The Local Multiplier: Theory and Evidence

Brock Mendel*

February 24, 2013

First Version: November 3, 2011

Abstract

I show that 1) the policy-relevant “global multiplier” can be written as the sum of a spending component and a taxation component, all scaled up by spillover effects, 2) the “local multiplier” is exactly the spending component, and 3) if trade is anonymous, the local effects of a shock to federal government purchases in a county will be identical to the effects of a shock to consumer demand for the exports of that locality. I estimate a bound for the local multiplier and consider spillover effects to contiguous counties. I find that a shock of $48,000 creates at least one job-year locally. Analysis at a monthly frequency suggests that these jobs are more persistent than previously estimated. Evidence of higher multipliers in recessions is mixed.

1 Introduction

The debate over the effect of exogenous government spending on economic outcomes has been stymied by a lack of exogenous variation in the time-series of spending. A recent wave of the literature has considered government spending at a local level and conducted instrumental variables analysis on the cross-sectional distribution of this spending. This approach relies on the strength and validity of the instruments used. I get around this limitation by showing that the coefficient on government spending in any regression, if properly identified, will match the coefficient on certain firm-level revenue shocks. The latter is readily identified using OLS.

To see this intuitively, consider two thought experiments. In the first, the federal government exogenously and unexpectedly decides to increase spending on iPods by one dollar. In the second, the same decision is made by a randomly chosen American consumer. The

---

*jbmandel@fas.harvard.edu. I would like to thank the National Science Foundation for financial support. I would like to acknowledge helpful comments from Adam Clark-Joseph, Gabriel Chodorow-Reich, Brad DeLong, Emmanuel Farhi, Greg Mankiw, Adam Guren, David Laibson, Christina Romer, David Romer, Andrei Shleifer, Alp Simsek, Jenny Tang, Laurence Wilse-Samson, and participants at the Harvard Macro Lunch.
national effects of these two shocks may be different (particularly depending on what the
government does with its iPods in the first example), but Apple cannot tell the difference.
Apple only observes that its demand exceeded expectations by a dollar.

Up to an approximation ignoring the changes in aggregate variables in these scenarios,
Apple’s endogenous response is identical in these two cases. Apple’s hiring decisions, invest-
ment decisions, and pricing decisions are all the same in these two cases. This argument
then extends to the agents that Apple directly interacts with, particularly the segment of
the labor market it hires in. This line of logic breaks down as I consider ever larger subsets
of the economy, but holds as an approximation locally.

I formalize this argument in a non-parametric structural model. It holds whenever trade
is anonymous, or more generally whenever the persistence of shocks to government spending
matches the persistence to shocks to non-government spending.

The local multiplier literature has not identified the precise relationship between the local
multiplier and the more policy-relevant global multiplier. Many of these papers have conjec-
tured that spillover effects will make the global multiplier larger than the local multiplier. I
examine the matter theoretically and find that this may only be true in certain special cases
depending on the network structure of spillovers.

Intuitively, think of the net marginal effect of government spending as the sum of two
terms. The first term is the Keynesian term. A marginal dollar of spending increases the
demand faced by some firms, who then in turn have to increase their labor demand. The
newly employed workers have a high marginal propensity to consume, and their consumption
stimulates other businesses in the area, continuing the process.

The second term is a Ricardian term. The marginal dollar of spending implies a marginal
dollar of taxes that the government is going to have to levy on someone eventually. Agents
with rational expectations see this, so in response to the marginal government spending,
save more and spend less. This counteracts some portion of the Keynesian term. The sign
and magnitude of the multiplier depends on the sum of these two terms.

The central appeal of the local multiplier literature is that it decouples the agents im-
pacted by the Keynesian term in the multiplier– the agents in a local economy– from the
agents impacted by the Ricardian term in the multiplier– the agents in the global econ-
omy. As I show, the transition from the local multiplier to the global multiplier does involve
spillover effects, but it also involves adding back in the Ricardian term. These spillover ef-
tects also apply to the Ricardian term, so there is no free lunch in transitioning to the global
multiplier.

After considering the theory of the local multiplier, I proceed to estimate it using local
data. I map firm-level revenue surprises into the counties where those firms are headquar-
tered. I then regress employment growth in those counties on these revenue surprises. Under
the restrictions from the theoretical section, this regression consistently identifies the impact
of surprise government spending on employment growth. I include the effects of surprises
to revenues for firms located in a neighboring county. I find that one year of one job is
created for every $48,000 in spending. The validity of this procedure implicitly assumes that
firms conduct all of their operations in the county containing their headquarters, which is
obviously untrue. I therefore think of my estimate an upper bound and not a point estimate.

The jobs created by a revenue surprise in a county are not isolated to the firm that received the shock, nor even to the industry. Roughly half of the job growth due to a revenue shock in a county occurs in non-shocked industries. The jobs created by these shocks are temporary, but not as temporary as identified in the previous literature (see Shoag ([45], 2011)). I find that an appreciable portion of the jobs created persist six years after the original shock.

The literature has hypothesized (see Auerbach and Gorodnichenko ([2], 2011)) that multipliers may be higher in recessions than they are in normal times. This may be in part due to the zero lower bound on nominal interest rates (see Woodford ([48], 2010)). I examine the multiplier separately during recessions and find that the initial job impact of spending is higher in recessions, but then that these extra jobs disappear faster and in fact turn negative about two years after the spending occurs. This may be a figment of the short sample size, but is at best mixed evidence for the hypothesis of a higher multiplier in recessions.

1.1 Many Different Multipliers

There is no precise definition of “the” government spending multiplier. There are many different ways in which the government can spend, tax, or rebate money. There is no theoretical reason to think that these will have equivalent effects, and strong empirical reason to think that they won’t.

The traditional method of modeling government expenditure is as fully wasteful spending, goods bought and then thrown into the sea. For some portion of spending this may be an appropriate modeling device, but a much larger share of the federal budget goes into direct transfers to individuals, state and local governments, investment in public capital, or investment in public goods. Adjustment along any of these margins, or along any margins of the tax code, could reasonably be called “stimulus” and have its own distinct multiplier.

The literature has identified multipliers for at least six different types of expenditures; debating which of these is “the” correct multiplier misses the point. There is a literature identifying the effects of federal government spending on the military, another using a narrative approach to identify the effects of exogenous changes to the tax code, a third using a VAR approach to identify innovations in total government spending, a fourth looking at the stimulative impact of tax rebates delivered in different formats, a fifth examining the effects of a liquidity shock to a state governments, and a sixth considering exogenous changes in federal grants to states.

I review each of these below. The important point here is that there is no ex-ante reason to expect all of these to identify the same multiplier. In fact, it is more surprising that so much of the local multiplier literature does manage to find consensus.

My approach to estimating the local multiplier is locally equivalent to buying goods and throwing them into the sea but globally robust to several different types of government spending. Any type of government spending that buys goods from one county and then transfers them elsewhere will fall under the category of the local multiplier I estimate.
1.2 Measurement and Theory

This paper is, to the best of my knowledge, the first to prove results linking the local multiplier to the global multiplier in a theoretical framework. The model setup requires only that trade be anonymous and that the structural model governing the world be continuously differentiable.

Under this setup, I prove that the effect of surprise consumer spending on the goods and services exported by a county must be the same as the effect of surprise government spending on those same goods and services. Under sticky prices, the surprise to consumer spending is unrelated to county conditions, and is in fact an idiosyncratic variable uncorrelated with every other pre-determined variable that could go into a regression. It therefore has the appropriate correlation properties to identify its effect on county outcomes, and through it to identify the effect of government spending on county outcomes.

Most attempts to estimate a government spending multiplier look for exogenous variation in government spending. In some cases this is taken as residuals from a VAR, in some cases this is identified manually through the narrative approach, and in many cases this is done by instrumenting. My approach uses a theoretically motivated identifying restriction to get around the need to find exogenous government spending. Because data on firm-level revenue expectations are directly available, I do not have to proxy for or estimate expected levels of spending. So long as these expectations are reasonably close to rational, revenue surprises will be uncorrelated with all pre-determined variables. This obviates the need to worry about omitted variables bias for the vast majority of potential variables. Moreover, because I directly observe market expectations, I do not need to proxy for or estimate the unexpected portion of spending. I can directly estimate the result of the surprise.

2 Literature Review

2.1 The Global Multiplier

Ramey ([37], 2011) gives a thorough review of the literature on the government spending multiplier, highlighting the caveats about the different multipliers corresponding to different types of government spending or tax changes. The review that follows is necessarily incomplete.

Several papers focus specifically on the response of the economy to the American Recovery and Reinvestment Act of 2009 (ARRA). Romer and Bernstein ([5], 2009) use a large-scale econometric model to estimate the effect of the ARRA. They use a time-path of multipliers plateauing at 1.57 after eight quarters for spending and .99 for tax cuts. Denes and Eggertson ([17], 2009) use Bayesian techniques to estimate that the 2009 stimulus increased output by 3.6% in 2009 and 2010. Blinder and Zandi ([8], 2010) use a large-scale econometric model to argue that the 2009 stimulus and other programs like the Troubled Asset Relief Program averted a “Great Depression 2.0.” On the opposite side of this debate are Cogan and Taylor ([13], 2011), who find that as of the 2nd quarter of 2010, only 2% of the 2009 stimulus had
actually been used to increase government purchases.

A popular strand of the literature uses VAR analysis to identify exogenous shocks to taxes and spending and examine their dynamic effects. Blanchard and Perotti ([7], 2002) use institutional information about tax and transfer systems to construct a structural VAR and estimate the effects of spending and tax changes. They find that a positive shock to spending increases consumption, consistent with a Keynesian model. They also find that both increases to taxes and spending decrease investment, which is inconsistent with Keynesian models. Their identifying assumption, now standard in the literature, is that fiscal policy reacts to economic shocks only with a lag. Perotti ([35], 2004) uses a VAR approach to estimate the effect of fiscal policy on outcomes in OECD countries. He finds a multiplier that has decreased over time and no evidence that tax cuts work faster or more effectively than spending increases. Mountford and Uhlig ([30], 2005) use a VAR analysis to find that the most effective way to stimulate the economy is a deficit-financed tax cut.

A more recent approach is the “Narrative Approach,” where researchers look directly at historical documents to ascertain the motivations of policy changes. Romer and Romer ([39], 2010) comb the narrative record for plausibly exogenous changes in tax policy. They find that an exogenous tax increase of 1% of GDP decreases GDP by 3%, a very large multiplier. Ramey ([36], 2007) constructs a measure of exogenous defense spending shocks using a narrative approach to measure expectations. She finds that the government spending shocks identified in the literature are often anticipated, invalidating the VAR results on the economy’s dynamic response to these shocks.

Many authors consider military spending to be at least partially exogenous, on the theory that wars are not caused by domestic economic conditions. Ramey and Shapiro ([38], 1998) examine the effect of military buildups on macroeconomic variables, find evidence supporting the predictions of a two-sector neoclassical model. Barro and Redlick ([4], 2009) examine defense spending in a sample including World War II and find a multiplier of .6-.7. They find the multiplier rises with the unemployment rate. The results in these exercises are often sensitive to inclusion of the World War II and Korean War eras.

There is also a significant literature examining the multipliers predicted by different types of models. Barro ([3], 1981) analyzes the theoretical difference between temporary and permanent innovations to government purchases. Hall ([24], 2009) concludes based on military spending evidence that the multiplier is at least .5, discusses the relationship between variable markups and the multiplier. Leeper, Traum, and Walker ([25], 2011) conduct a Bayesian estimation of five structural models to identify which aspects of these models drive their predictions about the multiplier. Cogan, Cwik, Taylor, and Wieland ([12], 2009) examine the difference between new and old Keynesian multipliers, find that the multipliers implied by new Keynesian models (Smets and Wouters ([44], 2007) in particular) are much smaller than those implied by old Keynesian models (Romer and Bernstein ([5], 2009)). Nekarda and Ramey ([33], 2010) find that an increase in government purchases in a sector increases output and hours in that sector but decreases real wages. They argue this is consistent with a neoclassical model and inconsistent with the standard New Keynesian model.
The Keynesian viewpoint suggests that multipliers will be higher in times when there is slack in the economy. This cannot be identified by a purely linear model, so requires a more sophisticated analysis. Auerbach and Gorodnichenko ([2], 2011) use a regime-switching model to estimate multipliers both in and out of recessions. They find that multipliers are higher in recessions. They also include data from the Survey of Professional Forecasters and find that controlling for these forecasts increases the size of estimated multipliers in a recession. Mian and Sufi ([28], 2010) examine the effect of the ‘Cash for Clunkers’ program, find that car purchases increased by 360,000 in July and August of 2009, but that nearly all of these would have been purchased in the near future anyway. They find no effect on employment or other variables of interest. Gordon and Krenn ([22], 2010) use a VAR analysis to find a multiplier of .88 for capacity-constrained periods and 1.8 for unconstrained periods.

Christiano, Eichenbaum and Rebelo ([9], 2011) argue theoretically that the government spending multiplier will be large when the zero lower bound on nominal interest rates is binding. Woodford ([48], 2010) shows that the multiplier depends crucially on the response of monetary policy to government spending, and that the multiplier may be well over unity when monetary policy is constrained by the zero lower bound.

One of the subtleties of the multiplier literature is that realistic fiscal policy has many dimensions. Several papers have analyzed the differences in the impacts of different types of stimulus. Mertens and Ravn ([27], 2011) break down the response of economic outcomes to tax changes based on those that are anticipated and unanticipated. They find that tax cuts reduce output on announcement, but increase output on implementation. Sahm, Shapiro, and Slemrod ([40], 2010) ask whether the stimulative effect of tax cuts depends on the delivery method. They find that a single salient payment has a larger effect on marginal consumption than reductions in paycheck withholdings. Shapiro and Slemrod ([42], 2001) find that only 22% of households receiving a tax rebate in the 2001 stimulus actually spent it whereas the others saved it or used it to pay off debt. Parker, Souleles, Johnson, and McClelland ([34], 2010) find that households spent an average of 50-90% of their 2008 stimulus payments in the three month period after the payments were received.

A strand of the literature has considered the hypothesis that fiscal consolidations– budget cuts and tax increases designed to improve a country’s debt to GDP ratio– may have a stimulative impact, counter to the Keynesian intuition. Giavazzi and Pagano ([21], 1990) find that the 1980s fiscal consolidations of Denmark and Ireland led to increases in current and planned consumption. Alesina and Ardagna ([1], 2009) find in OECD countries that fiscal stimulus based on tax cuts is more effective than stimulus based on spending, and that fiscal adjustments based on cutting spending are more likely to decrease deficits and less likely to create recessions than those based on tax increases. Guajardo, Leigh, and Pescatori ([23], 2011) use a narrative approach to examine the historical record and find that fiscal consolidations based on a desire to reduce budget deficits result in declines in consumption and GDP. They find that a 1% fiscal consolidation causes a decline in real GDP of .62%.
2.2 Local Multipliers

A more recent literature considers the impact of government spending on local economic outcomes. While my review of the global multiplier literature was necessarily a partial one, the following review covers every paper I have found on local multipliers as of December 21, 2011.

Each of these papers uses some instrument for state or federal spending at a state or local level, and measures subsequent economic outcomes at that disaggregated level. Each measures a slightly different type of government spending, each financed in different ways. It is therefore no surprise that they come to slightly different conclusions about the multiplier. These papers each include the warning that it is not clear how to translate the local multiplier they identify into the global multiplier identified by the traditional literature. In the theory section below, I show how to perform that translation.

A strand of this literature finds costs per job-year in the $25,000-$35,000 range. Shoag ([45], 2011) uses variation in state pension fund returns to instrument for state government spending. He finds a multiplier of 2.12 and that $35,000 in spending generates an additional job-year. He finds that these job gains are largely transitory. Nakamura and Steinsson ([32], 2011) use variation in military spending across states to instrument for federal government spending at the state level. They find a local multiplier of 1.5. Serrato and Wingender ([41], 2011) instrument for local government spending using the fact that population is estimated differently in non-census years and census years and spending responds to population estimates. They find that federal spending has a local multiplier of 1.88 and creates jobs at a cost of $30,000 per job-year. Clemens and Miran ([11], 2010) use the stringency of state balanced-budget rules to estimate the effect of state budget cuts on state-level employment. They find that avoiding a $25,000 mid-year spending cut saves one job-year, corresponding to a local multiplier of 1.7. Chodorow-Reich, Feiveson, Liscow, and Woolston ([10], 2011) investigate the effect of the ARRA on local employment using pre-recession variation in Medicaid spending to instrument for fiscal transfers. They find that a hundred thousand dollar transfer creates 3.8 job-years locally, 3.2 of which are outside of the government, health, and education sectors. This corresponds to $26,000 per job-year. They note that in 2008 the average compensation of employees was $56,000, though data limitations prevent them from ascertaining the pay of the jobs created by the ARRA.

A second strand of this literature finds much less encouraging local multipliers. Cullen and Fishback ([16], 2006) find that World War II spending had very little effect on the consumption growth rates of local economies at a state and county level. Fishback and Kachanovskaya ([20], 2011) instrument using the political competitiveness of different states to examine annual federal spending by state during the Depression and find a state multiplier of 1.1. Cohen, Coval and Malloy ([14], 2011) use changes in congressional committee membership to instrument for federal government spending at a state level. They show that when a politician ascends to the chairmanship of a powerful committee, her state’s federal transfers increase by 9-10% and federal contracts in her state increase by 24%. They find that spending shocks tend to dampen corporate investment and employment. They find that their seniority shocks lead to declines in GDP, employment, and personal income on the order
of 5%. These declines are concentrated in Retail, Construction, and Manufacturing. Wilson ([47], 2011) finds that the ARRA created jobs in the first year at a cost of $100,000 per job, but that these jobs were short-lived. Feyrer and Sacerdote ([19], 2011) use Congressional seniority to instrument for state and county level spending in the 2009 stimulus,. They estimate that the stimulus created jobs at a cost per job-year of $170,000, but note that excluding education spending their estimate is below $100,000 corresponding to a multiplier of 2. Conley and Dupor ([15], 2011) use highway funding from the ARRA to estimate that the act created/saved 450 thousand public sector jobs and destroyed/forestalled one million private sector jobs. They instrument based on formulas for Department of Transportation spending defined before The Great Recession.

Moretti ([29], 2010) finds that for each job created in manufacturing in a given city, 1.6 jobs are created in the non-tradable sector in the same city. Broken down into skilled and unskilled labor, the effects are 2.5 and 1, respectively. He claims that the multiplier for the non-tradable sector measured locally is an upper bound for the global multiplier, while the multiplier for tradables is a lower bound. I show that these statements need not hold because of potential spillover effects.

3 Model

3.1 Time

Time is discrete and indexed by $t = 1, 2, \ldots$ An econometrician has a sample of length $T$. In the empirical work below, these units are months.

3.2 Geography

A closed economy is separated into $C$ geographic units called “counties” and indexed by $c = 1, 2, \ldots, C$. A county is a unit small enough so that I may consider it vanishingly small as compared to the entire economy. A Law of Large Numbers applies whenever I sum variables across counties.\footnote{Technically, Laws of Large Numbers apply to averages, not sums. This is one of many wrinkles that can be ironed out, but only at the cost of considerable additional algebraic detail.}

3.3 Variables and Shocks

Three aggregate variables are determined each period. First is an exogenous shock to the economy $\epsilon_t$. Second is a shock to consumer spending, $\hat{d}_t$, which may be correlated with $\epsilon_t$. Third is a shock to federal government spending, $\hat{g}_t$, which may be correlated with both $\epsilon_t$ and $\hat{d}_t$. These shocks may also be autocorrelated.

Three county-specific variables are determined each period. First is an exogenous shock to the local economy $\epsilon_{ct}$. Second is a shock to federal government spending in the county $\hat{g}_{ct}$, which may be correlated with $\epsilon_{ct}$. The third and most interesting shock, $\hat{d}_{ct}$, is a shock
to the desirability of the goods and services produced by firms located in county \( c \). This is a shock to the preferences of consumers located outside the county, and so is independent of the idiosyncratic conditions of the county. Fiscal policy cannot react quickly and precisely enough for \( \hat{g}_{ct} \) to respond to \( \hat{d}_{ct} \), so these two shocks are uncorrelated.\(^2\)

The shocks \( \epsilon_{ct} \) and \( \hat{g}_{ct} \) may be spatially correlated, but are globally mean-zero each period.

### 3.4 Market Structure

A subset of firms in county \( c \) are large national firms, producing locally but exporting goods and services to the entire economy. Local conditions in county \( c \) may affect the cost curve faced by a firm, but will not affect its demand curve because the county is vanishingly small compared to the entire economy.

Prices are sticky in the short run. The exact details of the price setting mechanism are irrelevant so long as prices for time \( t \) are set no later than the end of time \( t - 1 \). This is empirically reasonable. Bils and Klenow ([6], 2004) find that the median duration of a price is 4.3 months, or 5.5 months if sales are excluded. Nakamura and Steinsson ([31], 2008) find that the median frequency of non sale price change is 9-12% per month. Modeling prices as predetermined at a monthly frequency is therefore a mild restriction.

Given the prices set by the large national firms in county \( c \), they face total consumer demand tautologically given by:

\[
d_{ct} = \bar{E}_{t-1}[d_{ct}] + \hat{d}_{t} + \hat{d}_{ct} \tag{1}
\]

and total federal government demand given by:

\[
g_{ct} = \bar{E}_{t-1}[g_{ct}] + \hat{g}_{t} + \hat{g}_{ct} \tag{2}
\]

To simplify the math below, I assume that \( \bar{E}_{t-1}[d_{ct}] = \bar{d}, \bar{E}_{t-1}[g_{ct}] = \bar{g} \) for all \( c, t \). This assumption can be dispensed with at the cost of considerable algebraic complexity.

Trade is anonymous, so that agents can only observe the sum \( d_{ct} + g_{ct} \) but not the individual components\(^3\). This is the key structural assumption driving the results below.

### 3.5 Federal Government

The government’s total spending in period \( t \) is \( \sum_{c} g_{ct} \). For simplicity I assume that the government has to run a balanced budget each period, so must raise taxes \( T_{ct} \) from agents in each county \( c \) totaling to \( \sum_{c} T_{ct} = \sum_{c} g_{ct} \). Taxes on agents in county \( c \) are given by:

\[
T_{ct} = \bar{g} + \hat{g}_{t} \tag{3}
\]

\(^2\)I confirm this restriction empirically below.

\(^3\)This assumption can be weakened significantly. It is sufficient for the two shocks to have the same persistence and informational properties.
Because the shocks $\hat{g}_{ct}$ are globally mean-zero, this tax policy satisfies the government’s budget constraint. The results below are invariant to deterministic redistributive policies. Loosening the government’s budget constraint to a more standard infinite-horizon budget constraint changes the character of $F_4$ below, but does not change the intuitive results.

### 3.6 Equilibrium

I make no attempt to model the endogenous behavior of a county economy. I assume only that there exists a well-behaved structural model. The endogenous behavior of the economy is described by a vector $x_{ct}$ of outcome variables. These outcomes depend on 1) $\epsilon_t$, 2) $\epsilon_{ct}$, 3) $d_{ct} + g_{ct}$, and 4) $T_{ct}$. A smooth structural model will have an equilibrium given by the solution to a system of functional equations. The implicit solution to such a system can be written as the output of a black box:

$$x_{ct} = F(\epsilon_t, \epsilon_{ct}, d_{ct} + g_{ct}, T_{ct})$$  \hspace{1cm} (4)

Linearizing around $(0, 0, \bar{d} + \bar{g}, \bar{g})$ and rearranging, this becomes:

$$\hat{x}_{ct} = F_3(\hat{d}_{t} + \hat{d}_{ct} + \hat{g}_{ct}) + (F_3 + F_4)\hat{g}_t + F_1\epsilon_t + F_2\epsilon_{ct}$$ \hspace{1cm} (5)

The estimability of this equation depends on the correlation structure between the observables $d_t$, $d_{ct}$, $g_{ct}$, $g_t$ and the unobservables $\epsilon_t$, $\epsilon_{ct}$. $\hat{g}_t$ is correlated with $\epsilon_t$, so OLS on $\hat{g}_t$ may give a biased estimate for $F_3 + F_4$. $\hat{g}_{ct}$ may be correlated with $\epsilon_{ct}$, so OLS on $\hat{g}_{ct}$ may give a biased estimate for $F_3$.

But by definition $\hat{d}_{ct}$ is uncorrelated with every other term in this equation. This gives the first main result: OLS on $\hat{d}_{ct}$ will give an unbiased estimate of $F_3$. This is assuming that the econometrician has data on $\hat{d}_{ct}$. I show how to construct this data below.

### 3.7 Extension: Cross-County Spillovers

Consumers and factors of production are mobile across county lines. It is then important to consider the effects of outcomes in one county on outcomes in another county. A completely general formulation of these spillovers would require $F$ to be county-specific:

$$x_{ct} = F_c(x_{1t}, x_{2t}, ..., x_{Ct}, \epsilon_t, \epsilon_{ct}, d_{ct} + g_{ct}, T_{ct})$$ \hspace{1cm} (6)

The downside of this formulation is that, once linearized, the estimation allows for too many degrees of freedom. To circumvent this problem, I need to sacrifice some generality in the functional form. It is reasonable to think that counties that are contiguous have stronger links than counties on opposite sides of the country. I therefore write instead:

$$x_{ct} = F(g_c(x_{1t}, x_{2t}, ..., x_{Ct}), \epsilon_t, \epsilon_{ct}, d_{ct} + g_{ct}, T_{ct})$$ \hspace{1cm} (7)
where the function $g_c$ is allowed to vary across counties to reflect the fact that different counties are connected in different ways. For simplicity suppose that $x_{ct}$ is one-dimensional\(^4\). Then this equation linearizes into:

\[
\hat{x}_{ct} = F_1 \sum_{c'} g_{cc'} \hat{x}_{c't} + F_2 \epsilon_t + F_3 \epsilon_{ct} + F_4 (\hat{d}_t + \hat{d}_{ct} + \hat{g}_t + \hat{g}_{ct}) + F_5 \hat{g}_t
\]  

(8)

Define the vector $x_t = (\hat{x}_{1t}, \hat{x}_{2t}, ..., \hat{x}_{Ct})'$. Let $G$ be the $C \times C$ matrix with $g_{cc'}$ in its $(c, c')$ position. Let $y_t$ be the $C \times 1$ vector with $F_2 \epsilon_t + F_3 \epsilon_{ct} + F_4 (\hat{d}_t + \hat{d}_{ct} + \hat{g}_t + \hat{g}_{ct}) + F_5 \hat{g}_t$ in the $c^{th}$ row. Then this last equation becomes:

\[
x_t = F_1 G x_t + y_t
\]  

(9)

Let $I_C$ be the $C \times C$ identity matrix. Then there exists a unique solution for $x_t$ if and only if the matrix $I_C - F_1 G$ is invertible. In this case, $x_t$ is given by:

\[
x_t = (I_C - F_1 G)^{-1} y_t
\]  

(10)

This suffices to show the second main result: depending on where government spending is targeted, spillover effects may cause the global multiplier to exceed the local multiplier. To see this, consider an example with $C = 3$ where the global variable of interest is $\hat{x}_t = \sum_c \hat{x}_{ct}$. County 1 is connected only to county 2, county 2 is connected to counties 1 and 3, and county 3 is connected only to county 2. The matrix $G$ is given by:

\[
\begin{pmatrix}
0 & 1 & 0 \\
1 & 0 & 1 \\
0 & 1 & 0
\end{pmatrix}
\]

and I let $F_1 = .5$. Then $I_C - F_1 G$ is given by:

\[
\begin{pmatrix}
1.5 & 1.0 & 0.5 \\
1.0 & 2.0 & 1.0 \\
0.5 & 1.0 & 1.5
\end{pmatrix}
\]

For ease, suppose that $\epsilon_t = \epsilon_{ct} = \hat{d}_t = \hat{d}_{ct} = 0$, so I can drop these terms. Then the last equation becomes:

---
\(^4\)In the case when $x_{ct}$ is a higher-dimensional vector, linearizing gives a system of equations that can be stacked and transformed into a matrix equation called a Sylvester Equation. This type of equation often has a solution, but not in closed form. As far as I can tell, no new economic intuition is gained from this case, so I exclude it from the exposition.
\[
\begin{pmatrix}
  x_{1t} \\
  x_{2t} \\
  x_{3t}
\end{pmatrix} = \begin{pmatrix}
  1.5 & 1.0 & 0.5 \\
  1.0 & 2.0 & 1.0 \\
  0.5 & 1.0 & 1.5
\end{pmatrix} \begin{pmatrix}
  F_4(\hat{g}_t + \hat{g}_{1t}) + F_5\hat{g}_t \\
  F_4(\hat{g}_t + \hat{g}_{2t}) + F_5\hat{g}_t \\
  F_4(\hat{g}_t + \hat{g}_{3t}) + F_5\hat{g}_t
\end{pmatrix}
\]

Summing across counties gives:

\[
\hat{x}_t = 10(F_4 + F_5)\hat{g}_t + 3F_4\hat{g}_{1t} + 4F_4\hat{g}_{2t} + 3\hat{g}_{3t}
\tag{11}
\]

Recall that the balanced budget constraint requires \(\hat{g}_{1t} + \hat{g}_{2t} + \hat{g}_{3t} = 0\). This equation shows two things: first, if \(F_4 + F_5\) does not give a positive net multiplier on total government spending at the local level, then spillover effects will not flip that result at the global level. Second, if some counties are more connected than others, targeted spending at those counties can produce positive spillover multipliers. In this case, taking a marginal dollar from county one or three and spending it in county two—assuming \(F_4 > 0\)—will generate enough spillover gains to increase \(\hat{x}_t\). Of course, this holds only in the region of small shocks where the linearization is valid.

This property cannot hold in a two-county model unless the counties are asymmetrically connected.

4 Data

4.1 I/B/E/S and Compustat

I begin with data from the Institutional Brokers’ Estimate System (IBES). IBES collects data from professional security analysts on forecasts of firm earnings per share, revenues, and other variables. These data are contributed voluntarily by analysts or their employers, primarily for reputational purposes (see Ertimur, Mayhey, and Stubben ([18], 2009)). As a result, not all analysts contribute their forecasts to the database. This is still the most comprehensive database of firm-specific forecasts available. The data begin in 1975; I take data from 1990-2010.

IBES was a popular dataset for research in finance up until a scandal involving analysts retroactively changing their recommendations. Lungqvist, Malloy, and Marston ([26], 2009) downloaded the database seven times between 2000 and 2007 and found that between 1.6% and 21.7% of observations differed from one download to the next. The changes systematically made analyst performance look better, and cast doubt on the reliability of the database. The process has since been fixed to correct these discrepancies.

IBES does not contain data on all firms, but only on those that are publicly traded, have analysts that follow them, and have those analysts provide their forecasts to the database. Each of these screens tilts the available data toward large, well-known firms. In the regressions, I include dummy variables for counties that are not home to the headquarters of any of these analyst-followed-firms. Without a structural model to predict when this censoring will occur, there is no better way of dealing with this missing data.
Analysts are more interested in some variables than others. In particular, data is most readily available on estimates of earnings per share. The revenue forecasts I use are relatively sparse in the data, so another approach might be to use earnings forecasts instead of revenue forecasts so as to have more observations and more statistical power. This approach runs into a confound of cost shocks. If firms face cost shocks related to the shocks affecting their home counties (such as labor supply shocks), then using earnings surprises as a proxy for revenue surprises will include a term that is correlated with the error in the county-level regression. This would result in a biased estimate, so I cannot use the more readily available earnings forecast data.

There is some evidence that analysts forecasts are biased predictors of earnings (see So ([45], 2011) and the references contained therein). I do not know of an analogous literature concerning bias in forecasts of revenues. As I show below, revenue surprises are slightly predictable. This inserts a wedge between the idealized regressor $d_{ct}$ in the model and the actual regressor available. The predictability of revenue surprises is sufficiently small that this wedge is negligible.

IBES provides data broken down by individual analyst and in summary statistic form. I use the summary statistics, particularly the mean of all analyst forecasts. The data are available for different forecast horizons. I use the forecasts at the one and two quarter horizons. The two quarter horizon data is very sparse, but where it exists I can use it to calculate the change in expectations for month $t + 3$ revenues between quarters $t - 3$ and $t$.

Firms are identified by trading ticker and eight-digit CUSIP (Committee on Uniform Security Identification Procedures) code. There are multiple time variables available; I use the earnings announcement date. In many cases there are multiple observations associated with the same company and earnings announcement date corresponding to analyst forecasts made at different times. In these cases I use only the earliest forecast. This introduces no bias but leaves the largest residual variance, maximizing the exogenous variation in the data. I drop all observations with either an empty CUSIP, missing (or zero) mean forecast for revenues, or missing (or zero) realized revenues.

I then match the CUSIP and month to the matching CUSIP and nearest month in Compustat. With this I convert from CUSIPs to GVKEYs, which are the identifiers Compustat uses to index firms. GVKEYs have the advantage of being constant over time for a given firm, whereas CUSIPs may change. CUSIP matching is imperfect, so I lose a handful of firms in the matching process. Again there are a handful of cases in which I have repeated month-GVKEY observations, and I drop all but the first such observation.

Compustat has data on firms’ headquarters zip code\(^5\) and NAICS (North American Industrial Classification System) industry. I drop those observations where the matched line in Compustat is missing either of these data items. I then find the FIPS (Federal Information Processing Standards) code matching the county in which the zip code is located. I drop the few cases in which I cannot find a match.

\(^5\)Cohen, Coval and Malloy ([14], 2011) note that this data is backfilled, so firms that have relocated will have inaccurate historical data. They note that these moves are extremely rare. They will be even rarer in the shorter sample I use.
This process leaves a database of 129,914 observations on months, FIPS counties, NAICS industries, GVKEYs, mean revenue estimates, and revenue surprises\(^6\). For a small subset of these, I also have the update \(\Delta E\) in expectations for the next announcement period’s revenues. That is, the expectation as of month \(t\) for revenues announced in month \(t+3\), minus the expectation as of month \(t-3\) for revenues announced in month \(t+3\).

### 4.2 Connected Counties

I take data on “connected” counties from the “Contiguous County File, 1991 [United States]” made available by the Inter-University Consortium for Political and Social Research at the University of Michigan ([46], 1991). This was constructed by the Census Bureau in 1991 and lists all pairs of counties that are either 1) contiguous, 2) nearby, not contiguous, but connected by a major road, or 3) connected by “significant economic ties,” defined as “having one-way commuting flows of at least 2,000 people per day based on data from the 1980 census.”

There are 3,142 counties and 9,535 pairs of connected counties. The average county is connected to 6.07 other counties.

### 4.3 Federal Contracts

usaspending.gov provides data on every\(^7\) contract signed by an agency of the federal government and any other organization. The data run from the year 2000 through 2010 and contain an incredible level of detail. I have never before seen this dataset used in the literature. This is likely due in part to the fact that the time series is short and in part to the fact that the data totals 129 GB, making it cumbersome to work with.

The data include 27,542,324 contracts totaling 4.27 trillion dollars over ten years. This is by no means the lion’s share of federal spending. Results using this data will then be valid for overall government spending only if other types of government spending in a county are uncorrelated with contract-based spending.

Each contract in the data contains the total dollar value of the contract, the date on which it was signed, the primary NAICS industry of the provider, the zip code of the provider, and the zip code in which the work was performed. I map these zip codes into their corresponding counties. I truncate the date and consider only the month in which each contract was signed. Then for each industry \(n\), month \(t\), and county \(c\), I calculate the sum of dollar values of all contracts signed in month \(t\) with providers located in county \(c\) and operating in industry \(n\). I call this variable \(\text{spending}_{nct}\). I also calculate the sum of all contracts to be performed in each industry, month, and county. I call this variable \(\text{performance}_{nct}\).

It is important to differentiate between \(\text{spending}_{nct}\) and \(\text{performance}_{nct}\) because they may have different levels of endogeneity. The federal government may be assigning contracts to be performed in counties that are in particular need of stimulus, but may be indifferent

---

\(^6\)This is the realized value of revenues minus the expected value.

\(^7\)Presumably the data exclude classified contracts, but the Department of Defense is well-represented.
as to the home location of the contractor running the project. In this case, \( \text{performance}_{net} \) would be more endogenous\(^8\) than \( \text{spending}_{net} \). This story only applies to a subset of contracts, as many of them are transfers to state and local governments, so the location of the provider and the location of performance are the same.

There are 178,967 county-month-industry triplets with non-zero \( \text{spending}_{net} \) and 126,464 with non-zero \( \text{performance}_{net} \). I sum \( \text{spending}_{net} \) and \( \text{performance}_{net} \) across industries to form \( \text{spending}_{ct} \) and \( \text{performance}_{ct} \). There are 29,029 county-months with non-zero \( \text{spending}_{ct} \) and 19,985 county-months with non-zero \( \text{performance}_{ct} \).

Using the contiguous county file described above, I form for each county \( c \) and month \( t \) the sum of \( \text{spending}_{ct} \) across counties \( c' \) that are connected to county \( c \). I call this \( \text{spending}_{spillover ct} \). I do the same for \( \text{performance}_{ct} \) and call the resulting variable \( \text{performance}_{spillover ct} \). There are 35,077 county-months that have non-zero \( \text{spending}_{spillover ct} \) and 27,216 county-months that have non-zero \( \text{performance}_{spillover ct} \).

4.4 QCEW

Finally I use the Quarterly Census of Employment and Wages (QCEW). This dataset is made available by the Bureau of Labor Statistics and compiled using data from unemployment insurance. As such, it covers roughly 98% of employment. The available data begin in 1990, and I use data up through the end of 2010. The QCEW in its current form has been available only since 2004. As such, it has only occasionally been used in the literature\(^9\).

The data are available at many levels of disaggregation. Geographically, they are broken down nationally, by state, county, MSA, and MicroSA. They are also broken down by ownership sector (federal government, state government, local government, private) and at the industry level by supersector, sector, 3, 4, 5, and 6-digit industry. At the state and national level they are also broken down by the size of establishment class. I use sectoral (2-digit) county data broken down by ownership sector.

For each calendar quarter in the sample, each county in the country, and each 2-digit sector (these sectors are described below), the data give the employment headcount in the first, second, and third months of the quarter, the number of establishments, total wages, and average weekly wages.

There are at least two drawbacks to this data. First, it measures employment headcounts instead of hours. This means that it misses the entire intensive margin of hours. Moreover, if an individual holds two jobs, she will be counted twice in the data. Second, it excludes the self-employed. There were 21,351,320 self-employed individuals in 2008 according to the Small Business Administration\(^{10}\). This accounted for approximately one sixth of all employment. Nevertheless, the level of detail in the QCEW makes it the best dataset.

\(^8\)In this case, \( \text{spending}_{net} \) would still not be exogenous because presumably providers closer to the performance site have an advantage in the bidding process.

\(^9\)Shoag ([45], 2011), Serrato and Wingender ([41], 2011), Wilson ([47], 2011) and Feyrer and Sacerdote ([19], 2011) each use some slice of the data.

\(^{10}\)http://www.sba.gov/sites/default/files/files/data_uspdf.xls
available.

For each month \( t \), county \( c \), and industry \( n \), I find all of the companies in the filtered IBES database that are headquartered in county \( c \), operate in industry \( n \), and announce revenues in month \( t \). I then sum revenue surprises across all these firms and call this variable \( \text{surprise}_{nt} \). For those firms that have nonzero \( \Delta E \), I sum these and call this variable \( \Delta E_{nt} \).

There are 18,604,536 month-county-industry-ownership quadruples. The vast majority of these month-county-industry-ownership quadruples have no associated firms with analyst coverage. There are 66,698 ownership-county-month-industries with nonzero forecasted revenue, 66,648 with nonzero revenue surprises, and 8,334 with non-zero \( \Delta E_{nt} \). Below I sort industries into “tradable” and “non-tradable” categories. I sum \( \text{surprise}_{nt} \) and \( \Delta E_{nt} \) across tradable industries \( n \) to form \( \text{tradable_surprise}_{ct} \) and \( \text{tradable}_\Delta E_{ct} \). There are 13,039 county-months with non-zero \( \text{tradable_surprise}_{ct} \) and 2,162 with nonzero \( \text{tradable}_\Delta E_{ct} \).

For each county \( c \) and month \( t \), I sum \( \text{tradable_surprise}_{ct} \) across all those counties \( c' \) connected to county \( c \). I call this new variable \( \text{tradable_spillover_surprise}_{ct} \). I similarly sum \( \text{tradable}_\Delta E_{ct} \) to form \( \text{tradable_spillover}_\Delta E_{ct} \). This gives 65,603 county-months with non-zero \( \text{tradable_spillover_surprise}_{ct} \) and 13,787 county-months with non-zero \( \text{tradable_spillover}_\Delta E_{ct} \).

5 Identification

My identification strategy relies on the fact that the coefficients on \( \hat{d}_{ct} \) and \( \hat{g}_{ct} \) will be equal in a properly identified regression. To make use of this fact, I need data on \( \hat{d}_{ct} \). In this section I show how to take the collection of series \( \text{surprise}_{nt} \) and form the series \( \hat{d}_{ct} \), then perform several identification checks to show that these series do in fact behave like those from the model.

Suppose for the moment that I had data on \( \hat{d}_{t} + \hat{d}_{ct} \). Then I could form the series \( \hat{d}_{ct} \) by regressing \( \hat{d}_{t} + \hat{d}_{ct} \) on time dummies and taking the residuals as the idiosyncratic series \( \hat{d}_{ct} \). This observation reduces the problem of identifying \( \hat{d}_{ct} \) to the simpler problem of identifying \( \hat{d}_{t} + \hat{d}_{ct} \).

Recall that \( \hat{d}_{t} + \hat{d}_{ct} \) is the unexpected component of revenue among all businesses in county \( c \) with the following three properties:

- These businesses have sticky prices in the short run.
- These businesses conduct (significant portions of) their operations in county \( c \).
- These businesses sell to a predominantly national market.

Not all of the businesses in the IBES dataset will satisfy all of these criteria, so several of the \( \text{surprise}_{nt} \) series need to be excluded from the analysis. I classify an industry as “tradable” if businesses in the industry predominantly satisfy these three characteristics, and as “non-tradable” otherwise.

The choice of tradable and non-tradable industries is non-trivial. For instance, manufacturing is traditionally considered tradable, but many manufacturing firms have their plants in different counties than their headquarters. Including them would tend to bias estimates
towards not finding an employment effect of the revenue shock. Conversely, the industry “Other Services (Except Public Administration)” includes businesses like barbershops, which certainly do not cater to a national customer base. Table 2 summarizes my classification choices. These are largely subjective. At this level of aggregation a researcher could choose an “objective” classification rule to justify almost any choice of tradable industries; since this choice would itself be subjective, I skip that step.

Table 1: Tradable and Non-Tradable Industries

<table>
<thead>
<tr>
<th>Industry</th>
<th>NAICS Code</th>
<th>Tradable</th>
<th>Reason</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agriculture, Forestry, Fishing, and Hunting</td>
<td>11</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Mining, Quarrying, and Oil and Gas Extraction</td>
<td>21</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Utilities</td>
<td>22</td>
<td>No</td>
<td>Often serve local markets</td>
</tr>
<tr>
<td>Construction</td>
<td>23</td>
<td>No</td>
<td>Often serve local markets</td>
</tr>
<tr>
<td>Manufacturing</td>
<td>31-33</td>
<td>No</td>
<td>Plants are often not located in the same county as headquarters</td>
</tr>
<tr>
<td>Wholesale Trade</td>
<td>42</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Retail Trade</td>
<td>44-45</td>
<td>No</td>
<td>Often serve local markets</td>
</tr>
<tr>
<td>Transportation and Warehousing</td>
<td>48-49</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Information</td>
<td>51</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Finance and Insurance</td>
<td>52</td>
<td>No</td>
<td>Revenues data unreliable</td>
</tr>
<tr>
<td>Real Estate and Rental and Leasing</td>
<td>53</td>
<td>No</td>
<td>Often serve local markets</td>
</tr>
<tr>
<td>Professional, Scientific, and Technical Services</td>
<td>54</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Administrative and Support [...]</td>
<td>56</td>
<td>No</td>
<td>Often serve local markets</td>
</tr>
<tr>
<td>Educational Services</td>
<td>61</td>
<td>No</td>
<td>Often serve local markets</td>
</tr>
<tr>
<td>Health Care and Social Assistance</td>
<td>62</td>
<td>No</td>
<td>Often serve local markets</td>
</tr>
<tr>
<td>Arts, Entertainment, and Recreation</td>
<td>71</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Accommodation and Food Services</td>
<td>72</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Other Services (Except Public Administration)</td>
<td>81</td>
<td>No</td>
<td>Often serve local markets</td>
</tr>
</tbody>
</table>

I sum $\text{surprise}_{\text{net}}$ over the tradable industries and call this new variable $\text{tradable\_surprise}_{ct}$. I similarly form $\text{non\_tradable\_surprise}_{ct}$, $\text{tradable\_spillover\_surprise}_{ct}$, $\text{non\_tradable\_spillover\_surprise}_{ct}$, $\text{tradable\_employment}_{ct}$ and $\text{non\_tradable\_employment}_{ct}$. Before considering the impact of a tradable revenue surprise on county outcomes, I perform several checks to ensure that the tradable surprise does in fact correspond to $d_t + d_{ct}$.

5.1 Censoring

As noted above, many counties have no firms with analyst coverage. As a result, they have zeros for all of their surprise variables. In many cases there are some businesses located in these counties, but the lack of analyst coverage means their data are censored. The censoring is non-random, but also does not conform to a simple model like a Tobit. To control for this, I include a dummy variables for censoring of the $\text{tradable\_surprise}_{ct}$ and $\text{tradable\_spillover\_surprise}_{ct}$ series in all the regressions in this section.

In the next section I exclude these dummies because they are by their nature uncorrelated with the surprise variable that is the regressor of interest. Once I verify in this section that the regressor has the desired properties, that renders the censoring dummies extraneous.

5.2 Correlations of the Shocks Series

The shocks $d_{ct}$ in the model are consumer preference shocks to the goods and services of individual firms. Part of the identifying assumption is that they are unrelated to local
conditions in the county where these firms are headquartered. This leads to the prediction that the surprise series for different industries should be uncorrelated.

I therefore compute the correlation matrix of the surprise series across each of the NAICS industries. I do not report the entire matrix here, but note that the largest absolute values produced are .0922, -.0685, and .0599. These are statistically significant under a Spearman Rank test, but small enough that I consider this evidence consistent with the hypothesis that these shocks are idiosyncratic demand shocks.

5.3 Own-Employment Response

Consider a firm satisfying all three criteria for tradability and suppose its demand curve is shifted outward by an exogenous shock. With sticky prices in the short run, the firm will have to increase production, which requires increasing inputs. This gives a prediction: tradable industries should have a positive employment response to their own revenue shocks.

Consider the same business, but suppose a competitor in a neighboring county receives the positive shock. The competitor increases its labor demand. This has two effects. First, the increase in labor demand increases the wealth of agents in the area, increasing their spending on local products. Second, the increase in labor demand puts positive pressure on wages in the industry in the local labor market. A tradable firm will only be affected by the second of these effects, because the first one will only benefit non-tradable industries in the area. This gives a second prediction: tradable industries should have a negative employment response to revenue shocks in the same industry in neighboring counties.

I test these two hypotheses. If tradable_surprise_{ct} does in fact correspond to \( \hat{d}_t + \hat{d}_{ct} \), then a regression of the change of tradableEmployment_{ct} on tradable_surprise_{ct} and tradable_spillover_surprise_{ct} should give a positive coefficient on the former and a negative coefficient on the latter. I run this regression on the change of employment from time \( t - 12 \) to \( t + 6 \), along with month, year, and censoring dummies. The coefficients produced are 1.335493 (\( t = 3.98 \)) and .1515785 (\( t = 2.83 \)), respectively.

This is consistent with the prediction that a revenue shock to a county’s tradable industries will increase employment in the effected companies. It is inconsistent with the prediction that labor supply effects would decrease employment in tradable industries in neighboring counties. There are at least three possible explanations for this last piece of evidence. First, it may simply be that the identification is incorrect. More likely, the firms receiving revenue surprises conduct some of their employment in neighboring counties, so these are the firms increasing tradable employment in these neighboring counties. Finally, there could be some super-Keynesian effect in which the increase in employment in one county invigorates the local labor market in a way that makes the underlying matching function more efficient for all firms.
5.4 Shocks Are Uncorrelated With Government Spending

The next requirement for identification of $\hat{d}_{ct}$ is that it be uncorrelated with $\hat{g}_{ct}$. I do not have a measure of $\hat{g}_{ct}$, but I can consider the change in $g_{ct}$. I regress the twelve-month change in government spending $spending_{ct}$ in a county on $tradable\_surprise_{ct}$, month and year dummies, and censoring dummies.

The endogeneity concern is that government spending will respond negatively to the revenue surprise variable, that it will be a locally counter-cyclical force. In fact the realized coefficient is positive and insignificant (4955.294, $t = 1.37$). This is further evidence that I have identified the desired set of shocks.

5.5 Are Surprises Really Surprises?

The final characteristic of revenue surprises is that they are actually surprises. If analysts have fully rational expectations and accurately report them, then revenue surprises should be unpredictable econometrically. To check this, I regress $tradable\_surprise_{ct}$ and $tradable\_spillover\_surprise_{ct}$ on the first three lags of each of these, along with the first three lags of the non-tradable surprises.

Nine out of twelve of the coefficients in the $tradable\_surprise_{ct}$ regression are insignificant, and the overall $R^2$ is .1602. Seven of twelve are significant in the $tradable\_spillover\_surprise_{ct}$ regression, and the overall $R^2$ is .2127. In each case these represent non-trivial degrees of predictability.

Below I report the results of regressions using the raw surprises. In unreported results I take the residuals from these two regressions and use them as the surprise series for the multiplier regressions. The results are largely similar to the results using the raw data.

5.6 Persistence of Revenue Shocks

The optimal dynamic response of a firm to a shift in its demand curve depends on the expected persistence of that shift. A permanent increase in demand may justify increases in capacity that are not justified by a temporary increase. All of the endogenous responses of the county economy will then depend on the expected persistence of the revenue shock.

The stimulus policies of interest are generally temporary demand increases. If the anonymous trade assumption of the model applies literally, then the persistence of revenue surprises is unimportant. But in in a world in which firms do know who their customers are, the endogenous response of firms to their revenue shocks will only be invariant to their source if the persistence of those shocks is the same regardless of where they originate from. It is therefore important to test the persistence of the identified revenue surprises.

Consider a firm for which data is available on revenue forecasts at horizons of one and two quarters. Suppose this firm receives a positive revenue surprise in quarter $q$. If this demand shock is persistent, this will increase in a positive revision in expectations for revenues in quarter $q + 1$. That is, there should be a positive relationship between $R_q - E_{q-1}[R_q]$ and $E_q[R_{q+1}] - E_{q-1}[R_{q+1}]$. 


For those counties for which I have at least one tradable firm with data on $E_{q-1}[R_{q+1}]$, I regress $E_q[Tradable\_Revenue_{c,q}, q + 1] - E_{q-1}[Tradable\_Revenue_{c,q}, q + 1]$ on $Tradable\_Surprise_{ct}$. This gives a coefficient of $0.0038$ ($t = 3.18$) and an $R^2$ of $0.0006$. This persistence is statistically significant but economically negligible. It is therefore reasonable to consider these revenue surprises as approximately transient, like traditional fiscal stimulus.

6 Results

Firms located in county $c$ have information about their revenues before these revenues are announced to the market. I therefore look at the employment response to a revenue shock starting eleven months before the shock is announced, and continue up through sixty months after the shock is announced. Examining too short a horizon would lead to underestimating the effect of the shocks, as the shocks still have measurable effect after five and a half years. On the other hand, as the horizon grows longer, more periods at the end of the sample are dropped and statistical power decreases. I chose sixty months as a middle-ground.

Figure 1 plots the estimated employment effect on county $c$ of a million dollar shock to the revenues of tradable firms located in county $c$ when that shock is announced in period $0$. 95% confidence intervals are plotted as dotted lines. All standard errors are clustered at the county level.

The pattern of jobs creation is hump-shaped, starting at about third of a job in the -11 month and growing to 3.77 jobs in month 7. The number of jobs created then declines to 1.12 by month 24, only to grow again and hit another trough in month 49. This is suggestive of a two-year cycle, or may just be statistical noise. It will be interesting to see if this pattern remains robust as more data become available.

The proper unit for assessing the effect of this shock is not the number of jobs created at any single horizon, but the total number of jobs weighted by how long they last. That is equal to the area under the graph in Figure 1 divided by twelve, which is 13.4976 job-years,
corresponding to a cost per job-year of $74,087. This is a lower bound on the number of job-years and upper bound on cost per job-year created for three reasons. First, it does not account for jobs created at horizons longer than 60 months out. Second, it does not account for spillovers to neighboring counties. Third, it does not account for effects from shocked firms hiring in faraway counties in response to the shock. I examine spillovers to neighboring counties below.

6.1 Recessions

There is a recent literature arguing that multipliers are higher in recessions than they are in normal times. To test this hypothesis, I run the same regressions as above using an interaction of the shock with an indicator for a recession. These results are plotted in Figure 2.

The initial response in the county is much larger in a recession than in the overall sample, growing to 9.53 jobs by month 5. This job creation drops off quickly eighteen months after the shock. Up through the eighteen month mark, the shock generates 12.3561 job-years, roughly the same amount as is created over the entire horizon in the whole sample. But the drop off in jobs continues and goes as low as -9.01 by month 27. The net gain in job-years over the entire horizon is -3.1294.

Had I reported only the first portion of the results, this would seem like strong evidence in favor of the hypothesis that the multiplier is higher in recessions than it is in normal times. The strong negative jobs effect in the latter portion of the results pose a challenge for this view. There are at least three reasons why the latter portion of the results should be taken with a grain of salt. First, the confidence errors are much wider there. Second, the data based on the 2007 is no longer taken into account after about the two year horizon, so the entire latter portion of the results is being driven by the 1991 and 2001 recessions. Third, because recessions generally have shorter horizons than 60 months, the end part of the results is an application of a recession multiplier in a non-recession regime.
Looking at only the 18 month horizon, the cost per job-year in a recession is $80,932. This is larger than the $74,087 above, but much smaller than the $176,055 I would have calculated above had I restricted to the shorter horizon.

6.2 Spillovers Across Counties

Counties are fairly small geographic units, and people frequently cross county lines either to go to their jobs or to consume. An increase in labor demand in one county will therefore have two effects on neighboring counties. First, some of the workers living in neighboring county \( c' \) will start to commute across county lines, lowering labor supply in county \( c' \) and decreasing job growth in county \( c' \). Second, the newly employed workers in both counties \( c \) and \( c' \) will spend a portion of their new income in county \( c' \). This drives up the demand curves for non-tradable firms located in county \( c' \) and then increases labor demand in the county. Which of these two effects dominates is an empirical question.

I run the same set of regressions as above, adding as a regressor the total tradable revenue shock summed across counties contiguous with county \( c \). Figure 3 plots the estimated coefficients on the tradable_spillover_surprise\(_ {ct} \) over the -11 to 60 month horizon.

![Graph showing employment response to million dollar shock in connected county](image)

The initial response of employment to a shock in a connected county is insignificant, consistent with the narrative that firms in the shocked county hire and these shocks slowly diffuse. After the diffusion kicks in, there is a single hump-shaped response climbing to four tenths by month 8, then staying flat for the next three years and slowly declining. The decline is associated with a loss in precision as the sample becomes smaller. This evidence suggests that the consumption spillovers dominate the labor supply effects. The shock to the neighboring county lifts all boats.

The net effect on this neighboring county is 1.1879 job-years in response to a million dollar revenue shock. Each county \( c \) is connected to an average of 6.07 other counties, so the total employment change due to a million dollar revenue surprise in county \( c \) is \( 13.4976 + 6.07 \times 1.1879 = 20.7201 \) job-years. This remains an underestimate because of the additional spillovers when firms headquartered in county \( c \) hire directly in unconnected counties.
counties $c'$. This corresponds to a cost per job-year of $48,262$, which I view as an upper bound\(^{12}\). This evidence is broadly in line with the estimates of Shoag ([45], 2011), Chodorow-Reich et al ([10], 2011), and Serrato and Wingender ([41], 2011).

I can do the same county-spillover analysis including an interaction term for a recession. Figure 4 plots the estimates response of employment to a million dollar revenue shock in a connected county in a recession.

![Figure 4](image)

These results mirror the previous recession results. The employment response is hump-shaped and positive for about 18 months, then drops below zero and stays there. The total number of jobs-years created over the 60 and 18 month horizons are $-3.3036$ and $0.2171$, respectively. As I argued above, the latter is the relevant horizon for considering the response in a recession. Adding these jobs to the jobs created in the home county, the total jobs created in response to a million dollar revenue surprise in a recession is $12.3561 + 6.07 \times 0.2171 = 13.6739$, corresponding to a cost per job-year of $73,132$. The analogous number over the whole sample is $110,838$, again suggesting that the multiplier is higher in a recession than in normal times.

### 6.3 Spillovers Across Industries

To better understand the mechanism through which jobs are created in response to these revenue shocks, I look at job growth within non-tradable industries. By construction these jobs have to be created through some sort of spillover channels. Figure 5 shows the number of jobs created in non-tradable industries in response to a shock to tradable industries.

\(^{12}\)The inversion problem makes calculation of standard errors for this estimate cumbersome.
At its peak, a million dollar shock to a tradable industry in county $c$ produces about one and a half extra jobs in non-tradable industries in the county, about half the total effect. This is indicative of strong spillover channels. One hypothesis is that newly employed workers in tradable industries have a high marginal propensity to consume which increases demand for non-tradable firms located in the county. Testing this hypothesis is a topic for future research.

6.4 Comparison with OLS

It is well-known that the OLS estimate of the effects of government spending on economic outcomes is biased. This mitigates but does not eliminate the usefulness of this estimate as a baseline against which to compare other estimates. An econometrician unaware of the correlation between $\hat{g}_{ct}$ and $\epsilon_{ct}$ would run the panel regression:

$$\Delta Emp_{ct} + \tau,t_{-12} = a_0 + a_c + a_t + b_t g_{ct}$$ (12)

I run these regressions for $\tau = -11, ..., 60$. The estimated coefficients $b_t$ are graphed in Figure 6. I also include as a regressor the total government spending in neighboring counties to assess spillovers. Instead of time dummies I use month and year dummies. The results are almost identical with time dummies.
Two things are striking about these graphs. First, a million dollars spending contract is associated with a strong negative response of employment, totaling -29.0415 job-years over the full horizon. For a performance contract, the associated number is -79.9747 job-years. Second, there is strong pattern at the 5-6 month frequency. I have no compelling hypotheses for what is causing this.

Again to consider the possibility that multipliers may be higher during recessions than during normal times, I run the same regression with an added interaction term for the presence of a recession:

\[
\Delta \text{Emp}_{ct+r,t-12} = a_0 + a_r + a_t + b_r g_{ct} + c_r 1_{\text{Recession},g_{ct}}
\]  

These coefficients \(b_r + c_r\) are plotted in Figures 8 and 9 for \(spending_{ct}\) and \(performance_{ct}\), respectively.
The naive estimate suggests that spending in a recession makes things much worse, and again is likely due to the bias introduced by ignoring the correlation between \( \hat{g}_{ct} \) and \( \epsilon_{ct} \). The net increase in jobs associated with a million dollars spending contract in a recession is -52.7515 job-years over the full horizon. For a performance contract the associated number is -331.5022 job-years. This strongly suggests that government spending is being targeted effectively at regions that need it most.

The estimated coefficients drop off sharply after 24 months. This is likely driven by the fact that after the 24 month horizon, the 2009-2010 period of the sample falls out of the regressions.

### 6.5 Robustness Checks

In unreported results, I perform the same analysis with the following sets of small changes to the regressors:
• Used time fixed effects instead of year and month dummies
• De-meaned firm-level surprises at the industry level for each time period
• Included county-level unemployment, both as a regressor and interacted with the surprises
• Regressed the surprises on their lagged values and used the residuals as the surprise series

In each of these cases, the estimated dynamic impacts are similar to those using the raw data.

7 Conclusion

We can think of the multiplier on government spending as the sum of two terms. The first term is a Keynesian term. Increased spending causes firms to hire more workers, those workers to go out and consume more, and a virtuous circle to ensue. The second term is a Ricardian term. Increased spending increases the discounted present value of future taxes, causing consumers to cut back on consumption in expectation of having to pay those taxes. Identifying the net effect has proven difficult because of a lack of exogenous variation in government spending.

Looking at the effect of local government spending gives us extra traction. An extra dollar of federal government spending at a local level increases demand by a dollar, but increases the discounted present value of taxes by a negligible fraction of a dollar. The effect on the local economy is given only by the Keynesian term. Splitting the problem up into smaller pieces gives us a better chance of identifying the multiplier of interest. But we are again plagued by a lack of exogenous variation. Spending at a local level can be driven by local conditions, so OLS gives biased results for the Keynesian term of the multiplier.

The observation that motivates this paper is that a business which sees a shift in its demand curve does not know or necessarily care what caused that shift. As a result, the local effects of that shift do not depend on its source. This observation implies that we can estimate the effect of local government spending surprises using non-government spending surprises. This is an ideal regressor because it is known to be unexpected and is not caused by county conditions. This lets us get around the usual endogeneity problem with estimating the causal effects of government spending.

The next question which naturally arises is whether spillover effects cause the global multiplier to be larger than the local multiplier. This question has never answered in a general theoretical framework. The answer is intuitive: the Keynesian term does get amplified by spillover effects, but so does the Ricardian term. It is untrue that the local multiplier is always a lower bound for the global multiplier.

It is worth noting that while the multiplier equality result holds whenever trade is anonymous, the multiplier need not be policy invariant. Consider an example in which local government spending shocks are purely transient while consumer spending shocks are persistent. Then when a firm receives a demand shock, its optimal response depends on the likelihood
that the shock came from government spending as opposed to consumer spending. Roughly speaking, increasing the variance of the government spending shock increases the likelihood that the demand shock came from the government. This pushes the firm’s optimal behavior toward the behavior they would take if they knew the demand shock came from the government. This changes the estimated multiplier $F_3$. In this way, the multiplier is invariant to where the spending comes from, but not invariant to government spending policy. I think of this as a kind of Lucas caveat.

7.1 Suggestions for Future Research

The usaspending.gov dataset has a tremendous level of detail available; my use of it only began to scratch the surface. As the sample available grows longer, researchers will be able to analyze in detail the effects of different types of government spending. For instance, researchers can find which contracts are grants to state and local governments, so could test the hypothesis that these grants prevent layoffs at lower levels of government. Each contract is also matched with the department of the federal government that issued the contract, so researchers could test the hypothesis that the Department of Defense has more stimulative impact than the Department of Energy.

My classification of industries as tradable and non-tradable suffers from several drawbacks. Most importantly, it was a subjective choice based on an assessment of whether firms in these industries both produce locally and sell nationally. Second, I only used the two-digit level of NAICS industries because of computing capacity constraints. Both of these could be improved by future researchers. An industrious researcher could comb through the IBES data and directly identify firms that satisfy the exclusion restrictions, getting a significantly better measure of tradable revenue surprises. This would improve both the statistical performance and the plausibility of the results.

As noted above, the QCEW does not capture self-employment or the hours margin. If data becomes available on high-frequency movements in either of these, future research could significantly refine my estimates.

My regressions did not include any county-level controls. Finding a set of controls which predict employment growth in ways orthogonal to the identified revenue shocks would reduce residual variance and increase the precision of the estimates. The social marginal value of having a more precise estimate of the multiplier is enormous.

References


[9] Lawrence Christiano, Martin Eichenbaum, and Sergio Rebelo, 2011. When is the Government Spending Multiplier Large?


[29] Enrico Moretti, 2010. Local Multipliers. AER.


[40] Claudia R. Sahm, Matthew D. Shapiro, and Joel Slemrod, 2010. Check in the Mail or More in the Paycheck: Does the Effectiveness of Fiscal Stimulus Depend on How it is Delivered?


