Estimating Local Fiscal Multipliers

Juan Carlos Suárez Serrato*
Department of Economics
University of California, Berkeley

Philippe Wingender
International Monetary Fund

November 2011

JOB MARKET PAPER # 2
(See also JMP # 1)

Abstract

We propose a new identification strategy to measure the causal impact of government spending on the economy. Our methodology isolates exogenous cross-sectional variation in government spending using a novel instrument. We use the fact that a large number of federal spending programs depend on local population levels. Every ten years, the Census provides a count of local populations. Since a different method is used to estimate non-Census year populations, the discontinuous change in methodology leads to variation in the allocation of billions of dollars in federal spending. Our IV estimates imply that government spending has a local income multiplier of 1.88 and an estimated cost per job of $30,000 per year. We also show that the local effects are not larger than aggregate effects at the MSA and state levels. Finally, we characterize the heterogeneity of the impacts of government spending and find that it has a higher impact in low growth areas.

Keywords: Government spending, fiscal multiplier, instrumental variables.

*Corresponding author: jcsuarez@berkeley.edu. We are very grateful for guidance and support from our advisors Alan Auerbach, Patrick Kline and Emmanuel Saez. We are also indebted to Daron Acemoglu, David Albouy, Charles Becker, David Card, Raj Chetty, Gabriel Chodorow-Reich, Colleen Donovan, Daniel Egel, Fred Finan, Charles Gibbons, Yuriy Gorodnichenko, Ashley Hodgson, Shachar Kariv, Yolanda Kodrzycki, Zach Liscow, Day Manoli, Steve Raphael, Ricardo Reis, David Romer, Jesse Rothstein, John Karl Scholz, Dean Scrimgeour, Daniel Wilson, numerous seminar and conference participants, and four anonymous referees for comments and suggestions. All errors remain our own. We are grateful for financial support from the Center for Equitable Growth, the Robert D. Burch Center for Tax Policy and Public Finance, IGERT, IBER and the John Carter Endowment at UC Berkeley. The latest version of this paper can always be found at `http://www.jcsuarez.com/`.

1 Introduction

The impact of government spending on the economy is currently the object of a critical policy debate. In the midst of the worst recession since the 1930s, the federal government passed the American Recovery and Reinvestment Act (ARRA) in February 2009 at a cost of more than $780 billion in the hopes of stimulating a faltering US economy. The bill contained more than $500 billion in direct federal spending with a stated objective to "... save or create at least 3 million jobs by the end of 2010" (Romer and Bernstein, 2009). Despite the importance of this debate, economists disagree on the effectiveness of government spending at stimulating the economy. The endogeneity of government spending makes it difficult to draw a causal interpretation from empirical evidence. We contribute to this important discussion by proposing a new empirical strategy to identify the impacts of government spending on income and employment growth.

In this paper we propose a new instrumental variable that isolates exogenous variation in government spending at the local level. We use the fact that a large number of direct federal spending and transfer programs to local areas depend on population estimates. These estimates exhibit large variation during Census years due to a change in the method used to estimate local population levels. Whereas the decennial Census relies on a physical count, the annual population estimates use administrative data to measure incremental changes in population. The difference between the Census counts and the concurrent population estimates therefore contains measurement error that accumulated over the previous decade. We use the population revisions which occurred following the 1980, 1990 and 2000 Censuses to estimate the effect of an exogenous change in federal spending across counties.[1] While we use this identification strategy to estimate fiscal multipliers, one of the contributions of this study is the careful documentation of an instrument that can be used to analyze the impact of government spending on other outcomes as well.

In a first step, we document a strong statistical relationship between changes in

[1]Similar identifications strategies can be found in the literature. Gordon (2004) uses the changes in local poverty estimates following the release of the 1990 Census counts to study the flypaper effect in the context of Title I transfers to school districts. In contrast to Gordon (2004), our identifying variation emanates from measurement error rather than from a decadal discontinuity. In a paper looking at political representation in India, Pande (2003) uses the difference between annual changes in minorities' population shares and their fixed statutory shares as determined by the previous Census.

population levels due to Census revisions and subsequent federal spending at the county level. This is consistent with the fact that a large number of federal spending programs use local population levels to allocate spending across areas. This dependence operates either through formula-based grants using population as an input or through eligibility thresholds in transfers to individuals and families.[2] We also document the fact that it takes several years for different agencies in the federal government to update the population levels used for determining spending. Thus, even though the instrument we propose occurs once every decade, it provides many years of exogenous variation in federal spending. The fact that our empirical results are consistent with the timing of the release of Census counts provides a very strong test for the validity of our identification strategy.

We use the exogenous variation in federal spending identified by our instrument to measure the causal impact of spending on economic outcomes at the local level. We find an estimate of the income multiplier, the change in aggregate income produced by a one dollar change in government spending, of 1.88 and a estimated cost per job created of $30,000 per year. The IV results imply a return to government spending that is ten times larger than the corresponding OLS estimates. This shows that not accounting for the endogeneity of federal spending leads to a large downward bias as we strongly reject the equality of the OLS and IV coefficients in all our main regression results. This highlights the obvious concerns for endogeneity and reverse causality between government spending and local economic outcomes. A number of robustness checks also show that our estimates are not confounded by known predictors of population changes such as local demand shocks.

The difficulty of finding a valid instrument for federal spending at the local level could explain why cross-sectional variation has not been used more extensively in the empirical literature thus far.[3] An OLS approach even using fixed effects to control for time-invariant local characteristics will typically yield biased estimates. For example, some categories of government spending are automatic stabilizers so that spending increases when the local economy experiences a slowdown. An OLS approach would

[2]A review by the Government Accountability Office (GAO 1990) in 1990 found 100 programs that used population levels to apportion federal spending at the state and local level. Blumerman and Vidal (2009) found 140 programs for fiscal year 2007 that accounted for over $440 billion in federal spending; over 15% of total federal outlays for that year.

[3]Recent examples addressing the endogeneity of government spending include Busso et al. (2010), Clemens and Miran (2010), Chodorow-Reich et al. (2010), Fishback and Kachanovskaya (2010), Shoag (2010) and Wilson (2010).

thus produce downward-biased estimates. The comparison of our OLS and IV results suggest this is the case.[4]

Since our main results are at the county level, we replicate our estimation methodology at the metropolitan statistical area (MSA) and state levels of aggregation. It is not clear a priori how the local multiplier relates to its national counterpart. Positive spillovers across counties would lead us to underestimate the national multiplier. On the other hand, if government spending is crowding out private demand for labor and this effect is operating differently in the recipient and neighboring counties, our estimates at the local level could be overestimating the total impact of government spending. We find that our estimates of the return to government spending do not decrease as a result of aggregation.[5]

Our estimation strategy differs from many papers in the empirical macroeconomics literature in that we rely on cross-sectional instead of time-series variation to measure the causal impact of government spending on the economy (e.g. Ramey and Shapiro, 1997, Fatás and Mihov, 2001, Blanchard and Perotti, 2002, Ramey, 2010). This approach has many advantages. Foremost, it allows us to clearly identify the source of exogenous variation in government spending. Exploiting cross-sectional variation also allows for research designs with potentially much larger sample sizes. This can increase statistical power and the precision of our estimates. We show that a cross-sectional approach is particularly amenable to the study of the effects of government spending on local outcomes and can yield new results. In particular, we characterize the heterogeneity in the impact of government spending using a new method that uses instrumental variables in a quantile regression framework (Chernozhukov and Hansen 2008). We show that government spending decreases income growth inequality across counties.

One further difference with time-series analysis is that nation-wide effects of policy changes cannot be identified in cross-sectional regressions.[6] One candidate for such a general equilibrium effect is the additional tax burden for individuals in all regions that comes from the increase in spending in a single area. If, for example, forward-

[4]On the other hand, OLS estimates could be upward-biased if infrastructure spending was targeted to counties with high complementarity between public and private capital.

[5]Davis et al. (1997) find positive spillovers of demand shocks across states. Glaeser et al. (2003) develop a model in which the presence of positive spillovers leads to larger social multipliers than those implied by lower level estimates.

[6]See Acemoglu, Finkelstein and Notowidigdo (2009) for a discussion in the context of health spending and local area income.

looking agents decrease consumption and investment as a result of higher expected future taxes, this behavioral response would go undetected by our empirical analysis. However, since we rely on the redistribution of federal spending across local areas and not on absolute changes in the level of spending, our natural experiment might not induce this Ricardian-type response. Another national general equilibrium effect is the impact of the monetary policy response. Nakamura and Steinsson (2010) show that the cross-sectional estimate of the fiscal multiplier coincides with the national multiplier when nominal interest rates are unresponsive to a fiscal expansion such as when they are constrained by the zero-lower bound.

The following section provides background into the source of variation in population levels. Section 3 describes the data used in the study. Section 4 provides a framework for thinking about the conditions for identification in the context of our natural experiment. Section 5 discusses the implementation of the empirical strategy and characterizes the variation in the instrument. Sections 6 and 7 present the first stage and instrumental variables results, respectively. Section 7 also compares the IV and OLS results and conducts several robustness checks while Section 8 relates the local multipliers with estimates at the MSA and state levels. Section 9 analyzes heterogeneity in the impact of government spending and Section 10 concludes.

2 Population Levels and Government Spending

As mandated by the Constitution, the federal government conducts a census of the population every ten years. These population counts are used to allocate billions of dollars in federal spending at the state and local levels. The increased reliance on population figures has also led to the development of annual estimates that provide a more accurate and timely picture of the geographical distribution of the population. Due to the prohibitive cost of conducting a physical count every year, the US Census Bureau developed alternative methods for estimating local population levels. For the last thirty years, it has relied on administrative data sources to track the components of population changes from year to year. These components are broadly defined as natural growth from births and deaths as well as internal and international migration.[7]

A crucial feature of these estimates is that they are "reset" to Census counts once these data become available. This revision process leads to a break in population

[7]See Long (1993) for details.

trends at all levels of geography. The difference between the two population measures in Census years is called "error of closure." The Census Bureau's objective is obviously to produce population estimates that are consistent over time. However, the use of two different methods for producing population figures necessarily leads to some discrepancy due to measurement errors in both the annual estimates and the physical Census counts.[8]

The error of closure has been substantial in the past three Censuses. In 1980, the Census counted 5 million more people than the concurrent population estimate. The 1990 Census counted 1.5 million fewer people than the national estimate. This was apparently due to systematic undercounting of certain demographic groups. In 2000, the Census counted 6.8 million more people than the estimated population level.[9] These errors of closure are relatively more important at the local level due to the difficulty of tracking internal migration. In Figure 1 we show the average county population growth rate across all counties by year. The series shows clear breaks in 1980, 1990 and 2000. We also show in Figure 2 the full distribution of county population growth rates for 1999 and 2000 separately. The figure clearly shows that the Census revisions affect the whole distribution of growth rates: the variance is also larger as more counties experience very high positive and negative growth in 2000 than in 1999.

Local population levels are used in the allocation of federal funds mainly through formula grants that use population as an input and through eligibility thresholds for direct payments to individuals (e.g. Blumerman and Vidal 2009, GAO 1987). Federal agencies use annual population estimates or Census counts depending on the availability and timeliness of the latter. The release of new Census counts therefore creates a discontinuity in population levels used for allocating spending that we exploit in our empirical design. However, this change does not occur in the year of the Census since it usually takes two years for the Census Bureau to release the final population reports.[10] The specific timing of the release of the final Census counts allows for a powerful test of our identification strategy as the Census shock should be uncorrelated with federal spending before the release of the final Census counts.

[8] A large literature acknowledges the measurement errors in the physical Census counts. The statistical adjustment of the physical count has also been the subject of a sharp political debate for many decades. See, for example, West and Fein (1990).

[9] See Census Bureau (2010a).

[10] See Census Bureau (2010b,c) and Census Bureau (2001).

Figure 1: Average County Population Growth Rate by Year

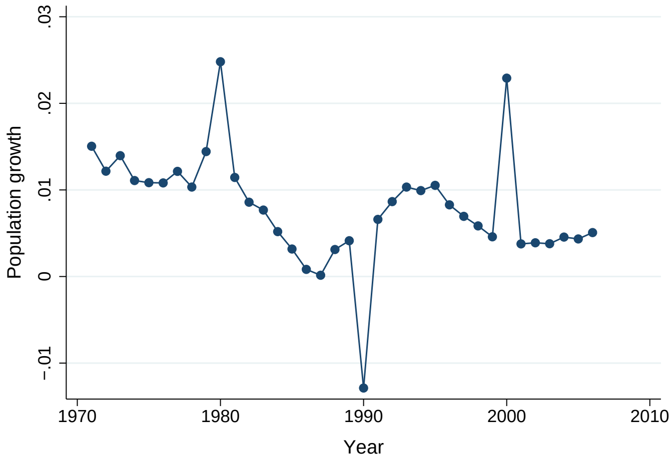
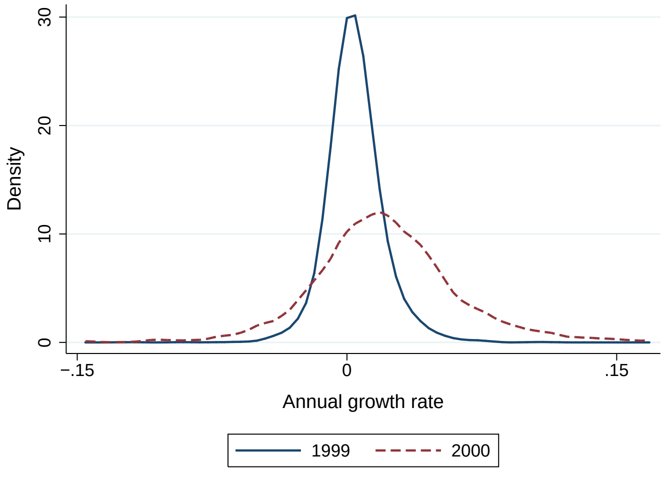
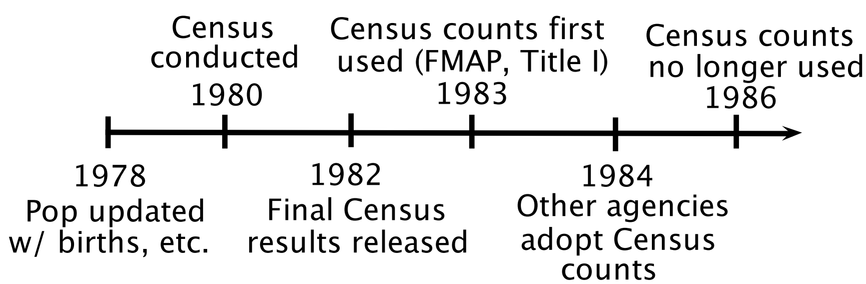


Figure 2: Distribution of County Population Growth Rates 1999-2000



Federal agencies have some discretion in updating the population levels used to allocate spending. Variation in the year of adoption of Census counts across agencies suggests that the Census shock influences federal spending several years after the release of the final counts. One example is the Federal Medical Assistance Percentage (FMAP) used for Medicaid and Temporary Assistance for Needy Families (TANF) transfers to states. This percentage is a function of a three year moving average of the ratio of states' personal income per capita to the national personal income per capita.[11] The three year moving average is also lagged three years so that the 2009 FMAP, the last year in our dataset, relies on population estimates dating back to 2004 (Congressional Research Service, 2008). We therefore would not expect the Census population shock to affect FMAP spending until three years after the Census is conducted. The moving average used in the FMAP implies that the population revision will be correlated with changes in the FMAP up until five years after the Census year. We illustrate a simplified timeline for the 1980 Census in Figure 3.

Figure 3: Timeline



3 Data

Counties are a natural starting point for our analysis because of their large number and stable boundaries for the period under study. There are over 3000 counties when excluding Hawaii and Alaska, which we do throughout the analysis. We use contemporaneous county population estimates published by the Census Bureau from 1970 to 2009. These are called postcensal estimates.[12] There were no postcensal estimates

[11] Per capita income depends on population estimates only through the denominator. See the Data section for further details.

[12] The Census Bureau also releases intercensal estimates, which are revised after new Census counts are available. See Census Bureau (2010a) for details on the revision procedure.

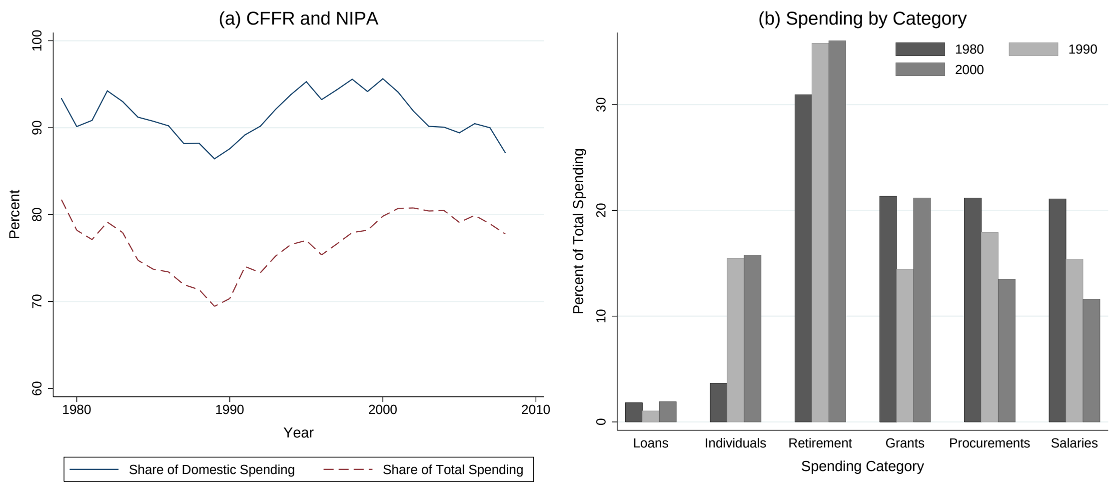
released in 1979, 1980, 1989, 1990 and 2000 because of the upcoming Censuses. Since our empirical strategy requires the comparison of estimated population levels and Census counts, we produce these estimates for the five missing years using publicly available data in an attempt to replicate the Census Bureau's methodology. This methodology involves tracking population changes using administrative data. Natural growth in population is estimated using data on births and deaths while migration is estimated using data on tax returns, Medicare, school enrollment, and automobile registration.[13] We use annual county-level births and deaths from the Vital Statistics of the U.S. to generate our own estimates of county natural growth. Data used to estimate internal and international migration are from the County-to-County Migration Data Files form the IRS's Statistics of Income.

Data on federal spending come from the Consolidated Federal Funds Reports (CFFR) published annually by the Census Bureau.[14] This dataset contains detailed information on the geographic distribution of federal spending down to the city level. In cases where federal transfers are passed through state governments, the CFFR estimates the sub-state allocation by city and county. Spending is also disaggregated by agency (from 129 agencies in 1980 to 680 in 2009) and by spending program (from 800 programs in 1980 to over 1500 in 2009). The specific programs are classified into nine broad categories based on purpose and type of recipient. We restrict our analysis to the following categories: *Direct Payments to Individuals* (which includes Medicare payments), *Direct Payments for Retirement and Disability*, *Grants* (Medicaid transfers to states, Highway Planning and Construction, Social Services Block Grants, etc.), *Procurement and Contracts* (both Defense and non-Defense), *Salaries and Wages* of federal employees and *Direct Loans*. Given the high variance of spending across years at the county level and the fact that some of the data represent obligations that are often subsequently revised, we use a three year moving average of total spending in these categories. We exclude *Direct Payments Other than for Individuals* which consist mainly of insurance payments such as crop and natural disaster insurance. We exclude these types of spending as they are not relevant in the context of our natural experiment and decrease the statistical power of our first stage. Finally, we exclude the *Insurance* and *Guaranteed Loans* categories since they

[13] See Long (1993) for details.

[14] The CFFR was first published by the Census Bureau in 1983. Predecessors to the CFFR are the Federal Outlays series from 1968 to 1980 and the Geographic Distribution of Federal Funds in 1981 and 1982.

Figure 4: Federal Spending in the CFFR



represent contingent liabilities and not actual spending. Panel (a) in figure 4 shows how our measure of federal spending compares to federal spending in the National Accounts. On average, we capture between 70 and 80% of total spending and over 90% of total domestic spending (total spending minus debt servicing and international payments). Panel (b) breaks down total federal spending by the broad categories used in the analysis for the three Census years.

Data on county personal income, salaries and wages and employment are taken from the Bureau of Economic Analysis' Regional Economic Information System (REIS). This data is compiled from a variety of administrative sources. Employment and earnings come from the Quarterly Census of Employment and Wages (QCEW) produced by the Bureau of Labor Statistics. The QCEW contains the universe of jobs covered by state unemployment insurance systems and accounts for more than 94% of total wages reported by the BEA. Personal income which also includes proprietors' and capital income, transfer receipts and supplements to salaries and wages uses IRS, Social Security Administration and state unemployment agencies data among other sources. An important feature of these data is that they do not depend on the discontinuity in population estimates that is the basis of our instrument (BEA 2010).

Finally, we also extract several county characteristics from the 1970, 1980, 1990 and 2000 Censuses and we express all dollar values in dollars of 2009 using the national Consumer Price Index published by the BLS.

4 Identification Strategy

This paper uses an instrumental variables strategy to estimate the impacts of government spending on the local economy. Taking advantage of cross-sectional identifying variation, our estimates circumvent endogeneity concerns that can bias an OLS approach. The identifying conditions for our strategy are the usual inclusion and exclusion restrictions of the IV framework. In Section 6 we show that our instrument satisfies the inclusion restriction by demonstrating that it is a strong predictor of government spending, verifying statutory requirements of federal spending programs (GAO 2006, Murray 1992). This section provides a framework for thinking about the source of variation in our instrument and provides conditions under which the untestable exclusion restriction can be a reliable working assumption.

Population levels used to allocate federal spending are updated with a rule that changes discontinuously in Census years. When final counts are released, previous population estimates are replaced with the new Census counts. In other years, population estimates are updated annually using data on births, deaths and migration to account for population growth. This change of data source creates a shock to the population levels used in calculating federal spending. The exclusion restriction for our instrument is that the discrepancy in population estimates between the two methodologies is not related to factors that would, independently of federal spending, influence employment and income.

The timing of the release of the new Census counts is a crucial feature of our identification strategy. As mentioned in Section 2, the final population counts for the 1980, 1990, and 2000 Censuses were released two years after they were conducted. A powerful test of our identification strategy leverages this timeline to examine the validity of the identification strategy. Government spending should not be correlated with the Census shocks in the years before the final counts are released. A correlation here would indicate that confounding factors might be the source of the correlation between the instrument and government spending. A lack of dependence is consistent with the assumption that the instrument is working through the statutory channels that we enumerate in Section 2.

We now present a framework that formalizes the source of variation in the Census shock. This model relates the instrument to specific factors that could potentially challenge the exclusion restriction. A general model of the postcensal (PC) and

Census (C) estimates of population can be written as follows:

$$Pop^i_{c,t} = g^i(Pop^*_{c,t}, u^i_{c,t}) \text{ for } i = C, PC,$$

for county c and year t where $Pop^*_{c,t}$ is actual population and $u^i_{c,t}$ are measurement errors. A specific yet flexible model of the population estimates is obtained by the following log-linear model

$$\log(Pop^i_{c,t}) = \alpha^i + \lambda^i \log(Pop^*_{c,t}) + u^i_{c,t} \text{ for } i = C, PC,$$

where the measurement error $u^i_{c,t}$ is independent of $\log(Pop^*_{c,t})$. This model allows both population estimates derived from administrative data and Census counts to have specific level-biases of magnitude α^i and growth-biases given by λ^i.

The Census shock is defined as the difference between these estimates in Census years

$$CS_{c,t} = \log(Pop^C_{c,t}) - \log(Pop^{PC}_{c,t}) = \Delta\alpha + (\lambda^C - \lambda^{PC}) \log(Pop^*_{c,t}) + \Delta\mu_{c,t} \qquad (1)$$

where $\Delta\alpha = \alpha^C - \alpha^{PC}$ and $\Delta u_c = u^C_{c,t} - u^{PC}_{c,t}$.[15] We can then express the exclusion restriction in the context of an IV regression as

$$\begin{aligned} 0 &= \mathbb{C}ov(CS_{c,t}, \varepsilon_{c,t}) \\ &= \mathbb{C}ov(\Delta\alpha + (\lambda^C - \lambda^{PC}) \log(Pop^*_{c,t}) + \Delta u_{c,t}, \varepsilon_{c,t}) \\ &= (\lambda^C - \lambda^{PC})\mathbb{C}ov(\log(Pop^*_{c,t}), \varepsilon_{c,t}) + \mathbb{C}ov(\Delta u_{c,t}, \varepsilon_{c,t}) \\ &= (\lambda^C - \lambda^{PC})\mathbb{C}ov(\log(Pop^*_{c,t}), \varepsilon_{c,t}), \end{aligned}$$

where $\varepsilon_{c,t}$ is the structural error term from a given outcome equation on income or employment such as in equation (3) below. The third line assumes $\Delta\alpha$ is constant. The fourth line uses the fact that $\Delta u_{c,t}$ is the difference between measurement errors that are uncorrelated with the true population and the IV error term. The exclusion restriction is then satisfied when $\lambda^C - \lambda^{PC} = 0$ or when $\mathbb{C}ov(\log(Pop^*_{c,t}), \varepsilon_{c,t}) = 0$.

A world where both estimation methodologies approximate true population with

[15] Note, however, that the source of variation is coming from differences in population estimates and not from changes in actual population. This is important as population can be endogenous to economic factors that might confound the estimation strategy.

added classical measurement error would have $\alpha^i = 0$ and $\lambda^i = 1$ for $i = C, PC$. In such a world, the Census shock would be the combination of two classical measurement errors and would be unrelated to any other factors that could confound the identification strategy. The model in equation (1) suggests that the classical measurement error model, while sufficient, can be overly restrictive. A sufficient, yet less restrictive condition, for the Census shock to be unrelated to true population and any other confounding factors is that $\lambda^C = \lambda^{PC}$. That is, both estimation methodologies may be level-biased ($\Delta\alpha \neq 0$) but the degree of growth-bias would have to be the same across methodologies. If this condition were satisfied it would be the case that the source of variation in the Census shock is exogenous to factors that would affect the outcomes of interest.

This condition is not directly testable as it relies on knowledge of the true population $Pop^*_{c,t}$. We therefore provide a number of robustness checks by including in our regressions several demand and supply shocks to the local economy that are believed to influence true population movements. In Section 7.2 we use local labor demand shocks obtained from the unobserved component of an autoregressive model used by Blanchard and Katz (1992), an industry share-shift instrument proposed by Bartik (1991), and a measure of supply shock of immigrants developed by Card (2001) as potential drivers of true population. We show that our estimates are robust to the inclusion of these factors in our specifications. We also provide in Section 7.4 an alternative construction of the instrument in a GMM framework that implements the model of this section. This procedure minimizes the correlation between the generated instrument and the supply and demand shocks we consider using optimal GMM weights.

5 Identifying Variation

The previous section motivated the source of variation in the Census shock as the difference between measurement errors from two population estimates and provided general conditions under which the exclusion restriction is satisfied. This section discusses the implementation of our conceptual experiment and describes the variation of the instrument.

To implement our strategy, we need both Census counts and concurrent population estimates. Unfortunately, the postcensal population estimates are not available in Census years. Even without population estimates, we can still gauge the amount of variation between population estimates and Census counts by referring to the population growth rates presented in Figures 1 and 2. This evidence indicates that resetting population estimates to Census count levels generates a large amount of cross-sectional variation. While the amount of variation is visible from the average county population growth rates, it is important to notice that population growth rates cannot be used as instruments for government spending as these are a combination of measurement error, which is a valid source of identifying variation, and true population growth, which is endogenous to economic factors that could confound the identification strategy. In order to implement the identification strategy outlined in the previous section, we need to isolate the component of population change that is due to measurement error. To do this, we need to calculate the counterfactual postcensal population estimates.

We produce population estimates for Census years using data on the components of change of population. Because we do not have access to all the data used by the Census Bureau, we estimate the following regression with the aim of approximating the methodology used to produce the estimates:

$$\Delta Pop_{c,t}^{PC} = \phi_1 Births_{c,t} + \phi_2 Deaths_{c,t} + \phi_3 Migration_{c,t} + u_{c,t}.$$

This calibration equation ensures that we can adequately replicate year-to-year population changes using the publicly-available data. The regression is estimated separately by decade on years for which population estimates are available (which excludes Census years). Birth and mortality data taken from Vital Statistics and County-to-County Migration Data from the IRS are used to estimate the change. This procedure gives us estimated population growth rates from which we can extrapolate population levels for Census years using the latest available data This means, for example, that we calibrate the components of change across counties to the average population growth over the 1990s. We then use population estimates for 1999 to produce population estimates for 2000. The resulting estimates are then used to produce the counterfactual postcensal population levels $\widehat{Pop_{c,\text{Census}}^{PC}}$.[16] Using these estimates, we

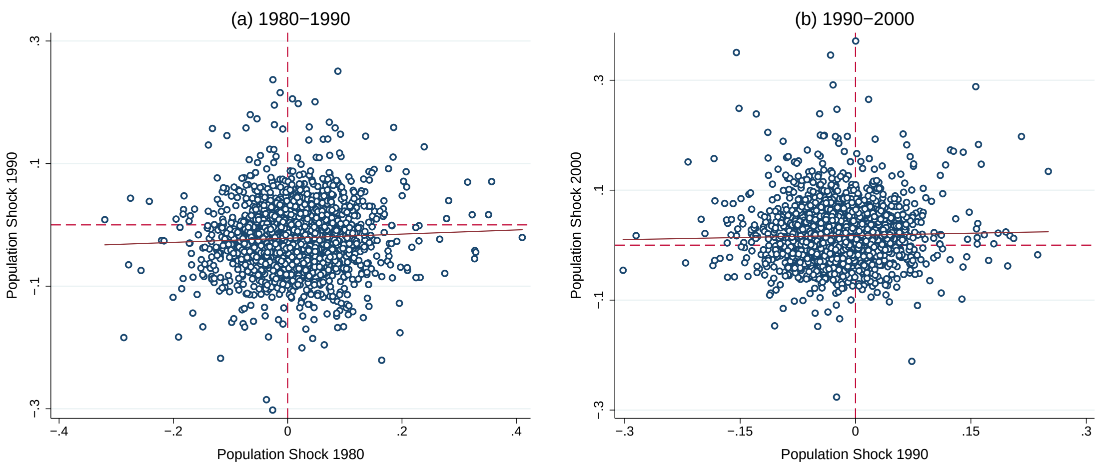
[16] Alternative methods of estimating the counterfactual postcensal population estimates, including a raw sum of the components of change (i.e. $\Delta Pop_{c,t}^{PC} = Births_{c,t} - Deaths_{c,t} + Migration_{c,t}$) and using an AR(3) time series model, produce similar estimates and do not alter our main results.

define the Census shock as[17]

$$CS_{c,\text{Census}} = \log(Pop^{C}_{c,\text{Census}}) - \log(\widehat{Pop^{PC}_{c,\text{Census}}}).$$

In order to characterize the source of variation of the instrument, we first consider whether the instrument is geographically correlated. If the Census shock is positively correlated for counties in a given region this might be evidence that the Census shock is related to a region-wide shock that might also explain the outcomes of interest. An analysis of variance (ANOVA) shows that only 5% of the variation can be explained by MSA and state indicators, ruling out concerns of geographic correlation. Since most of the variation is at the county level, this also shows it is the right level of analysis for our natural experiment.

Figure 5: Serial Correlation of the Census Shock



Second, we consider whether the instrument is serially correlated. Figure 5 presents the scatter plots of the Census shocks across decades. These plots demonstrate that there is virtually no serial correlation in the shocks across Censuses. In both graphs, the slopes of the correlation are very flat and not statistically different from zero. This feature of the Census shocks is consistent with measurement error being the source of the variation in the instrument. Importantly, it is evidence against confounding

[17]Notice that while our instrument has been generated in an estimation step prior to the main estimations, it is not necessary to adjust the standard errors of our instrumental variable estimates (see e.g. Wooldrige 2002).

factors that could be driving the variation across areas and that are known to be strongly serially correlated such as illegal immigration.

6 Census Shock and Government Spending

This section documents the first-stage relationship between our instrument and federal spending. We focus on three particular aspects of this correlation. First, we show that the Census shock is a strong predictor of growth in federal spending. Second, the timing of this growth is consistent with the timeline presented in Section 2 and that the Census shock is not related to growth in federal spending before the final Census population counts are released. Third, we present two falsification tests that show that the Census shock works only through spending programs that actually use population levels in allocating spending.

As mentioned above, a large number of federal spending programs depend on local population estimates. There is a delay in the adoption of new population levels since federal agencies have some discretion in the way new population figures are adopted in the allocation of funds (GAO 1990). These two factors suggest that the change in population due to the Census shock might affect spending for several years after the new Census count are released. We estimate this dynamic relationship with the following regression

$$\Delta F_{c,t} = \alpha_{s,t} + \gamma_t CS_{c,\text{Census}} + \Gamma X_{c,\text{Census}} + e_{c,t}, \tag{2}$$

where $\Delta F_{c,t}$ is the growth rate in federal spending, $\alpha_{s,t}$ are state-year fixed effects and $X_{c,\text{Census}}$ is a vector of control variables that includes the population predictors discussed in the previous section as well as demographic covariates that are also available in Census years. The full list of controls includes the value of the Blanchard-Katz employment shock (B-K) in the Census year as well as two of its lags, the Bartik industry share-shifter in the Census year and two lags, and the Card immigration supply shock in the Census year. The B-K shocks are constructed from the residuals of an AR(3) process using log changes in county-level employment. The industry share shifter relies on predicted changes in total county employment from national changes at the 3-digit industry level and base year industry composition of employment. We use employment data from the BEA for both measures of local demand shocks. The immigration supply shock is constructed in a similar fashion but relies on the predicted changes to immigrant population based on national changes in immigration levels by country of origin. We define base year foreign-born population composition as the number of foreign-born individuals by country of birth from the previous Census. If, for example, there was a large influx of Eastern European immigrants in the US between 1970 and 1980, counties with large Eastern European-born populations in 1970 would be likely to experience a large influx of immigrants. Card (2001) shows this proxy is a predictor of changes in total population.

The demographic covariates in Census years we use include the share of urban, black, Hispanic and foreign-born populations. We also include the share of individuals who moved into the county within the last five years, the share of families beneath the official poverty threshold, the log real median household income within the county, the average number of persons per household as well as the share of the population between the ages of 20 and 34 and over 65. Finally, notice in equation (2) that while $CS_{c,\text{Census}}$ is realized every ten years, this relationship allows for an impact on federal spending that is specific to each year relative to the Census year.

Figure 6 plots the individual γ_t's with a 95% confidence interval with year 0 being the year in which the Census is conducted. Importantly, this graph shows that the Census shock does not impact federal spending growth in the years before the Census counts are released. This feature of the relation between the shock and subsequent federal spending is an important test of the validity of our identification strategy. The graph shows that a positive Census shock is related to an increase in federal spending growth for the following four years. Once all agencies have adopted the new population counts, these counts become obsolete and no longer affect federal spending. This graph demonstrates that the instrument we develop provides exogenous variation in federal spending for several years even though the shock only occurs every ten years.

Figure 7 plots the cumulative effect on federal spending of the Census shock up to ten years after the Census is conducted.[18] This graph shows that, once the new Census counts are released, federal spending growth increases for the following four years, then levels off. The cumulative effect is statistically significant and has a large magnitude. A shock of 10% leads to an increase of 3% in the growth of federal spending in a given county over the next ten years. This elasticity implies the average

[18] The cumulative effect and the variance for this efect are obtained by adding up the individual γ_t's. So the cumulative effect for year T is given by $\sum_{t=1}^{T} \gamma_t$.

county will receive an additional $2,000 in federal spending per person "found" over the following decade.[19]

Figure 6: First Stage Effect by Year

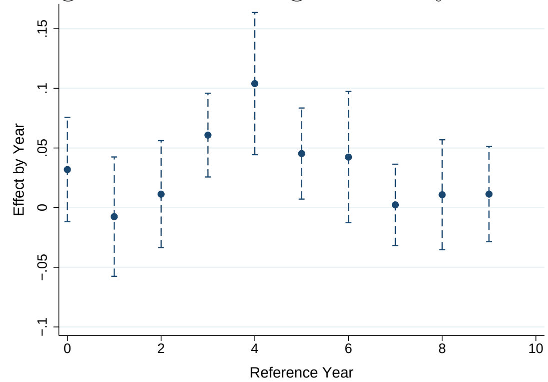
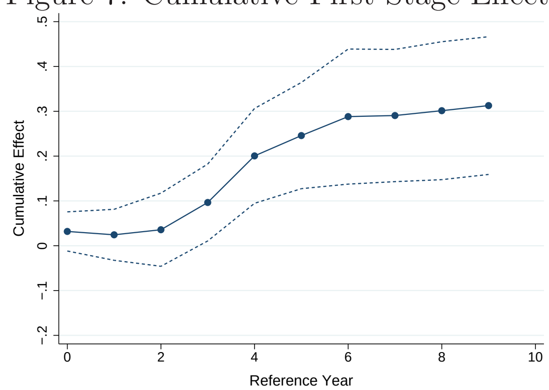


Figure 7: Cumulative First Stage Effect

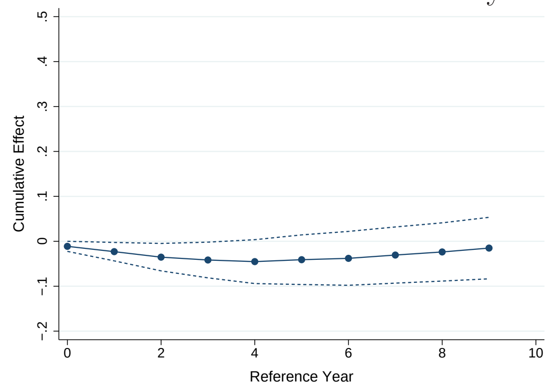


The dynamics shown in these graphs are a hallmark of the identification strategy of this paper. The timing of the effects can be tested against the alternative hypothesis that all of the effects occur during a single year. The hypothesis that all of the coefficients except the first are zero is tested and rejected at standard levels of significance. We rely on the dynamics in this graph in our instrumental variables specification and restrict the estimation to reference years 2 through 5 (i.e. 82-85)

[19] A GAO review of the 15 largest formula grant programs for fiscal year 1997 found that federal spending in a given state would increase by $480 per person per year had the 1990 Census state populations been adjusted for undercount (GAO 1999).

as these are the years in which our exogenous source of variation has a significant impact on the growth of federal spending.

Figure 8: Falsification Test: Social Security Payments



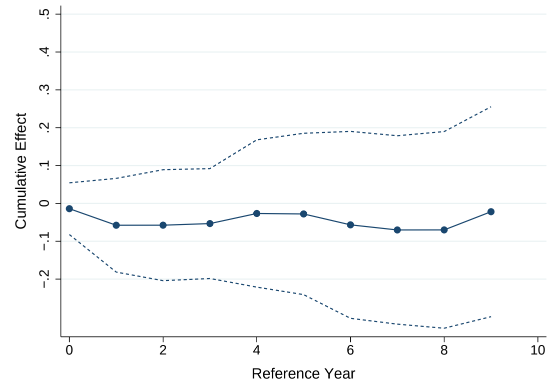
Two falsification tests provide further evidence that the relationship in Figure 7 is not due to statistical coincidence and indeed reflects the natural experiment described in Section 2. Figure 8 presents the estimates of the cumulative impact of the Census shock on Social Security spending at the county level. Since Social Security spending consists of direct payments to individuals, this category of spending should not depend on population estimates and should be uncorrelated with the Census shock.[20] This intuitive feature is borne out in the data.

Figure 9 plots the cumulative impact of a future Census shock on government spending growth. If a shock that has not been realized is a predictor of government spending then it might be the case that the instrument is identifying local areas with time-invariant, county-specific characteristics that are associated with increases in the growth of government spending. One example of such a characteristic would be a powerful congressional representative. The graph, however, shows that future shocks do not predict growth in government spending.

Finally, Figure 10 plots the cumulative effect of the Census shock on the different categories of federal spending in the CFFR. Consistent with statutory and narrative evidence, the *Direct Payments to Individuals* and *Grants* categories are the most responsive to the population shock. The Grants category increases gradually all

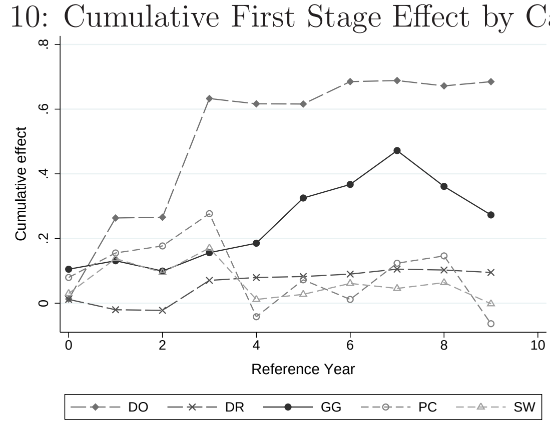
[20]Notice that an indirect positive relationship could arise if beneficiaries of social security moved to locations with growth in federal spending that is related to a large census shock. While migration is responsive to increases in government spending, social security beneficiaries are unlikely to be sensitive along this margin given their underlying low mobility.

Figure 9: Falsification Test: Future Census Shock



through year 7 whereas the *Direct Payments* jumps discontinuously after year 2 and remains flat afterwards. These two categories account for around 35 percent of total domestic spending as measure by the CFFR. Our natural experiment therefore captures variation in spending programs that account for 50% of the government budget excluding Social Security. As the graph shows, the other spending categories do not show long run responses and are not statistically different from zero.

Figure 10: Cumulative First Stage Effect by Category



Note: DO is Direct Payments to Individuals, DR is Retirement Payments, GG is grants, PC is Procurement and Contracts and SW is Salaries and Wages. The annual spending growth for each category is topcoded at ±100%, which affects 11.5% of the observations.

7 Estimates of Local Fiscal Multipliers

This section presents our main estimates. We first present a reduced form version of the results that shows that our identification strategy is borne out in the raw data. We then present OLS and IV regressions and interpret these results in terms of elasticities and multipliers.

7.1 Reduced Form Results

The estimates in Table I provide evidence of the impact of federal spending on local economic outcomes that does not rely on statistical models. The main idea is to compare growth in federal spending and economic outcomes across counties with large and small Census shocks. To this end, we group counties in each decade into bins based on quantiles of the Census shock. We then relate how each of these bins perform in terms of growth in federal spending, income, earnings and employment and calculate the implied federal spending elasticities of income, earnings and employment. To produce Table I, we computed the average growth in spending and outcomes relative to all other counties in the same state for a given decade.

Table I: Reduced Form Estimates of Growth Rates
(a) Percentage Changes

Census Shock		Average Percentage Change by Bin			
Bin		Fed Spend	Income	Earnings	Employment
0-20%	-6.15%	-2.57%	-0.63%	-0.54%	-0.78%
20-40%	-1.85%	-0.45%	-0.37%	-0.30%	-0.28%
40-60%	-0.07%	-0.19%	-0.30%	-0.66%	-0.38%
60-80%	1.74%	0.42%	0.28%	0.35%	0.36%
80-100%	6.33%	2.44%	1.09%	1.38%	1.25%

(b) Implied Elasticities

Census Shock Bin	Pop Elast of Fed Spend	Fed Spend Elasticity of Income	Earnings	Employment
0-20%	0.42	0.25	0.21	0.3
20-40%	0.24	0.82	0.67	0.63
40-60%	2.96	1.56	3.4	1.94
60-80%	0.24	0.68	0.85	0.87
80-100%	0.39	0.45	0.57	0.51
Mean	0.8	0.7	1.1	0.9

Census shocks are ordered by quintile in the first column. Average county growth rates for outcome variables are relative to state-decade averages for reference years 1 through 5. See text for details

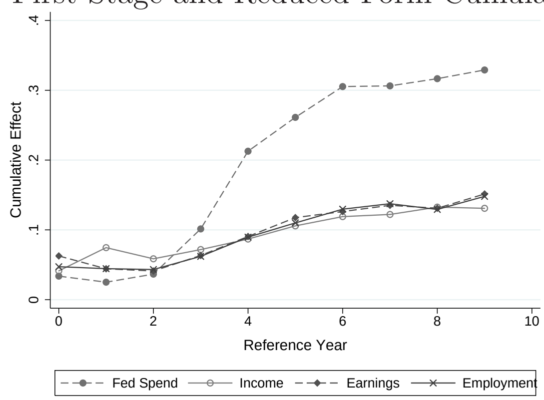
Panel (a) in Table I shows how the average growth rates of spending and the outcome variables vary by bin of the Census shock. The first column characterizes the variation in the instrument. Comparing the first and the last bin we see that the population shock can vary by up to 12.48 percentage points in our sample. The second column shows how this population shock translates into growth in federal spending. For the first bin, containing counties with a Census shock in the bottom quintile, a Census shock of -6.15% yields a decrease in spending of 2.57% over 6 years.

The monotone ordering of the averages in the first column is a mechanical effect from ranking the counties by Census shock. The fact that changes in federal spending in the second column are also ranked is evidence that our instrument is a strong predictor of federal spending. For all the outcome variables, it is also the case that a negative Census shock leads to negative spending growth and negative impacts on economic outcomes. The fact that the magnitudes of these changes are generally ranked in ascending order provides evidence that the identification strategy that we pursue in this paper does not rely on a particular statistical model.

Panel (b) of Table I shows the implied elasticities by taking ratios of the percentage changes in the first panel. These elasticities imply that a 10% Census shock leads to an increase in government spending of 0.8%. The last three columns compute the federal spending elasticity of each of the outcomes by dividing the change in the outcome by the change in spending. These elasticities are large in magnitude with median values of 0.68 for income, 0.67 for earnings and 0.63 for employment. As we show below, our instrumental variables estimates are close to these values. Furthermore, it is reassuring to find that, excluding the middle bin with an almost zero-valued shock, the spending elasticities of the outcomes are relatively stable across bins.

Finally, we provide in Figure 11 a graphical presentation of both the first stage relation between the Census shock and federal spending and the reduced form effect for all three economic outcomes. The dynamics of the reduced form results are similar to the first stage in that the first two years following the Census have a relatively flat profile before increasing between years two to six. The effects on income, earnings and employment are not as large as the effect on federal spending and the ratio of the reduced form to first stage curves at any point are the IV estimates themselves. However, the correlation between the Census shock and the economic outcomes is slightly higher than for federal spending before the release of the Census counts in year two.

Figure 11: First Stage and Reduced Form Cumulative Effect



7.2 OLS and IV Estimates

This subsection presents our main estimates of the impact of government spending on income and employment growth. As in the previous section, we restrict our analysis to reference years 2 though 5 as these are the years during which our instrument impacts government spending. Contrary to the raw estimates above, however, this section analyzes annual growth rates for both outcomes and government spending. The log-difference specification eliminates county fixed effects and provides estimates in the form of elasticities. We quantify the relationships explored in the previous section by linear models of the form:

$$\Delta y_{c,t} = \alpha_{s,t} + \beta \Delta F_{c,t} + \varepsilon_{c,t}, \tag{3}$$

where $\Delta y_{c,t}$ is the log change of a given outcome as a function of $\Delta F_{c,t}$, the log change in federal spending, and state-year fixed effects. We allow for correlation of the error term at the state level.

As a prelude to our causal estimates of the impact of federal spending on economic outcomes we present OLS regressions that do not address the potential endogeneity of federal spending. Table II reports the results from OLS regressions for income, earnings, and employment using federal spending as measured by the CFFR. The OLS estimates are statistically significant but of small economic magnitude.

Table II: OLS Estimates of the Impact of Federal Spending on Economic Outcomes

	(1) Income	(2) Earnings	(3) Employment
Federal Spending	0.041***	0.049***	0.036***
	(0.010)	(0.010)	(0.007)
Observations	36,410	36,410	36,410
R-squared	0.30	0.17	0.22

Regressions include years 1982-85, 1992-95 and 2002-05. Standard errors clustered at the state level in parentheses. All regressions include state-year fixed effects. * p<0.10, ** p<0.05, *** p<0.01

The central contribution of this paper is to provide causal estimates of the impact of federal spending on economic outcomes at the local level. We instrument for changes in federal spending in equation (3) using the most recent Census shock:

$$\Delta F_{c,t} = \alpha_{s,t} + \gamma CS_{c,\text{Census}} + e_{c,t}$$

Table III provides estimates from our IV specifications. The first column provides the estimates from the first-stage of our instrumental variables regression. A 10% Census shock leads to an increase of 0.7% of spending growth at yearly level. Over a period of four years this represent an increase of 2.6%. A concern in instrumental variables estimation is that weak instruments can lead to large biases in the estimand whenever the errors are correlated with the instrument (e.g. Bound, Jaeger, and Baker 1995). To address this issue, we provide the F-statistic of the test that the instrument has a zero coefficient in the first stage equation. An F-statistic of 20 is greater than conventional levels of acceptance, suggesting that our instrument is not subject to the weak instrument problem.

Table III: IV Estimates of the Impact of Federal Spending on Economic Outcomes

	(1) First Stage	(2) Income	(3) Earnings	(4) Employment
Federal Spending		0.527***	0.579***	0.561***
		(0.152)	(0.179)	(0.153)
Census Shock	0.066***			
	(0.015)			
Observations	36,410	36,410	36,410	36,410
R-squared	0.13	.	.	.
F-Stat Instr	20.13	20.13	20.13	20.13
IV = OLS (p-value)		0.00	0.00	0.00

Regressions include reference years 2 to 5. Standard errors clustered at the state level in parentheses. State-year fixed effects included. * p<0.10, ** p<0.05, *** p<0.01

Columns (2) through (4) in Table III present our baseline estimates of the impact of federal spending on local economic outcomes. For all three outcomes we find economically large and statistically significant estimates of the impact of government spending. The estimated elasticity imply that a 10% increase in federal spending causes a 5.3% increase in total personal income, a 5.8% increase in earned income, and a 5.6% increase in employment. These estimates are more than ten times larger than the corresponding OLS estimates and are statistically different. The direction of the bias in the OLS estimates suggests that federal spending might be directed towards counties with unobserved characteristics that are correlated with low economic growth.

Table IV: IV Results With Population Controls

	(1) First Stage	(2) Income	(3) Earnings	(4) Employment
Federal Spending		0.539***	0.569***	0.545***
		(0.184)	(0.194)	(0.158)
Census Shock	0.055***			
	(0.012)			
B-K Emp Shock	0.052***	0.014	0.004	0.012
	(0.018)	(0.017)	(0.016)	(0.013)
L1 B-K	0.043	0.049**	0.029	0.061***
	(0.027)	(0.019)	(0.021)	(0.013)
L2 B-K	0.075***	-0.007	0.011	0.032**
	(0.017)	(0.015)	(0.019)	(0.012)
Ind Share Shifter	-0.251**	-0.008	-0.068	0.018
	(0.110)	(0.059)	(0.080)	(0.052)
L1 Share Shifter	0.110	0.073	0.084	0.118**
	(0.070)	(0.048)	(0.062)	(0.044)
L1 Share Shifter	-0.017	0.000	0.071***	0.047***
	(0.017)	(0.015)	(0.015)	(0.012)
Migration Shifter	-0.016**	0.002	0.006*	0.003
	(0.007)	(0.004)	(0.003)	(0.003)
Observations	35,962	35,962	35,962	35,962
R-squared	0.14	.	.	.
F-Stat Instr	21.21	21.21	21.21	21.21
IV = OLS (p-value)		0.02	0.04	0.00

The B-K Emp shock variable is the Blanchard-Katz employment residual. Ind Share Shifter is the Bartik industry share-shifter and Migration shifter is the Card immigration shock variable. L1 and L2 denote lag operators. Regressions include reference years 2 to 5. Standard errors clustered at the state level in parentheses. State-year fixed effects included. * p<0.10, ** p<0.05, *** p<0.01

Before proceeding to interpret our results as fiscal multipliers, we consider the impact that other covariates might have on our estimates. Consider first the role of demand and supply shocks. As prefaced in Section 4, a potential confounder of our identification strategy is that the Census shock might be correlated with demand and supply shocks that can have a direct impact on the outcomes of interest. We address this concern in Table IV by including the employment and migration shocks. The IV regression now becomes

$$\Delta y_{c,t} = \alpha_{s,t} + \beta \Delta F_{c,t} + \Gamma X_{c,\text{Census}} + \varepsilon_{c,t},$$

where the vector $X_{c,\text{Census}}$ includes the local demand and supply shocks listed earlier. Table IV shows that the IV estimates are not sensitive to the inclusion of the additional variables even though some of the controls are themselves strongly correlated with the dependent variable. Furthermore, the first stage relationship becomes stronger with an F-statistic above 20.

Our final set of results accounts for observable county characteristics. These covariates include county demographic characteristics. Including these covariates is important for two reasons. First, while federal spending depends on population, the explicit formulas that compute spending are also a function of other characteristics such as income, proportion of people below the poverty line, and the age profile of the population. Including some of the covariates might better approximate the nonlinearities in the formulas that determine government spending. Second, given that these formulas link federal spending to demographic and income characteristics, controlling for these covariates provides estimates that are local to the communities that are most affected by our natural experiment. Table V presents the IV estimates with these covariates. For both income and employment, it is the case that the estimates are slightly smaller but they remain statistically significant.[21] The following section translates these estimates into parameters of policy interest: income multiplier and the cost per job created.

[21]The IV estimates are still 10 times larger than the OLS estimates with the full set of controls (available from the authors) and the differences remain statistically significant.

Table V: IV Results Controlling for Shocks and Covariates

	(1) First Stage	(2) Income	(3) Earnings	(4) Employment
Federal Spending		0.419**	0.320	0.397**
		(0.187)	(0.245)	(0.172)
Census Shock	0.051***			
	(0.012)			
B-K Emp Shock	0.025	0.003	-0.007	-0.002
	(0.015)	(0.013)	(0.016)	(0.010)
L1 B-K	0.004	0.028*	0.005	0.035***
	(0.029)	(0.015)	(0.020)	(0.012)
L1 B-K	0.053***	-0.012	0.010	0.024**
	(0.017)	(0.014)	(0.017)	(0.010)
Ind Share Shifter	-0.249*	-0.027	-0.120	-0.007
	(0.126)	(0.055)	(0.096)	(0.050)
L1 Share Shifter	0.026	0.034	0.027	0.069*
	(0.075)	(0.039)	(0.057)	(0.036)
L2 Share Shifter	-0.007	0.005	0.076***	0.054***
	(0.016)	(0.013)	(0.014)	(0.011)
Migration Shifter	-0.014*	-0.003	0.000	-0.001
	(0.007)	(0.003)	(0.004)	(0.003)
Urban	0.007**	-0.002	0.000	-0.001
	(0.003)	(0.002)	(0.003)	(0.003)
Black	-0.031***	-0.005	-0.019*	-0.023***
	(0.008)	(0.008)	(0.010)	(0.006)
Hispanic	-0.004	-0.002	-0.016***	-0.015***
	(0.009)	(0.007)	(0.005)	(0.005)
Foreign Born	-0.032	0.051**	0.061***	0.031
	(0.049)	(0.023)	(0.022)	(0.021)
Moved Last 5 Years	0.095***	0.046**	0.080***	0.059***
	(0.012)	(0.020)	(0.026)	(0.021)
Share Poor Families	0.020	0.001	0.018	0.019*
	(0.017)	(0.009)	(0.016)	(0.011)
Log Median HH Inc	0.016**	0.002	0.007	0.007
	(0.007)	(0.004)	(0.007)	(0.005)
Age 20-34	-0.038	0.020	0.036**	0.027
	(0.039)	(0.016)	(0.017)	(0.017)
Age 65+	0.005	−0.047*	-0.002	-0.028
	(0.046)	(0.024)	(0.031)	(0.023)
Observations	35,962	35,962	35,962	35,962
R-squared	0.14	.	.	.
F-Stat Instr	16.38	16.38	16.38	16.38
IV = OLS (p-value)		0.09	0.34	0.03

Urban, Black, Hispanic, Foreign Born, Moved Last 5 Years, Age 20-34 and Age 65+ are shares of total county population. All controls variables use values from Census years. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

7.3 Implied Multipliers

The income multiplier and the cost per job created have recently resurfaced as key parameters in the policy debate. This subsection provides estimates of these parameters by transforming our elasticities into marginal effects. These multipliers are interpreted as the total impact of policy interventions that include direct impacts of government spending (such as government purchases or government hires) as well as impacts through indirect channels (such as the economic activity created by new government employees).

Table VI: Marginal Effects

(a) Income Multiplier

Quantile					Mean
10	25	50	75	90	
1.19	1.49	1.88	2.38	3.04	2.02
(0.53)	(0.66)	(0.84)	(1.06)	(1.36)	(0.90)

(b) Cost per Job

Quantile					Mean
10	25	50	75	90	
$19,395	$24,055	$30,388	$38,650	$49,691	$32,914
(8,389)	(10,405)	(13,144)	(16,718)	(21,493)	(14,237)

Standard errors in parenthesis. Estimates computed using the same county-year observations as in Table V.

Table VI presents the marginal effects implied by our preferred specification that includes all controls as presented in Table V.[22] The transformation from elasticities into marginal effects is a non-linear transformation that relies on the ratio of economic

[22] We follow Cameron and Trivedi (2009) in this transformation. We first generate the expected level of the outcome

$$\mathbb{E}[y_{c,t}] = \exp\{\log(y_{c,t-1}) + \alpha_{s,t} + \beta\Delta\log(F_{c,t}) + \Gamma X_{c,\text{Census}}\}\mathbb{E}[\exp\{\varepsilon_{c,t}\}]$$

The income multiplier is now given by

$$\frac{\partial\mathbb{E}[Inc_{c,t}]}{\partial F_{c,t}} = \beta_{Inc}\frac{\exp\{\log(Inc_{c,t-1}) + \alpha_{s,t} + \beta\Delta\log(F_{c,t}) + \Gamma X_{c,\text{Census}}\}\mathbb{E}[\exp\{\varepsilon_{c,t}\}]}{F_{c,t}},$$

where $\mathbb{E}[\exp\{\varepsilon_{c,t}\}]$ is estimated by paired-bootstrapping of the exponentiated residuals. The cost per job created is given by

$$\left[\frac{\partial\mathbb{E}[Emp_{c,t}]}{\partial F_{c,t}}\right]^{-1} = \frac{F_{c,t}}{\beta_{Emp}\exp\{\log(Emp_{c,t-1}) + \alpha_{s,t} + \beta\Delta\log(F_{c,t}) + \Gamma X_{c,\text{Census}}\}\mathbb{E}[\exp\{\varepsilon_{c,t}\}]}.$$

outcomes to government spending. Cross-county variation in this ratio generates a distribution of multipliers rather than a single number. Table VI presents different quantiles of this distribution as well as the mean. These distributions are not symmetric and have extreme values that influence the mean due to the unequal distribution of economic outcomes and federal spending across counties. For this reason, we rely on the median value of the multiplier for the discussion below. Evaluating the multipliers at median values of these ratios give median multipliers of 1.88 for income and a cost per job created of $30,000. Computing the multiplier and cost per job using national averages gives slightly higher but very similar values that are not statistically different.

With these estimates, we can consider the impact of a marginal increase in government spending in a representative county. In terms of employment, the results in Table VI suggest that a $1 million increase in federal spending would create 33 new jobs at a cost of around $30,000 per job for the median county. In terms of income, a fiscal multiplier of 1.88 implies that a $1 increase in federal spending would raise personal income by $1.88 in the median county. While the multiplier interpretation is natural for the income and earnings multipliers, it is worth reconsidering the interpretation of the cost per job created. Our results do not imply that a new employee would be paid $30,000. Rather, it can be seen as the share of the cost per job that accrues to the government. The remaining share is paid by employers as a result of increased economic activity generated by government spending through direct and indirect channels. Combining the income and employment multipliers we could posit that the job created would have a total remuneration of 1.88*$30,000=$56,400.

It is worth noting that interesting patterns in the heterogeneity of the impacts of government spending may arise. The variation in the estimates of Table VI, however, is due solely to the non-linear nature of the transformation and the variation in the ratio of economic outcomes to government spending. In Section 9 we characterize the heterogeneity of outcomes using a quantile regression framework that describes how the impact of government spending differs throughout the distribution of county growth rates.

A simpler derivation that ignores the impact of the error term and uses actual, as opposed to predicted, outcome levels yields similar estimates.

7.4 Instrument Construction via GMM

Section 7.2 shows that our main estimates are robust to including measures of demand and supply shocks in the instrumental variables specification. This evidence validates the construction of the instrument and shows that our results are not due to shocks to the local economy that could otherwise confound our causal interpretation. This section presents an alternative and novel approach to generating the instrument. It relies on a GMM framework to implement the errors-in-measurement model presented in Section 4. The objective is to generate an instrument that is as close to being orthogonal to true population changes as possible and that only relies on variation from measurement error.

Recall the model in Section 4 defines the instrument as

$$CS_{c,t} = \log(Pop^{C}_{c,t}) - \log(Pop^{PC}_{c,t}) = \Delta\alpha + (\lambda^{C} - \lambda^{PC})\log(Pop^{*}_{c,t}) + \Delta\mu_{c,t}.$$

For our previous estimates, we acknowledged that λ^C and λ^{PC} might differ and that our instrument could be correlated with true population changes. We showed that including variables to control for this did not changes our main conclusions. We now propose a GMM procedure to estimate the ratio $\frac{\lambda^C}{\lambda^{PC}}$ and the difference in level biases $\Delta\alpha$. The intuition for this approach is that at the true values of these parameters, the instrument will be uncorrelated with factors that are correlated with true population.

To see this, suppose $\tilde{\lambda} = \frac{\lambda^C}{\lambda^{PC}}$ and $\Delta\tilde{\alpha} = \alpha^C - \tilde{\lambda}\alpha^{PC}$ are known. The instrument generated by

$$\begin{aligned}\widetilde{CS}_{c,t} &= \log(Pop^{C}_{c,t}) - \tilde{\lambda}\log(Pop^{PC}_{c,t}) - \Delta\tilde{\alpha} \\ &= (\lambda^{C} - \tilde{\lambda}\lambda^{PC})\log(Pop^{*}_{c,t}) + \Delta\tilde{\mu}_{c,t} \\ &= \Delta\tilde{\mu}_{c,t}\end{aligned}$$

is thus independent of true population $Pop^{*}_{c,t}$ and identifies exogenous changes in federal spending only through the difference in measurement errors $\Delta\tilde{\mu}_{c,t}$.

The GMM estimation minimizes the weighted sum of moments given by

$$\sum_{c}\left(\log(Pop^{C}_{c,j}) - \tilde{\lambda}_{r,j}\log(\widehat{Pop^{PC}_{c,j}}) - \Delta\tilde{\alpha}_{r,j}\right) Z_{c,j} = 0,$$

where $\widehat{Pop}_{c,j}^{PC}$ are generated as in Section 5. The parameters to be estimated $\tilde{\lambda}_{r,j}$ and $\Delta\tilde{\alpha}_{r,j}$ are specific to decades $j = 1980, 1990$ and 2000 and Census regions $r =$ Northeast, Midwest, South, and West. The model is estimated separately by decade and is pooled across Census regions.

Our vector of instruments $Z_{c,j}$ includes the Blanchard and Katz and Bartik shocks in Census years along with two lags as well as the Card immigration-supply shock. We also include the share of black, Hispanic and foreign born populations, the share of people who lived in a different county five years prior to the Census as well as the log median household income as additional instruments.[23] This gives us a total of 12 moments to identify two parameters for each decade and region and therefore provides a test of the over-identifying restrictions. Failing to reject this test implies that the GMM-adjusted instrument is not correlated with the control variables in $Z_{c,j}$.

Table VII presents the results from this estimation. Panel (a) shows that the estimated $\tilde{\lambda}$'s are very close to 1. They are very precisely estimated and the departures from 1 are statistically significant. We can also reject the equality of the $\tilde{\lambda}$'s across regions. However, as we will see below, those departures have minimal effects on our IV results. The estimates of the relative biases $\Delta\tilde{\alpha}$ are larger and vary by region and decade. No region has consistently the same sign for $\Delta\tilde{\alpha}$ across decades. However, for the 1990 Census, we see that for all regions Census counts were more downward biased than the Census Bureau estimates. This is consistent with aggregate evidence concerning the 1990 Census undercount. The decade-specific tests of over-identifying restrictions have p-values of 0.39 for 1980, 0.44 for 1990, and 0.43 for 2000. This implies that the adjusted Census shock is not correlated with factors that affect population movements such as local demand shocks.

Panel (b) shows the IV results using the GMM-adjusted instrument $\widetilde{CS}_{c,t}$ and including the full set of controls from Table V. The first stage relationship is somewhat weaker than Table V both in terms of the point estimate and the strength of the relationship between the Census shock and federal spending. The F-statistic of the instrument is still strong enough to rule out a weak instrument problem. The estimated elasticities of income and earnings in columns (2) and (3) are slightly smaller, with the coefficient for income being 15% smaller. Nonetheless, it is still statistically

[23] We select these covariates because they have the highest explanatory power in Table V. Including the full set of controls does not change our results.

significant and the median income multiplier implied by this estimate is 1.59, which is not statistically different from the multiplier presented in Section 7.3. The impact of federal spending on employment in column (4) is identical to the estimated impact using the baseline instrument in Table V.

The approach presented in this section provides an alternative construction of the instrument that ensures our identification strategy is not confounded by demand and supply shocks to the local economy that also affect true population changes. The estimated $\tilde{\lambda}$'s from the errors-in-measurement model are very close to 1 and confirm that the assumption that $\lambda^{C} = \lambda^{PC}$ is reasonable in this context. Our IV results are also very similar to those found in our baseline regressions. These results bolster our confidence that our estimates are in fact identified by random unsystematic differences in measurement between two sources of population estimates.

Table VII: Instrument Construction via GMM

(a) Estimation of $\tilde{\lambda}$ and $\Delta\tilde{\alpha}$ by Census and Region

		(1) 1980	(2) 1990	(3) 2000
Northeast	$\tilde{\lambda}$	0.992***	0.999***	1.007***
		(0.001)	(0.001)	(0.001)
	$\Delta\tilde{\alpha}$	0.107***	0.004	-0.066***
		(0.009)	(0.009)	(0.010)
Midwest	$\tilde{\lambda}$	1.007***	1.006***	0.997***
		(0.001)	(0.001)	(0.000)
	$\Delta\tilde{\alpha}$	-0.080***	-0.083***	0.038***
		(0.010)	(0.009)	(0.005)
South	$\tilde{\lambda}$	1.002***	1.000***	0.997***
		(0.002)	(0.001)	(0.001)
	$\Delta\tilde{\alpha}$	0.012	-0.029**	0.055***
		(0.023)	(0.012)	(0.006)
West	$\tilde{\lambda}$	1.018***	1.008***	0.998***
		(0.004)	(0.000)	(0.001)
	$\Delta\tilde{\alpha}$	-0.178***	-0.095***	0.043***
		(0.038)	(0.003)	(0.011)
Observations		2,991	3,012	2,999
OverID test (p-val)		0.388	0.443	0.432

(b) IV Results Using GMM-Adjusted Instrument

	(1) First Stage	(2) Income	(3) Earnings	(4) Employment
Federal Spending		0.355*	0.231	0.397*
		(0.208)	(0.287)	(0.205)
GMM Census Shock	0.043***			
	(0.012)			
Observations	35,962	35,962	35,962	35,962
R-squared	0.14	.	.	.
F-Stat Instr	12.21	12.21	12.21	12.21

Standard errors clustered at the state level in parentheses. Regressions include full set of controls from Table V. * p<0.10, ** p<0.05, *** p<0.01

8 Aggregation

In this section, we present the results when we aggregate our methodology at the MSA and state levels. The aggregated analysis is important since there might be spillovers in the effects of fiscal shocks across counties. Depending on the sign of these spillovers, we could be underestimating or overestimating the total effect of government spending on the economy. For example, if federal spending goes to building a road in a county and some of the workers are hired from other areas or materials are purchased elsewhere, the increased demand for inputs could have positive effects outside the targeted county. The county-level results would then be underestimating the total impact of federal spending. If, however, the increase in federal spending leads to in-migration from neighboring areas and higher wages due to a decrease in labor supply, this could potentially reduce the number of firms in other counties. This kind of effect could then lead to negative spillovers and our county-level results would overestimate the total impact. Note that we do not attempt to control for spillovers across states since aggregating at the national level would be irrelevant in this context. The rationale for the natural experiment is that the Census population shocks lead to a redistribution of federal funds across geographical areas, not to an increase in total spending.

The aggregated analysis is done by summing all relevant county-level variables within the larger geographic areas.[24] For example, we define the Census shock at the

[24]We also used a different approach where the spending shock for neighboring counties was included (and instrumented) directly in the county-level regression. Neighbors were defined as being within

MSA level as the percentage difference between the Census population count and the concurrent population estimate of the entire MSA. We grouped all counties not within an MSA to a rest-of-state area. When a county was located in more than one MSA, we assigned it to the MSA in which it had the largest share of its population. We used the 1993 OMB definition of MSA to be consistent over time. Our sample consists of 281 MSAs and a total of 328 areas (with 47 rest-of-state areas). The aggregation obviously leads to smaller sample sizes and less variation in our instrument. Whereas the average county experiences a population shock (in absolute value) of 3.9%, the average MSA's population shock is 2.7%. The variance of the shock at the MSA level is also one quarter of the county level variance. State level shocks are 2.2% on average for the contiguous states and the District of Columbia with a variance half the MSA-level shocks. We expect this aggregation to lead to a loss of power, which could weaken the statistical relationship between the Census shock and federal spending in the first stage. Indeed, the F-statistic for the instrument decreases significantly and is subject to a weak instrument problem when we aggregate. This can potentially lead to biased IV results in the MSA and state level regressions.

We present in Table VIII the results for the IV regressions of MSA-level personal income, total earnings and employment on federal spending. These regressions include indicator variables for the nine Census regions interacted with year fixed-effect and the full set of demand and supply shocks and demographic covariates we use in the county-level specification. Standard errors are clustered at the MSA level to account for possible autocorrelation in the error term. Similar to the county-level results, the IV estimates are larger and statistically different from the OLS estimates (available from the authors). More importantly, we note that the MSA-level point estimates are larger than the county-level estimates. This would indicate that the impact of government spending does not decrease as we aggregate and is consistent with positive spillovers across counties. The F-statistic on the instrument in the first stage, however, is 5.12 which does not rule out a weak instrument. The income multiplier for the median MSA implied by the elasticity in column two is 2.05, which is larger than the multiplier at the county level.

Table IX presents the same regressions at the state level. We use 48 states in three different Censuses for a sample size of 576 observations. The estimated effects are now much larger than at both the county and MSA levels. These elasticities

---

the same MSA or state. Results are very similar to the aggregated results.

represent implausibly large multipliers, but as we have mentioned this could be due to a weak first stage. Taken together, the aggregated results seem to suggest there are positive spillovers across counties and federal spending has a beneficial impact on the economic outcomes of areas beyond the initial recipient counties.

Table VIII: IV Results for MSA Aggregation

	(1) First Stage	(2) Income	(3) Earnings	(4) Employment
Federal Spending		0.432*	0.585*	0.645**
		(0.248)	(0.327)	(0.297)
Census Shock	0.121**			
	(0.054)			
Observations	3,924	3,924	3,924	3,924
R-squared	0.09	.	.	.
F-Stat Instr	5.12	5.12	5.12	5.12
OLS=IV (p-value)		0.09	0.04	0.00

Standard errors clustered at the MSA level in parentheses. Full set of controls from table V and region-year fixed effects included. * p<0.10, ** p<0.05, *** p<0.01

Table IX: IV Results for State Level Aggregation

	(1) First Stage	(2) Income	(3) Earnings	(4) Employment
Federal Spending		1.157***	1.209***	0.933***
		(0.299)	(0.378)	(0.248)
Census Shock	0.227*			
	(0.115)			
Observations	576	576	576	576
R-squared	0.39	.	.	.
F-Stat Instr	3.87	3.87	3.87	3.87
IV=OLS (p-value)		0.02	0.03	0.03

Standard errors clustered at the state level in parentheses. Full set of controls from table V and region-year fixed effects included. * p<0.10, ** p<0.05, *** p<0.01

We finally note that, beyond the issue of spillovers, the other fundamental difference between cross-sectional analyses (at any level of aggregation) and time-series designs is the fact that we cannot identify the effects of fiscal shocks common to all areas. For example, including year fixed effects in an attempt to control for unrelated macroeconomic shocks will also capture any nation-wide effect of the spending change itself in a particular year. As mentioned earlier, candidates for such nation-wide shocks related to our instrument are the impact of future taxes on the current behavior of consumers and firms and the effect of the monetary policy response to a fiscal expansion.

9 Heterogeneity

The heterogeneity of impacts discussed in Section 7.3 described the cross-sectional variation of multipliers resulting from the non-linear transformation of elasticities into multipliers as well as from the cross-sectional variation in spending, previous levels of outcomes, and other covariates. This section characterizes the heterogeneity of the impacts of government spending in terms of elasticities using an instrumental variable quantile regression approach recently developed by Chernozhukov and Hansen (2008).

Our main regression estimates show that government spending has large impacts on the conditional means of income and employment across counties. A more complete characterization of the impacts of government spending over the entire distribution of income and employment growth rates is also possible. This could answer the question as to whether faster or slower growing counties are more impacted by government spending. This could also address the potential for government spending to reduce inequality in economic outcomes across counties. Quantile regression provides an appealing approach to characterizing the impact of government spending on different parts of the outcome distribution. However, methods that combine quantile regression with instrumental variables have only recently been proposed in the literature.[25] We implement the instrumental variable quantile regression (IVQR) procedure developed by Chernozhukov and Hansen (2008) that takes advantage of our identification strategy to produce causal estimates.

Before introducing the IVQR approach, we consider a quantile regression estimate that does not account for the endogeneity of government spending. For a given quantile q of the outcome distribution of $\Delta y_{c,t}$, we estimate the conditional quantile function

$$Q_q(\Delta y_{c,t}) = \alpha_t^q + \beta^q \Delta F_{c,t} + \Gamma X_{c,\text{Census}} \tag{4}$$

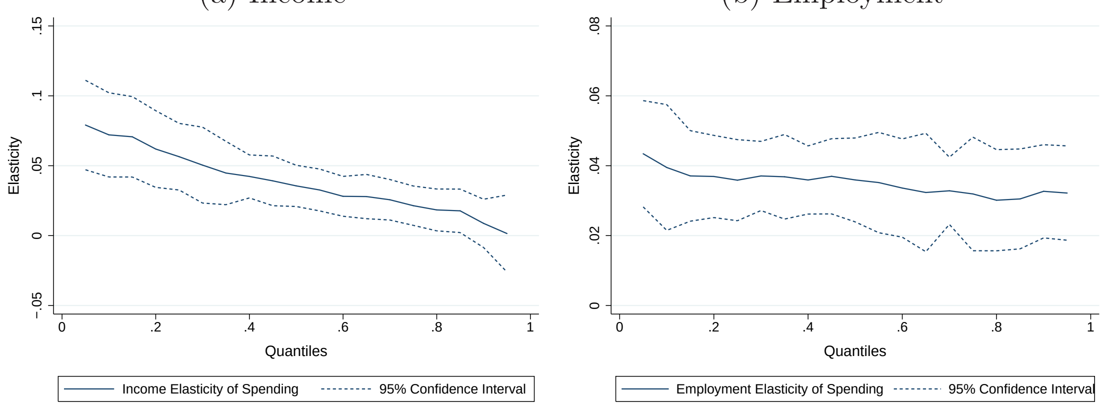
with α_t^q year fixed effects, $\Delta F_{c,t}$ the log change in federal spending and county covari-

[25]See Angrist and Pischke (2009) for a review of recent developments.

ates $X_{c,\text{Census}}$. We do not include state fixed effects as we are interested in comparing counties relative to the national distribution. Including state fixed-effects would change the interpretation of the results by limiting the comparison to counties within the same state. Figure 12 plots the β^q's from these estimations for 20 values of q for each of our main outcomes. Panels (a) and (b) show coefficients that are of a similar magnitude than the OLS estimates and have relatively flat profiles. These results would lead us to believe government spending has a modest impact across the distribution of outcomes and does little to reduce the inequality in income, earnings and employment across counties.

Figure 12: Quantile Effects - Endogenous Federal Spending

(a) Income (b) Employment



The IVQR we implement acknowledges the endogeneity of government spending and provides consistent estimates of the β^q's that are not subject to endogeneity bias. Consider the alternative quantile function

$$\tilde{Q}_q(\Delta y_{c,t}) = \alpha_t^q + \beta^q \Delta F_{c,t} + \gamma^q CS_{c,\text{Census}} + \Gamma X_{c,\text{Census}} \tag{5}$$

where we add the county-level Census shock $CS_{c,\text{Census}}$. The IVQR framework uses the insight that, at the true value of the structural parameter β^q, the instrumental variable will not influence the conditional quantile, so that $\gamma^q = 0$. To compute estimates of β^q, the IVQR framework finds values of β^q such that γ^q is as close to zero as possible. Distance from zero, in this context, is measured using the F-statistic for testing $\gamma^q = 0$.[26]

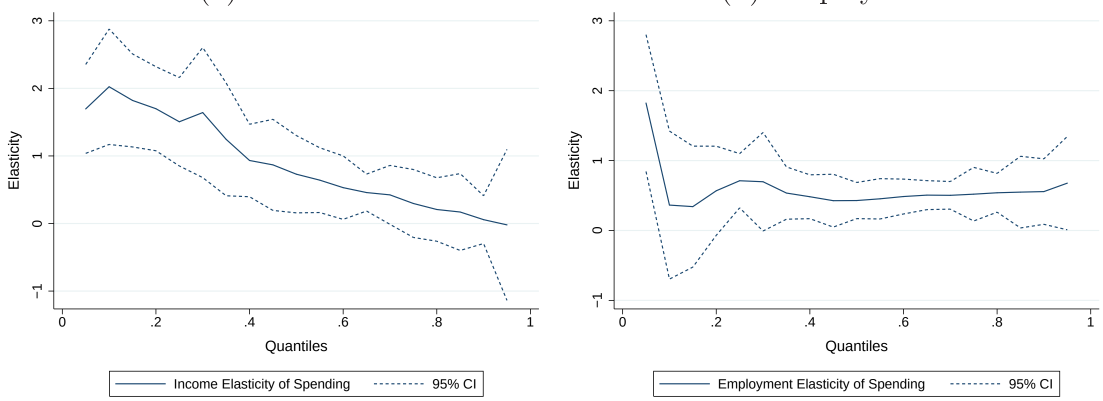
[26]For a given quantile q, the algorithm used in the estimation is as follows

1. Use a golden search method (see, e.g., Miranda and Fackler (2002)) to find the value of $\tilde{\beta}^q$

Figure 13 presents the result of these estimations for income and employment for 20 values of q. These figures confirm our previous findings that instrumental variable estimates suggest a much larger effect of government spending on income and employment than do methods that do not account for the endogeneity of government spending.

Figure 13: Quantile Effects - IVQR

(a) Income (b) Employment



These graphs further show that counties with lower income growth are more impacted by changes in government spending than counties with higher income growth. This differential effect can be interpreted either as a "redistributional effect," i.e. poor areas benefit more from federal spending, or as a "stabilizing effect." The latter highlights the view of fiscal federalism as providing insurance against local shocks. Because federal spending has such a large impact in low growth counties, it could be an effective way to help areas experiencing temporary negative shocks. Since we

---

that minimizes the F-statistic for testing $\gamma^q = 0$. The F-statistic is computed by first fixing a value of $\tilde{\beta}^q$, estimating the quantile regression

$$\tilde{Q}_q(\Delta y_{c,t}) = \alpha^q_{s,t} + \tilde{\beta}^q \Delta F_{c,t} + \gamma^q CS_{c,census} + \theta X_{c,census}$$

and testing $\gamma^q = 0$. Grid search methods were also implemented with similar, albeit computationally more intensive, results.

2. Confidence intervals and standard errors are computed using a paired-bootstrap of step 1 to account for inter-cluster correlation at the state level. The dual inference approach of Chernozhukov and Hansen (2008) was also implemented and yielded similar results.

Note that the inference procedure for the IVQR is robust to weak instruments. An important caveat, however, is that the results we estimate are consistent estimates of the structural parameters in equation (4) only if the model is correctly specified. Alternative methods that are robust to model misspecification have been proposed by Chen and Pouzo (2009).

do not include dynamics in our analysis, we cannot differentiate between counties which are experiencing temporary shocks and those which are permanently better-off. Regardless of these interpretations, the downward-sloping profiles in Figure 13 (a) shows that increasing government spending not only raises income but also decreases inequality of income growth rates across counties.

The results for employment growth seem to be constant across the distribution of outcomes. The estimates are similar to the IV estimates and are ten times larger than the quantile regression estimates. However, government spending does not seem to influence the relative inequality of employment growth. Note that this is evidence that our IVQR is not subject to some form of misspecification that is common to both income and employment. A challenge to the IVQR framework in equation (5) would have to explain the steep pattern for the effect on income and the flat pattern for employment.

10 Conclusion

The impact of government spending on the economy is one of the most important policy questions we face in the current macroeconomic context. The federal government is spending vast amounts of money in the hope of stimulating the economy, but many economists and policy analysts claim fiscal policy has a limited impact in the short term and cripples long term growth prospects. In this paper, we propose a new methodology to estimate critical parameters. We rely on cross-sectional instead of time-series variation and propose a new instrumental variable to identify the causal impact of federal spending. This new approach is a powerful yet transparent way to measure several important parameters such as the income multiplier, the cost per job created, and the inequality-reducing effect of government spending.

We find a large effect of government spending on local economic outcomes. The timing of the impact is consistent with the release of the new Census counts and our estimates are robust to the inclusion of potential confounders, thereby strengthening the case for causal identification. We have shown that aggregation of our methodology at the MSA and state levels does not cause our estimates to decrease. We also show that government spending provides higher returns in depressed areas and that it has contributed to reducing inequality in employment across counties.

Future work could focus on the interaction of federal spending with local business cycles, since recent papers have shown that the income multiplier might be larger during recessions (Auerbach and Gorodnichenko 2010, Christiano et al. 2009, Woodford 2010). It would also be of interest to document the dynamic relationship between the new measure of spending shocks and economic outcomes by using more flexible estimation specifications. This would make the current results more comparable to macroeconomic estimates of impulse response functions and would allow the estimation of the long term effects of fiscal shocks on local economies.

The instrument we introduce in this paper is also relevant for the field of urban and regional economics. The exogenous variation in government spending we propose constitutes a shock to local labor and housing markets that can be used to test general spatial equilibrium models where agents move across locations to benefit from higher wages or cheaper amenities (Roback 1982, Kline 2010). The empirical strategy we proposed can be used to further our understanding of agglomeration effects as well as migration, wages and housing price responses to government spending shocks. Such models can also be used to estimate the deadweight loss of federal spending as a place-based policy due to the potential distortions in the locational decisions of individuals (Glaeser and Gottlieb 2009, Glaeser 2008, Moretti 2010). We plan to address these questions in future work.

References

Acemoglu, D., Finkelstein, A. and Notowidigdo, M.J. (2009): "Income and Health Spending: Evidence from Oil Price Shocks" NBER Working Paper No. 14744.

Angrist, J. D. and Pischke, J.-S. (2009): *Mostly Harmless Econometrics: An Empiricist's Companion,* Princeton, Princeton University Press.

Auerbach, A. and Gorodnichenko, Y. (2010): "Measuring the Output Responses to Fiscal Policy," Working Paper, UC Berkeley, August 2010.

Bartik, T.J. (1991): *Who Benefits from State and Local Economic Development Policies?,* Books from Upjohn Press, W.E. Upjohn Institute for Employment Research.

Blanchard, O.J. and Katz, L. K. (1992): "Regional Evolutions," *Brookings Papers on Economic Activity*, Economic Studies Program, The Brookings Institution, 23, 1-76.

Blanchard, O.J. and Perotti, R. (2002): "An Empirical Characterization Of The Dynamic Effects Of Changes In Government Spending And Taxes On Output," *The Quarterly Journal of Economics*, 117(4), 1329-1368.

Blumerman, L.M. and Vidal, P.M. (2009): "Uses of Population and Income Statistics in Federal Funds Distribution - With a Focus on Census Bureau Data," U.S. Census Bureau, Government Division Report Series, Research Report #2009-1, Washington, DC.

Bound, J., Jaeger, David., and Baker, R. (1995): "Problems with Instrumental Variables Estimation When the Correlation between the Instruments and the Endogenous Explanatory Variables is Weak." *Journal of the American Statistical Association*, 90(430), 443-50.

Bureau of Economic Analysis (2010): "State Personal Income and Employment: Methodology," Washington, DC.

Busso, M, Gregory, J. and Kline, P. (2010): "Assessing the Incidence and Efficiency of a Prominent Place Based Policy," Working Paper.

Cameron, C. and Trivedi, P.K. (2009): *Microeconometrics Using Stata,* Stata Press.

Card, D. (2001): "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration," *Journal of Labor Economics,* 19(1), 22-64.

Chen, X. and Pouzo, D. (2010): "Efficient Estimation of Semiparametric Conditional Moment Models with Possibly Nonsmooth Residuals," *Journal of Econometrics*, forthcoming.

Chernozhukov, V. and Hansen, C. (2008): "Instrumental variable quantile regression: A robust inference approach," *Journal of Econometrics*, 142, 379-398.

Chodorow-Reich, G., Feiveson, L., Liscow, Z., and Woolston, W. (2010): "Does State Fiscal Relief During Recessions Create Jobs," Working Paper.

Christiano, L., Eichenbaum, M., and Rebelo, S. (2009): "When is the government spending multiplier large?," NBER Working Paper No. 15394.

Clemens, J., and S. Miran (2010): "The Effects of State Budget Cuts on Employment and Income," Working Paper.

Congressional Research Service (2008): "Medicaid: The Federal Medical Assistance Percentage (FMAP)," CRS Report for Congress #RL32950, Washington DC.

Davis, S.J., Loungani, P., and Mahidhara, R. (1997): "Regional Labor Fluctuations: Oil Shocks, Military Spending, and Other Driving Forces," *International Finance Discussion Papers*, 1997-578.

Fatás, A. and Mihov, I. (2001): "The Effects of Fiscal Policy on Consumption and Employment: Theory and Evidence," CEPR Discussion Paper No. 2760.

Fishback, P., and V. Kachanovskaya (2010): "In Search of the Multiplier for Federal Spending in the States During the New Deal," Working Paper.

Glaeser, E. (2008): *Cities, Agglomeration, and Spatial Equilibrium*, Oxford University Press.

Glaeser, E. and Gottlieb, J.D. (2009): "The Wealth of Cities: Agglomeration Economies and Spatial Equilibrium in the United States," *Journal of Economic Literature,* 47(4), 983-1028.

Glaeser, E., Sacerdote, B., and Scheinkman, J. (2003): "The Social Multiplier," *Journal of the European Economic Association*, 1(2-3), 345-353.

Gordon, N. (2004): "Do Federal Funds Boost School Spending? Evidence from Title I," *Journal of Public Economics*, 88(9-10), 1771-92.

Government Accountability Office (1987): "A Catalog of Federal Aid to States and Localities," GAO/HRD-87-28, Washington, DC.

——— (1990): "Federal Formula Programs: Outdated Population Data Used to Allocate Most Funds," GAO/HRD-90-145, Washington, DC.

——— (1999): "Effects of Adjusted Population Counts on Federal Funding to States," GAO/HEHS-99-69, Washington, DC.

——— (2006): "Illustrative Simulations of Using Statistical Population Estimates for Reallocating Certain Federal Funding," GAO-06-567, Washington, DC.

Kline, P. (2010): "Place Based Policies, Heterogeneity, and Agglomeration," *American Economic Review: Papers & Proceedings*, 100, 383-387.

Long, J.F. (1993). "Postcensal Population Estimates: States, Counties, and Places," U.S. Bureau of the Census, Population Division Working Paper No. 3.

Miranda, M.J. and Fackler, P.L. (2002): *Applied Computational Economics and Finance,* Cambridge, MA: MIT Press

Moretti, E. (2010): "Local Labor Markets," in *Handbook of Labor Economics*, forthcoming.

Murray, M.P. (1992): "Census Adjustment and the Distribution of Federal Spending," *Demography*, 29(3), 319-332.

Nakamura, E. and Steinsson, J. (2010): "Fiscal Stimulus in a Monetary Union: Evidence from U.S. Regions," Working Paper.

Ramey, V., and Shapiro, M. (1997): "Costly Capital Reallocation and the Effects of Government Spending," Carnegie-Rochester Conference Series on Public Policy.

Ramey, V. (2010): "Identifying Government Spending Shocks: It's All in the Timing," *Quarterly Journal of Economics,* forthcoming.

Roback, J. (1982): "Wages, Rents, and the Quality of Life." *Journal of Political Economy,* 90(6),1257-78.

Romer, C., and Bernstein, J. (2009): "The Job Impact of the American Recovery and Reinvestment Plan," Accessed October 25, 2010. `http://www.ampo.org/assets/library/184_obama.pdf`

Shoag, D. (2010): "The Impact of Government Spending Shocks: Evidence on the Multiplier from State Pension Plan Returns," Working Paper.

U.S. Census Bureau (2001): "Census 2000 Summary File 1: Technical Documentation," Washington, DC.

——— (2010a): "National Intercensal Estimates (1990-2000)," Accessed October 25, 2010. `http://www.census.gov/popest/archives/methodology/intercensal_nat_meth.html`

——— (2010b): "1980 Census of Population and Housing: History," Accessed October 25, 2010. `http://www2.census.gov/prod2/decennial/documents/1980/proceduralHistory/1980CPH_TOC.pdf`

——— (2010c): "1990 Census of Population and Housing: History," Accessed October 25, 2010. `http://www2.census.gov/prod2/decennial/documents/1990/history/Chapter1-14_TOC.pdf`

——— (2010d): "Consolidated Federal Funds Report," Accessed October 25, 2010. `http://www.census.gov/govs/cffr/`

Pande, R. (2003): "Can Mandated Political Representation Provide Disadvantaged Minorities Policy Influence Theory and Evidence from India," *American Economic Review*, 93(4), 1132-1151.

West, K. and Fein D.J. (1990): "Census Undercount: An Historical and Contemporary Sociological Issue," *Sociological Inquiry,* 60 (2), 127-141.

Wilson, D. (2010): "Fiscal Spending Multipliers: Evidence from the 2009 American Recovery and Reinvestment Act," Federal Reserve Bank of San Francisco Working Paper 2010-17.

Woodford, M. (2010) "Simple Analytics of the Government Expenditure Multiplier," NBER Working Paper No. 15714.

Wooldridge, J. (2002): *Econometric Analysis of Cross Section and Panel Data*, MIT Press, Cambridge, MA.

Data Appendix

In order to construct the panel of county population and the instrument, we use postcensal population estimates published by the Census Bureau from 1971 to 2009. This distinction between postcensal and intercensal is important. The latter are retrospectively revised to account for the error of closure in Census years whereas the former are the contemporaneous estimates produced every year to tract population growth. Intercensal population estimates are not relevant for our study since federal spending only depends on the contemporaneous estimates. Most of the earlier data are archived at the Inter-University Consortium for Political and Social Research (ICPSR) (http://www.icpsr.umich.edu/). For the years 1971 to 1974, we use the *Population Estimates of Counties in the United States* (ICPSR 7500). For years 1975 to 1978, we use the data from the *Federal-State Cooperative Program: Population Estimates* study (ICPSR 7841 and 7843). No postcensal population estimates were published for 1979, 1980, 1989, 1990 and 2000. For 1981 to 1988, we use population data from the *County Statistics File 4* (CO-STAT 4) (ICPSR 9806). Data for Census years and from 1991 onward were taken directly from the Census Bureau's website (http://www.census.gov/popest/estimates.html) since the postcensal estimates are still available. Local and state population estimates are produced jointly by the Census Bureau and state agencies. The Federal-State Cooperative Program has produced the population estimates used for federal funds allocation and other official uses since 1972.

Birth data from Vital Statistics are taken from the micro data files available at the NBER (http://www.nber.org/data/) for the years 1970 to 1978. We use the Centers for Disease Control and Prevention's (CDC) *Compressed Mortality Files (*http://wonder.cdc.gov/) for years 1979 to 1988 and tables published in the Vital Statistics, *Live births by county of occurrence and place of residence* for years 1989 and 1990. Data for 1991 to 2009 are taken directly from the Census Bureau's components of growth data files available on the Census website. Data on county level deaths are taken from the NBER's *Compressed Mortality* micro data files from 1970 to 1988 and from the CDC's *Compressed Mortality* tabulated files from 1989 to 2006. County level deaths for 2007 to 2009 were taken directly from the Census Bureau's components of growth files.

Migration data come from the IRS Statistics of Income. Years 1978 to 1992 were taken from the *County-to-County, State-to-State, and County Income Study Files, 1978–1992* (ICPSR 2937) and *Population Migration Between Counties Based on Individual Income Tax Returns, 1982-1983* (ICPSR 8477). The most recent years are available directly from the IRS SOI's website (http://www.irs.gov/taxstats/).

Data on Federal spending were taken from the Census Bureau's *Consolidated Federal Funds Reports*. These reports have been produced annually since 1983 and provide a detailed account of the geographic distribution of federal expenditures. 1983 and 1984 data are available on CD-ROM from the Census Bureau and for downloading from the SUDOC Virtualization Project housed at the University of Indiana's Department of Computer Science (http://www.cs.indiana.edu/svp/). Data from 1985 to 1992 are available for download individually by year at the ICPSR. The Census Bureau's website has CFFR releases from 1993 onwards. Data on federal spending prior to 1983 is available from the *Geographic Distribution of Federal Funds* for fiscal years 1981 and 1982 (ICPSR 6043 and 6064) and from the *Federal Outlays* dataset from 1976 to 1980 (ICPSR 6029). Note that debt servicing, international payments and security and intelligence spending are not covered in the CFFR. See Census Bureau (2010d) for further details.

The county demographic and economic covariates were downloaded from the Census Bureau's *American FactFinder* (http://factfinder.census.gov/) for the 1990 and 2000 Censuses. Data for the 1980 and 1970 Censuses were downloaded from the National Historical Geographic Information System (NHGIS) (http://www.nhgis.org/).