act." *American Economic Journal: Economic Policy* 9 (2): 253-92.

Leeper, Eric M., Todd B. Walker, and Shu-Chun Susan Yang. "Government Investment and Fiscal Stimulus." *Journal of Monetary Economics* 57 (2010): 1000-1012.

Leeper, Eric M., Todd B. Walker, and Shu-Chun Susan Yang. "Fiscal Foresight and Information Flows." *Econometrica* 81 (3): 1115-1145.

Leff Yaffe, Daniel. 2019. "The Interstate Multiplier." UCSD manuscript.

Miyamoto, Wataru, Nguyen, Thuy Lan, and Dmitriy Sergeyev. 2018. "Government Spending Multipliers Under the Zero Lower Bound." *American Economic Journal: Macroeconomics* 10 (3): 247-277.

Mountford, Andrew, and Harald Uhlig. 2009. "What are the Effects of Fiscal Policy Shocks?." *Journal of Applied Econometrics* 24 (6): 960-992.

Munnell, Alicia H. "Why has productivity growth declined? Productivity and public investment." *New England economic review* Jan (1990): 3-22.

Nakamura, Emi and Jón Steinsson. 2014. "Fiscal Stimulus in a Monetary Union: Evidence from U.S. Regions." *American Economic Review* 104 (3): 753-792.

Pereira, Alfredo M., and Rafael Flores De Frutos. 1999. "Public Capital Accumulation and Private Sector Performance." *Journal of Urban Economics* 46 (2): 300-322.

Pereira, Alfredo M. "Is all public capital created equal?." *Review of Economics and Statistics* 82.3 (2000): 513-518.

Ramey, Valerie A. 2011a. "Identifying Government Spending Shocks: It's All in the Timing." *The Quarterly Journal of Economics* 126 (1): 1-50.

Ramey, Valerie A. 2011B. "Can Government Purchases Stimulate the Economy?" *Journal of Economic Literature* 49 (3): 673-85.

Ramey, Valerie A. 2016. "Macroeconomic Shocks and their Propagation." In *Handbook of Macroeconomics*, eds. John B. Taylor and Harald Uhlig. Vol. 2, 71-162. Amsterdam: Elsevier.

Ramey, Valerie A., and Sarah Zubairy. 2018. "Government spending multipliers in good times and in bad: evidence from US historical data." *Journal of Political Economy* 126 (2): 850-901.

Swanson, Eric T., and John C. Williams. 2014. "Measuring the Effect of the Zero Lower Bound on Medium-and Longer-Term Interest Rates." *American Economic Review* 104 (10): 3154-85.

Wieland, Johannes F. "Are negative supply shocks expansionary at the zero lower bound?." *Journal of Political Economy* 127.3 (2019a): 973-1007.

Wieland, Johannes F. 2019b. "Zero Lower Bound Government Spending Multipliers and Equilibrium Selection." UCSD Manuscript.

Woodford, Michael. "Simple analytics of the government expenditure multiplier." *American Economic Journal: Macroeconomics* 3.1 (2011): 1-35.

Appendix

The following provides the first-order conditions and steady state conditions for the neoclassical model presented in Section II.

The social planner chooses sequences $\{C_t\}$, $\{N_t\}$, $\{I_t\}$, $\{Y_t\}$, and $\{K_t\}$ to maximize the lifetime utility of the representative household given in equation (2), subject to the economywide resource constraint in equation (1), the production function in (3), the capital accumulation equations in (4) and (5), as well as exogenous processes for the two types of government spending. The first-order conditions for the perfect foresight solution are:

(A-1) $\frac{\theta_n Y_t}{C_t} = \frac{\varphi N_t}{1-N_t}$ Marginal Rate of Substitution Condition

(A-2) $\frac{C_{t+1}}{C_t} = \beta\left[\theta_k \frac{Y_{t+1}}{K_t} + 1 - \delta\right]$ Consumption Euler Equation

If the social planner chooses government capital optimally, then we also have the first-order condition for that choice:

$$\frac{C_{t+1}}{C_t} = \beta\left[\theta_G \frac{Y_{t+1}}{K_t^G} + 1 - \delta\right]$$

The steady-state conditions are:

(A-3) $\frac{\theta_n Y}{C} = \frac{\varphi N}{1-N}$

(A-4) $\frac{K}{Y} = \frac{\theta_k}{\beta^{-1} - 1 + \delta}$

(A-5) $I = \delta \cdot K$

(A-6) $G^I = \delta \cdot K^G$

(A-7) $C + I + G^C + G^I = Y$

(A-8) $Y = A\, K^{\theta_k} N^{\theta_n} (K^G)^{\theta_G}$

If the social planner chooses public capital optimally, then in steady state,

$$\frac{K^G}{K} = \frac{\theta_G}{\theta_k}$$

Table 1. Long-Run Multipliers from a Quantitative Neoclassical Model

Experiment	**Present discounted value multiplier**	**Undiscounted integral multiplier**
Government consumption	0.3	0.3
$\boldsymbol{\theta_G}$**=0.05**		
Government investment No delays	1.7	2.4
Government investment 8-quarter time-to-build delay	1.5	2.2
$\boldsymbol{\theta_G}$**=0.10**		
Government investment No delays	3.1	4.4
Government investment 8-quarter time-to-build delay	2.6	4.2

Notes. These estimates are based on a calibration of the stylized neoclassical model presented in Section II.

Table 2.

Comparison of Government Investment and Consumption Multipliers

in DSGE Models

A. Neoclassical models

<table>
<tr><th>Model</th><th>Experiment</th><th>Government consumption</th><th>Government investment</th></tr>
<tr><td rowspan="2">Baxter-King (1993) – benchmark neoclassical model

Lump-sum taxation.</td><td>Impact multiplier for various durations of spending.

2-yrs of spending
Permanent spending</td><td>

0.4
0.9</td><td></td></tr>
<tr><td>Long-run Multipliers:

θ_G = 0 (e.g. gov consumption)
θ_G = 0.05
θ_G = 0.40</td><td></td><td>1.2
2.6
13</td></tr>
<tr><td rowspan="2">Leeper, Walker, Yang (2010) – time-to-build delay

Persistent govt spending - ρ = 0.95

Estimated finance response, including distortionary taxes.</td><td>PV multipliers θ_G=0.05

1 quarter time-to-build delay
3 year time-to-build delay
Across delay times</td><td></td><td>

SR: 0.5
SR: 0.3
LR: 0.3-0.4</td></tr>
<tr><td>PV multipliers θ_G=0.1

1 quarter time-to-build delay
3 year time-to-build delay
All delay times</td><td></td><td>

SR: 0.5
SR: 0.1
LR: 0.9-1.1</td></tr>
</table>

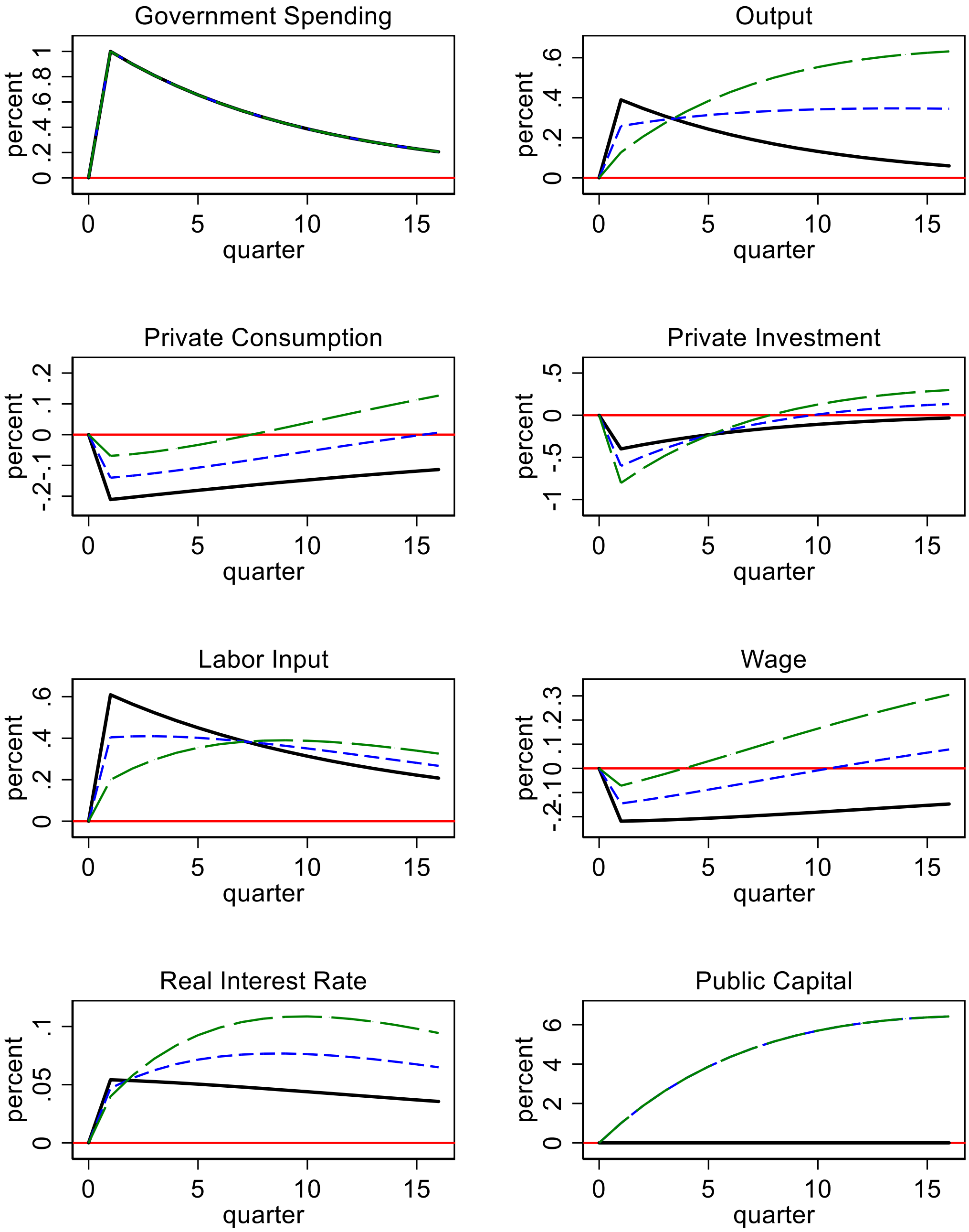
B. New Keynesian Models

Model	Experiment	Government consumption	Government investment
Coenen plus 17 co-authors (2012) Large scale policy models + 2 academic models. U.S. estimates	Avg. first year multipliers for a 2-year stimulus, financed with deficits No monetary accommodation 1 yr monetary accommodation 2 yrs monetary accommodation	 0.9 1.2 1.6	 0.9 1.1 1.6
Coenen, Straub, Trabandt (2013) Estimated ECB's Euro-area model	Estimated responses based on standard monetary and fiscal (e.g. tax) reaction Years 1 - 4 Long-run	 1 – 1.2 0.8	 0.7 – 1 1.1
	2-year stimulus, 2 yrs of monetary and fiscal accommodation (i.e. no tax increases in SR) Years 1 - 4 Long-run	 1.3 – 1.7 1.6	 1.1 1.6
Albertini, Poirier, Roulleau-Pasdeloup (2014) – Calibrated model. No private capital, but public capital with θ_G = 0.08.	Stylized stimulus package (AR), lump-sum taxes. Impact multiplier Outside the ZLB At ZLB	 0.6 1.5	 0.75 1
Drautzburg-Uhlig (2015) – Estimated medium scale model with additional elements.	ARRA, with distortionary taxation later. Short run Long run	 0.8 to 1 -0.1	 0.2 to 0.5 0.3
Boehm (forthcoming) – calibrated 2 sector model (consumption and investment), imperfect labor mobility in SR , with	AR(1) govt spending (ρ= 0.86) SR (0 to 20 quarters) LR	 0.4 to 0.7 1.1	 0.1 - 0.2 1.6

θ_G = 0.05, lump sum taxes.			
Bouakez, Guillard, Roulleau-Pasdeloup (2017) -Calibrated model with time to build in public capital, with θ_G = 0.08	AR(1) govt spending (ρ= 0.86), lump sum taxes Impact multiplier Normal times 1-quarter time to build delay 4-year time to build delay Liquidity trap 1-quarter time to build delay 4-year time to build delay	 0.9 0.9 2.3 2.3	 0.9 0.8 2.3 4.0

Note: θ_G is the elasticity of output to public capital in the aggregate production function.

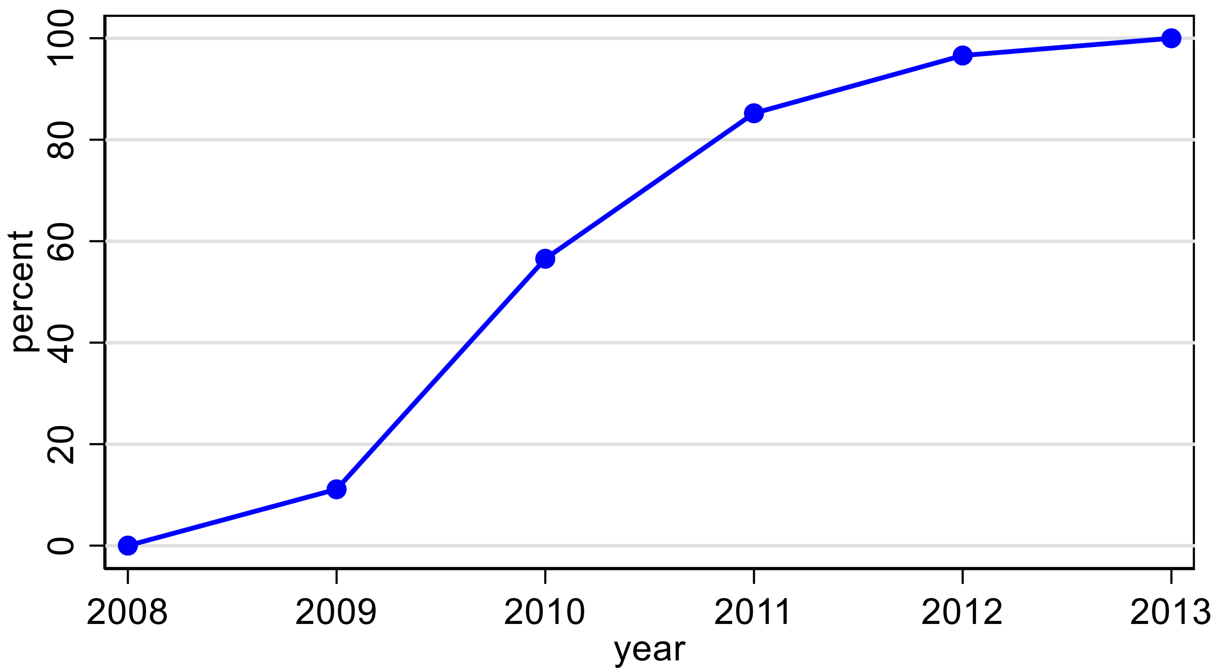
Figure 1. Effect of Increases in Government Consumption or Investment
Baseline Neoclassical Model



Notes. Black solid: government consumption shock; blue short dashed: government investment shock, $\theta_G = 0.05$; green long dashed: government investment, is $\theta_G = 0.1$. Government spending, output, consumption, private investment, and public capital are expressed in deviations from steady state value as a percent of output in steady state. Labor input and wages are percent deviations from their own steady state values. Real interest rate is percentage point deviations from its own steady state.

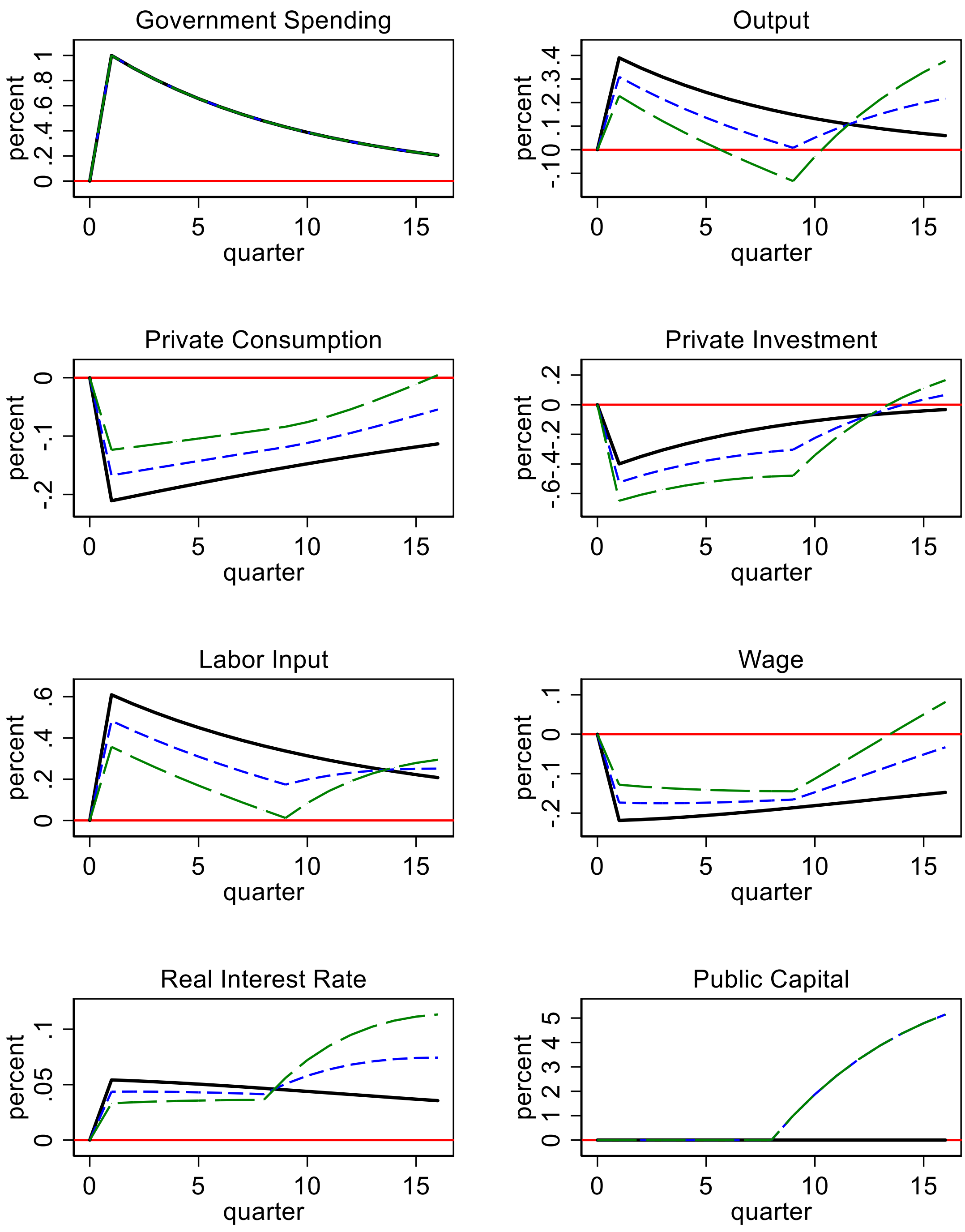
Figure 2. Time Pattern of Federal Highway Administration Outlays from the ARRA

Cumulative Percent Spent of Total Appropriation



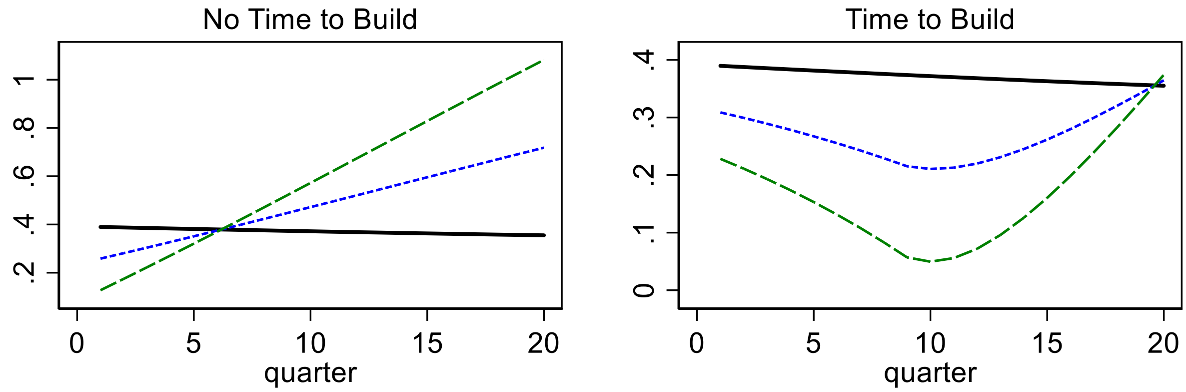
Notes. These data are from Leduc and Wilson's (2017) replication files. I aggregated their state-level data to the national level.

Figure 3. Effect of Increases in Government Consumption or Investment
8-Quarter Time To Build



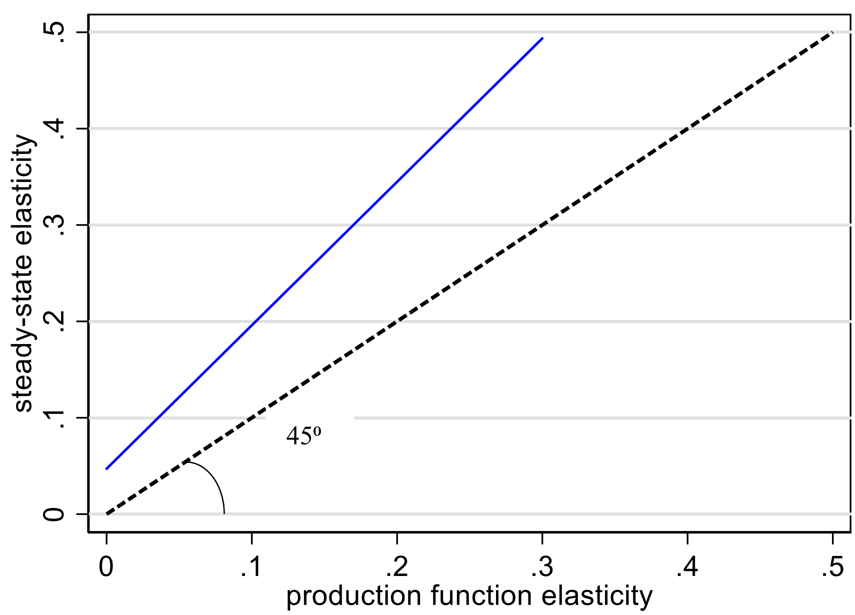
Notes. Black solid: government consumption shock; blue short dashed: government investment shock, $\theta_G = 0.05$; green long dashed: government investment, is $\theta_G = 0.1$. Government spending, output, consumption, private investment, and public capital are expressed in deviations from steady state value as a percent of output in steady state. Labor input and wages are percent deviations from their own steady state values. Real interest rate is percentage point deviations from its own steady state.

Figure 4. Present Discounted Value Integral Multipliers



Notes. Black solid: government consumption shock; blue short dashed: government investment shock, $\theta_G = 0.05$; green long dashed: government investment, is $\theta_G = 0.1$. These estimates are based on the calibrated neoclassical model of Section II.

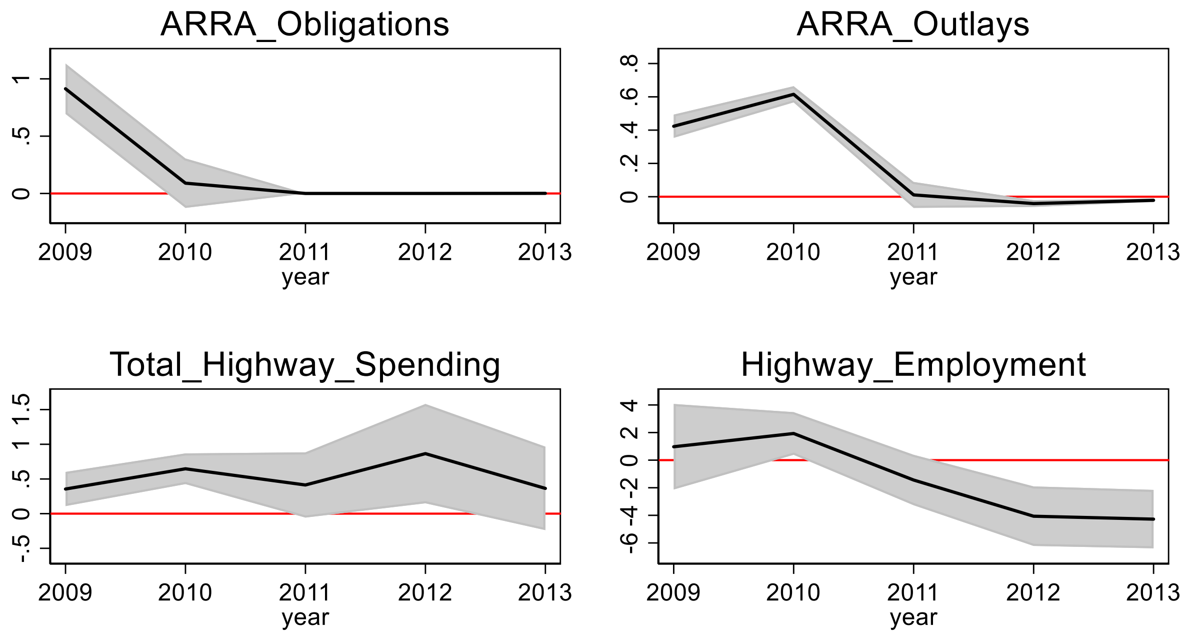
Figure 5. Relationship between the Elasticity of Steady-State Output to Public Capital and the Production Function Elasticity



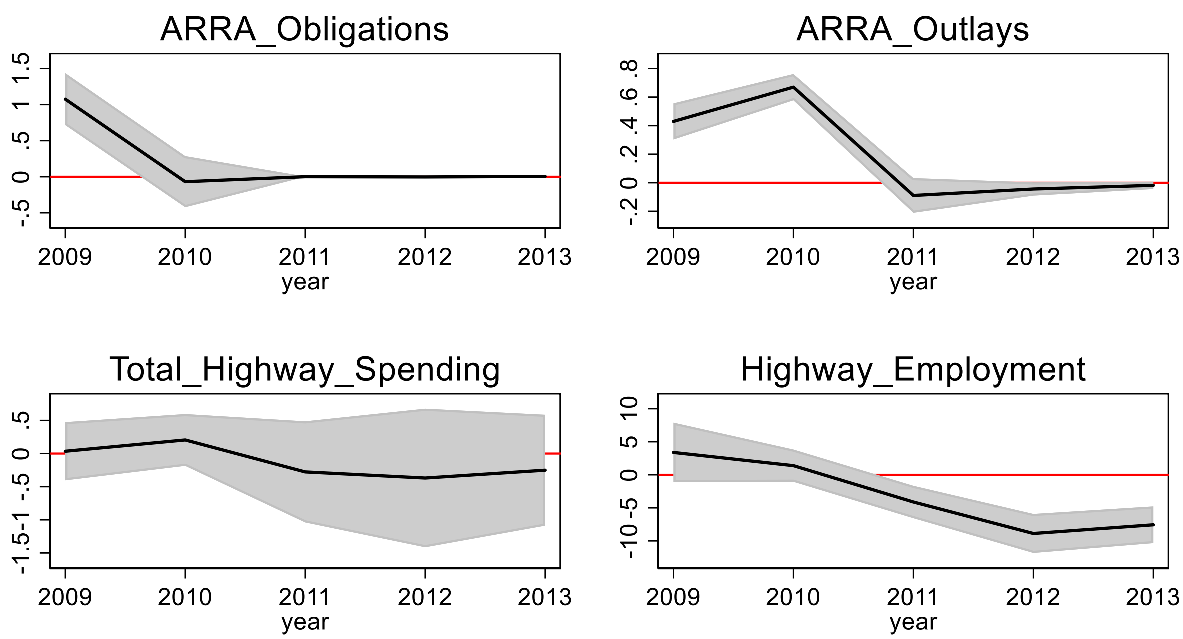
Notes. Based on simulations of the calibrated model in Section II. The relationship is: GE elasticity = 0.047 + 1.49·θ_G for this calibration.

Figure 6
Estimated Impulse Responses to Instrumented ARRA Highway Apportionments

A. **No controls for log state population.**



B. Control for log state population.



(See notes on next page.)

Notes to Figure 6: The obligations, outlays and total highway spending graphs show the impact in dollars of a dollar of ARRA highway apportionments in 2009. The highway employment graph shows the employment impact in highway, road, and bridge construction employment for each $1 million of ARRA highway apportionments in 2009. In both cases, the ARRA apportionments are instrumented by Leduc and Wilson's (2017) two road factors. Controls: LW political controls + LS.roadconstruction_emp_pc (Regressions are not weighted.) The confidence bands are 90 percent bands.