Heterogeneous Spending, Heterogeneous Multipliers*

Umberto Muratori†

Pedro Juarros‡

Daniel Valderrama§

July 11, 2022

Abstract

Do local fiscal multipliers depend on what the government purchases? To answer this question, we build a panel of military spending at the product-MSA level and implement an instrumental variables research design. Purchases of services have larger effects on employment than spending in goods. Differences in the labor intensity of industries producing goods or services are an important mechanism behind these results. Intuitively, the larger is the spending on labor-intensive industries, the larger is the amount of spending translating into increases in labor income and consumption, generating multiplier effects. Our estimates suggest spending in services drives the aggregate fiscal multiplier.

JEL Codes: E12, E32, E62, F33, H56, H57, R12

Keywords: Heterogeneous Local Fiscal Multipliers, Services and Goods Spending, Labor Intensity, Input-Output Linkages

*We are grateful to Isaac Baley, Laurent Bouton, Russell Cooper, Garance Genicot, Mark Huggett, Toshihiko Mukoyama, Martin Ravallion, Aysegul Sahin, and conference participants at the Max-Weber Programme June Conference for the helpful comments. We thank Digna Isabel Andrade Portilla for her excellent research assistance. We acknowledge the generous financial support from the Washington Center for Equitable Growth. The views in this paper are solely those of the authors and should not be interpreted as reflecting the views of the International Monetary Fund or World Bank. All errors are our own.

†European University Institute. *Corresponding Author*: umberto.muratori@eui.eu

‡International Monetary Fund.

§World Bank.

1 Introduction

The empirical government purchase multiplier literature commonly thinks government spending as a homogeneous good. Very little is known about the heterogeneity in the fiscal multipliers coming from the heterogeneity in the spending (Chodorow-Reich, 2019). This paper fills this gap by answering two questions. Do different types of government purchases —spending in goods and spending in services —generate different employment local fiscal multipliers? If yes, what are the mechanisms driving these differences?

The scarce research on this area may be explained by the lack of granular data on government purchases. The bulk of research on local fiscal multipliers has used variation in public spending coming from either the American Recovery and Reinvestment Act (Chodorow-Reich et al., 2012; Wilson, 2012; Conley and Dupor, 2013; Dupor and Mehkari, 2016) or from the government purchases made by the Department of Defense (Nakamura and Steinsson, 2014; Dupor and Guerrero, 2017; Demyanyk et al., 2019; Auerbach et al., 2020). These studies estimate local fiscal multipliers using aggregate spending, and do not delve into the heterogeneity of the estimates by type of spending.

We harmonize a rich military contract-level public spending dataset starting in 1979. Our data have two main features. First, we can identify the composition of government purchases using the products a contract requires to produce. The product codes are used to aggregate the fiscal spending into two categories: spending in goods and spending in services. Second, we can identify the MSA where a contract is performed. The geolocation of contracts allows us to quantify the causal effects of public spending on employment using standard cross-sectional instrumental variables research design (Nakamura and Steinsson, 2014).

The estimates of the employment local fiscal multipliers differ between shocks to spending in goods or services. One percent increase in services spending generates an on-impact employment response of 0.36 percent. The estimates increase as time passes, and become larger than one two years after the shock.[1] Goods spending employment multipliers are positive but non-significant at any horizon after the occurrence of the shock. The estimates for goods spending range between 0.06 percent and 0.19 percent, and they are between one-fourth and one-tenth smaller than the estimates for services spending.

As an external validity test, we compute the estimates for the employment fiscal multiplier using the aggregate spending. As we cover a longer time period than most studies, we want to verify our estimates for the spending by category are not the results of specific features in the aggregate spending. Our estimates for the aggregate spending are similar to ones reported in the literature. For example, our on-impact employment local fiscal multiplier for the aggregate spending is 0.14 percent compared to 0.19 percent of Auerbach et al. (2020) that is the closest study to ours in terms of geographic aggregation. It is likely to think the services component of the spending also drives the estimates in other studies, making crucial to understand the rationale behind the different patterns.

[1] Two-year fiscal multipliers are the benchmark estimates in the literature for three reasons. First, two years is considered the policy-relevant time frame for counter-cyclical policy packages. Second, it reduces the measurement error that comes from differences between the fiscal year, the unit of measurement of public spending, and the calendar year, the unit of measurement of economic activity. Third, the bias of not adopting a fully dynamic specification fades away as time elapses.

We explore other economic outcomes to better understand the mechanisms behind the differences between the effects coming from shocks to goods and services spending. Services spending leads to a significant increase in the household's personal income. This rise is the consequence of rises in the labor income (salary and wages). Shocks to goods spending have no significant effects on labor income, and as a consequence, the personal income remained unaffected.

We argue these differences are due to differences in labor intensity across industries that produce goods and services. Intuitively, in labor-intensive industries a larger share of the government dollars spent in production pass through the households. Precisely, public spending pushes up the demand for workers. In the new labor market equilibrium, a larger share of workers are hired, and labor income rises pushing up private consumption and production. The pass-through generates a virtuous circle leading to multiplier effects. In non-labor-intensive industries this pass-through mechanism is attenuated because labor demand reacts by less after a government shock.

We find evidence supporting labor-intensity as an important factor in driving the differences in the estimates by category of spending. Industries that produce the services demanded by the government are more labor-intensive than industries that produce goods. We divide each spending category into labor-intensive and non-labor-intensive sub-components, and we estimate the employment multipliers for the new sub-components. On the one hand, the effects from spending in non-labor-intensive industries are negligible or even negative. On the other hand, shocks to either type of spending directed to labor-intensive industries generate positive and significant increases in employment. We also document similar responses for labor income. We interpret these findings as evidence on the importance of labor intensity in determining the fiscal multipliers.

An interesting fact is that although the estimates of spending in services in labor-intensive industries and spending in goods in labor-intensive industries are both positive, their magnitude significantly differ. We provide some suggestive evidence that these differences could be partially explained by the production network. Industries with the largest share of services spending receive inputs from a larger number of industries than industries with the largest share of goods spending. A government spending directed to a more central industry in the production network generates an amplification of the shock through its input-output relationships, leading to greater estimates for the fiscal multiplier.

Differences in the size of the estimates could also signal alternative channels that act simultaneously to the labor intensity channel. We explore two alternative channels: tradeability and productivity gains. Goods are more tradeable than services. If the production of goods occurs in neighboring locations, one could observe small effects from goods spending on the local fiscal multiplier. Thus, the differences in the fiscal multipliers by category of spending could capture geographic spillovers. Our empirical results highlight geographic spillovers do not explain the heterogeneous fiscal multipliers by type of spending.

Government spending may improve productivity. Increases in spending may reduce the uncertainty about future profits, ease credit constraints leading to higher firms' turnover rates, or generate a faster growth of incumbents. If different types of spending generate different firm dynamics, the differences in the fiscal multipliers could be partially explain by productivity gains. Our tests suggest shocks to services spending increase the firms' entry rates, decrease the exit rates, and hurt the innovation activities. Shocks to goods spending positively affect the entry rates, but these effects are small. The results suggest the effects of government spending on productivity improvements are, if any, relatively small.

The main contribution of our work consists in documenting the differences in the fiscal multipliers by type of spending in opposition to the heterogeneity in economic conditions. Alonso (2017) and Dupor et al. (2021) study the effect of fiscal policy on the composition of consumption. Our paper investigates the other side of the coin. Instead of exploring the heterogeneity in the consumption responses to shocks, we focus on the heterogeneity of the shocks themselves, i.e. services and goods spending shocks. We show this heterogeneity is an important, yet overlooked, determinant of the spending multiplier. Boehm (2020) is a notable exception and studies the different reactions of output to government investment and consumption shocks. We classify spending into goods and services rather than consumption and investment. The importance of this classification is twofold. First, the US economy has been shifting from a goods economy to a services economy. Studying the differences between these two types of spending would help policymakers in designing effective fiscal interventions. Second, goods and services use different production technologies. Quantifying their effects separately would help in pinning down important determinants of the fiscal multiplier. Our findings provide complementary insights to the ones already known in the literature.

Our paper also relates to the vast literature on the determinants of the fiscal multiplier. This literature has highlighted the role of business cycle (Riera-Crichton et al., 2015; Suárez-Serrato and Wingender, 2016; Buchheim et al., 2020), trade openness (Ilzetzki et al., 2013; Corbi et al., 2019), exchange rate regime (Born et al., 2013), population demographics (Basso and Rachedi, 2021), households' heterogeneity (Hagedorn et al., 2019), labor market rigidity (Cole and Ohanian, 2004; Gorodnichenko et al., 2012), automatic stabilizers (Dolls et al., 2012; Galeano et al., 2021), public indebtedness (Ilzetzki et al., 2013), degree of monetary policy accommodation (Woodford, 2011), firm-size distribution (Juarros, 2020), and direction of the intervention (Barnichon et al., 2020) in amplifying the response of economic activity to public spending. This paper concentrates on the characteristics of the fiscal spending itself rather than the characteristics of the economy.[2] Our findings show the composition of the basket of products purchased by the government and their labor shares impact the effectiveness of fiscal policy. The paper provides new insights on the design of fiscal spending, and contributes to the policy debate on how to make fiscal stimuli more effective. Future works should further quantify the contribution of the structural decline in the labor share in explaining the decline in the effectiveness of fiscal policy at national-level (Blanchard and Perotti, 2002; Bilbiie et al., 2008).

Finally, our work is linked to a fast-growing literature on the transmission of sector-specific demand shocks to the aggregate economy and the relevance of the sector where a shock originates for its aggregate impact. Vom Lehn and Winberry (2022) show sector heterogeneity in the propensity to invest is a crucial for

[2]Studies on the infrastructure multiplier (see Ramey (2020) for references) also focus on the characteristics of the spending, but they do not simultaneously test the heterogeneity in the responses coming from different types of spending.

quantifying the amplification of a productivity shock. In contrast, Bouakez et al. (2021) and Bouakez et al. (2020) suggest network centrality in an input-output economy is the primary amplification mechanism of fiscal shocks. Flynn et al. (2021) highlight the importance of the consumer-side and find the heterogeneous incidence of fiscal shocks onto households with different marginal propensities to consume explain the heterogeneity in the fiscal multipliers. This paper focuses on the supply-side, and argues the sector where the fiscal spending is directed and how the industries are connected via their input-output relationship matter for the quantification of the multiplier. In this respect, we provide further support for the idea that fiscal authorities should design policy interventions by not only choosing how much to spend, but also where to direct that spending.

2 Empirical Strategy

2.1 Specification

Our empirical strategy builds on the works of Nakamura and Steinsson (2014), Dupor and Guerrero (2017), and Auerbach et al. (2020). We exploit the variation in military procurement spending across time and localities to estimate local fiscal multipliers.[3] The benchmark specification is defined as follows:

$$\frac{v_{l,t+k} - v_{l,t-1}}{v_{l,t-1}} = \beta^k \frac{G_{l,t+k} - G_{l,t-1}}{Y_{l,t-1}} + \alpha_l^k + \delta_{t+k} + \varepsilon_{l,t+k} \tag{1}$$

where v is an outcome of interest in location l at horizon $t+k$ with $k = \{0, \dots, 4\}$. The endogenous variable $G_{l,t+k} - G_{l,t-1}$ measures the change in military spending normalized by the first lag of the personal income of location l. The specification also includes locality fixed effects, α_l^k, to control for locality-specific trends, and the time fixed effects, δ_{t+k}, to account for any mechanical correlation between secular trends in military spending and unobserved local factors. We cluster the standard errors by locality.[4]

The coefficient β^k quantifies the multipliers in a window of k years. Since we normalize spending by personal income, the personal income multipliers correspond to the dollar amount of personal income produced by one dollar increase in government spending. This interpretation change for dependent variables different from the personal income, for example employment. In the case of the employment multiplier the interpretation is the change in the growth rate of employment produced by a 1-percentage point increase in government spending at local-level measured in personal income percentage points.[5]

We augment the benchmark specification to examine the heterogeneous effects of each government spend-

[3]In our estimation, the temporal dimension plays a more important, particularly for the computation of the shift-share for the instruments. As we disaggregated the military spending by sub-category, the within-locality volatility of the disaggregated spending is much higher than of the total spending. Thus, we need longer periods to compute representative local-level shares.

[4]Although Auerbach et al. (2020) use narrower geographic units, they cluster the error terms by state. For the specifications in which we use narrower geographic units, we have also checked the significance of our estimates by clustering the error terms at state-level. Our results (available upon request) remain unaffected by adopting their clustering decision.

[5]Our empirical strategy provides estimates for the local fiscal multiplier. Estimates of the local fiscal multiplier cannot be straightforwardly translated into a national multiplier. That's the case for two reasons. First, there could be spillover effects that propagate outside the locality's borders. Second, different assumptions and parameterizations of macroeconomic models could generate a broad range of estimates of the national multiplier. In the remainder of the paper, we will use local fiscal multiplier and fiscal multiplier interchangeably.

ing components. In this respect, we estimate the specification (2):

$$\frac{v_{l,t+k} - v_{l,t-1}}{v_{l,t-1}} = \beta_g^k \frac{G^g_{l,t+k} - G^g_{l,t-1}}{Y_{l,t-1}} + \beta_s^k \frac{G^s_{l,t+k} - G^s_{l,t-1}}{Y_{l,t-1}} + +\alpha_l^k + \delta_{t+k} + \varepsilon_{l,t+k} \tag{2}$$

where G^g represents the spending in goods, and G^s the spending in services, with $G^g + G^s = G$ for each period and locality.

2.2 Instrumental variables research design

Military spending is potentially endogenous due to political or economic factors. The observed changes in military spending may respond to unobserved local shocks also affecting the outcome variables. For example, private firms in locations with higher military contracts could have exerted more effort trough lobbying to win those contracts. As long as firm-specific shocks affect the outcome variables, our estimate may be biased.

To obtain casual estimates of the fiscal multipliers, we follow the instrumental variable research design proposed by Nakamura and Steinsson (2014). This approach consists of constructing a shift-share instrument in the spirit of Bartik (1991). The instrument is obtained by combining the nationwide changes in military procurement with a measure of comparative advantage that certain localities have in obtaining contracts. The instrument, $Z_{l,t+h}$, is defined as:

$$Z_{l,t+h} = s_l \frac{G_{t+k} - G_{t-1}}{Y_{l,t-1}} \tag{3}$$

where s_l is the average share of spending in locality l and captures the comparative advantage of certain locality of receiving military contracts. Since the empirical specifications are in changes and we use locality fixed effects, the conditional variation that is provided by the instrument comes from national-level changes in military spending, which is plausibly exogenous to local shocks that may affect the outcome variables.

The instruments for equation (2) are defined as:

$$Z^g_{l,t+h} = s^g_l \frac{G^g_{t+k} - G^g_{t-1}}{Y_{l,t-1}}; \quad Z^s_{l,t+h} = s^s_l \frac{G^s_{t+k} - G^s_{t-1}}{Y_{l,t-1}} \tag{4}$$

with Z^g being the instrument for changes in goods spending, and Z^s the instrument for changes in services spending. s^g_l and s^s_l measure the shares of spending in goods and services capture by location l, respectively.

2.3 Identification assumptions and threats

Although the previous instrumental variables research design has become the gold standard in the literature of cross-sectional fiscal multipliers, it is not free of identification threats that may bias our coefficients. Our identification assumption is that, conditional on the MSA and time fixed effects, the instrument is not systematically associated with any unobserved political or economic characteristics that may explain the outcomes of interest. This identification strategy may suffer from three main identification threats.

The first identification threat is related to measurement error because of the outsourcing of military procurement. A contract is assigned to its place of performance, defined as where the product is assembled or processed. Suppose sub-contractors outside the location of interest do the intermediate steps of the production. In that case, we would geographically misallocate the part of the outsourced military spending. This measurement error will create an attenuation bias in our estimates of the local fiscal multiplier, and therefore we should interpret them as a lower bound of the national fiscal multiplier. For the purpose of our analysis, a more serious threat in interpreting the relative size of the product-spending multipliers would come if the measurement error is systematically correlated with a specific spending category. As physical goods are more tradeable than services, it could be the measurement error is correlated with spending in goods. We rule out the importance of this measurement error in section 6.1.

The second identification threat is the potential presence of geographical spillovers. This would imply a violation of the stable unit treatment value assumption. That could be the case when defense spending in a specific MSA affects neighboring localities who did not receive any contracts. The importance of the bias in our estimates depends on the direction of the geographic spillovers. When the demand for final-consumption or intermediate goods increase in the neighboring localities due to input-output linkages, the spillovers are positive and our estimates would suffer from a downward bias. When the increase in spending affects factor prices and the allocation of production factors across MSAs,[6] the spillovers could be negative and our estimates would suffer from an upward bias. In section 6.1, we provide evidence that our estimates are downward bias, and they can be considered as a lower bound.

The third concern is related to a recent econometric literature on the drawbacks of shift-share instruments (Goldsmith-Pinkham et al., 2020; Borusyak et al., 2022; Adao et al., 2019). Our instrument consists in a shift-share instrumental variable with several shifts (e.g., as many as $G_{t+k} - G_{t-1}$ differences can be computed in the data), and one single share for each MSA (e.g., the time-invariant measure of comparative advantage). Borusyak et al. (2022) show the shift-share instrumental variable requires either the shares or the shifts to be uncorrelated with unobserved characteristics that may affect the local outcomes. In our view, the shifts are a reasonable source of quasi-random variation because military buildups respond to international geopolitical events rather than to unobserved factors such as automation, trade competition, or national fiscal policy that could have a heterogeneous impact across MSAs.

The exogeneity of the shifts assures the validity of the instrument, even when MSAs with a high comparative advantage in attracting procurement contracts are in a different economic trend than MSAs with a low comparative advantage. As both outcomes and instruments are measured in changes, our research design purges the instrument from the potential correlation between the shares and the unobserved local economic trends. Furthermore, the locality fixed effects absorb any MSA-specific secular trend. Thus, the variation provided by the instrument consists of deviations in spending from its long-term trend, which is plausibly exogenous to local economic characteristics.

[6]For example, when spending attracts workers from non-treated localities. Chodorow-Reich (2019) shows that limited factor mobility make the negative spillovers small.

3 Data

3.1 Military Spending

We assemble new data on military procurement contracts awarded by the U.S. Department of Defense (DoD). Our data have unique advantages compared to previous studies (Nakamura and Steinsson, 2014; Dupor and Guerrero, 2017; Auerbach et al., 2020). We collect and harmonize military procurement contracts data from two sources: National Archives and Records Administration (NARA) for the period 1966-2006, and USASpending.gov for the period 2007-2019.[7] The data from both sources are based on DD-350 and DD-1057 military procurement forms that accounts for about 96% of contracts awarded by the DoD.[8]

The data contain detailed information including the contract identification, the dates of action and completion, the transaction value, the location where the contract is performed, and the Product Service Codes (PSCs) or the Federal Supply Codes (FSCs). We construct the military spending in the following way. We define the year in which the government spending occurs as the year of the signature date that is the date when a contract is either awarded or modified.[9] The modifications of existing contracts could consist in downward revisions of the contract obligation. These modifications are reported as negative entries.[10] We follow Auerbach et al. (2020) and consider contracts with obligations and de-obligations with magnitudes within 0.5% of each other to be null and void. The contract dollar value is reported in nominal term. For comparability over time, we convert the nominal transaction value into real values by using the the US Bureau of Labor Statistics' Consumer Price Index.[11]

The Product Service Codes and the Federal Supply Codes play a crucial role in classifying the fiscal spending in either services or goods spending. These codes consist of a 4-digit alphanumeric classification representing the type of deliverable requested by the contract. Although some codes have been added or deleted over time, the existing codes refer to the same deliverable starting from 1979. The Product Service Codes refer to a service as the deliverable, and the first digit is a letter. The Federal Supply Codes request goods as deliverables, and the first digit is a number. Thus, if the first digit is a letter, we classify that spending as "Spending in Services." If the first digit is a number, we classify that spending as "Spending in Goods." Table A.2 in Appendix A.2 reports the major product codes included in the two categories.

Finally, the identification strategy exploits the geographic variation in military spending. As standard in the literature, the geographic allocation is based on the location where the tasks of the contracts are performed. The detailed location information permits us to geolocate contracts in narrow geographic areas. The available information differs between our two sources of data. In the USASpending, we know the city

[7] The data from USASpending.gov are available from 2001. The period 2001-2006 validates the quality of the NARA's data.

[8] Appendix A.1 tests the similarity of our data to the ones used in previous studies (Nakamura and Steinsson, 2014; Dupor and Guerrero, 2017; Demyanyk et al., 2019). We find our universe of military spending is very similar to one from previous studies in both the aggregate level and the geographic distribution.

[9] The government fiscal year has been defined from October 1^{st} to September 30^{th} since 1976. The mismatch between the fiscal year of the government and the calendar year could cause a time inconsistency between the the military spending and the other economic variables. Thus, we use the calendar year as reference year.

[10] For most years, the contract value is reported as an alphanumeric code. The last digit identifies whether the contract is an obligation or a de-obligation. We use the contract dictionaries to decode the alphanumeric strings into numeric values.

[11] The base period is between 1982 and 1984.

and the zip code of the performing firm.[12] We use these two pieces of information to identify the county in which a firm is operating. In NARA, the county in which a firm is performing the contract tasks is reported. We use the spatial crosswalks provided by the National Bureau of Economic Research to aggregate the county-level military contracts into Core-Based Statistical Areas.

3.2 Economic Outcomes

We use four outcome variables to proxy changes in the on economic activities after a local spending shock. The main dependent variable is the employment, and consists in the headcount of employed workers. The next three variables are the personal income that measures the income households get from salary and wages, Social Security and other government programs, dividends and interest, business ownership, and other sources, and two components of the personal income—salary and wages, and dividends and interest—that capture labor income and firm profits, respectively. These data are collected annually from the Bureau of Economic Analysis at MSA-level.[13] Personal income, salary and wages, and dividends are in nominal terms, and, for consistency with the other monetary variables, we convert them in real terms.

3.3 Sample Construction

We construct the final sample by applying two sets of filters. We drop from the universe of DoD procurement contracts the ones with missing information in at least one of these dimensions: the locality in which the contract is performed, the year in which the contract is signed, the product code,[14] or the industry in which the contractor operates.[15] We apply these filters because the absence of any of these variables prevent us from correctly assigning contracts to one of the spending categories. We restrict the sample to 2019 to avoid the effect of the COVID-19 pandemic on our estimates.[16]

The second set of filters are applied to the aggregated spending by MSA. These filters are used to minimize the effect of outliers or misreported values on our estimates. We exclude MSAs with incomplete histories in any outcome variable (personal income, salary and wages, dividends, or employment) or with growth rates between two consecutive periods greater than 100% or smaller than -50%. We remove MSAs with an average population smaller than $50,000$ inhabitants. We only include MSAs with non-negative ratios of aggregate military spending to personal income smaller than 1.5. Finally, as in Auerbach et al. (2020) and Demyanyk et al. (2019), the analysis is carried out at locality aggregated level rather than per capita.

The final sample of procurement contracts are allocated to 334 MSAs from 1979 to 2019. The data filters have a minor impact on the representativeness of the procurement contracts. The final sample includes 94%

[12] Following Demyanyk et al. (2019), if we know that a contract has been performed in the US, but we do not know the exact location where it was performed, we assign the location of the contract recipient as the performance location. Notice that, differently from Demyanyk et al. (2019), we adjust a marginal share of contracts. That's the case because we locate contracts not only using the postal code, but also the city. There are contracts for which information about the postal code is missing, but the information about the city is not missing.

[13] Both employment and employee earnings are derived from the Bureau of Labor Statistics' Quarterly Census of Employment and Wages (QCEW).

[14] Our analysis starts from 1979 because the product codes before 1979 are not consistent with the ones after that date.

[15] For a small share of contracts, we impute the industry in which the contractor operates by using the product codes. We exclude contracts for which the industry is missing after the imputation. Appendix A.4 describes the imputation procedure.

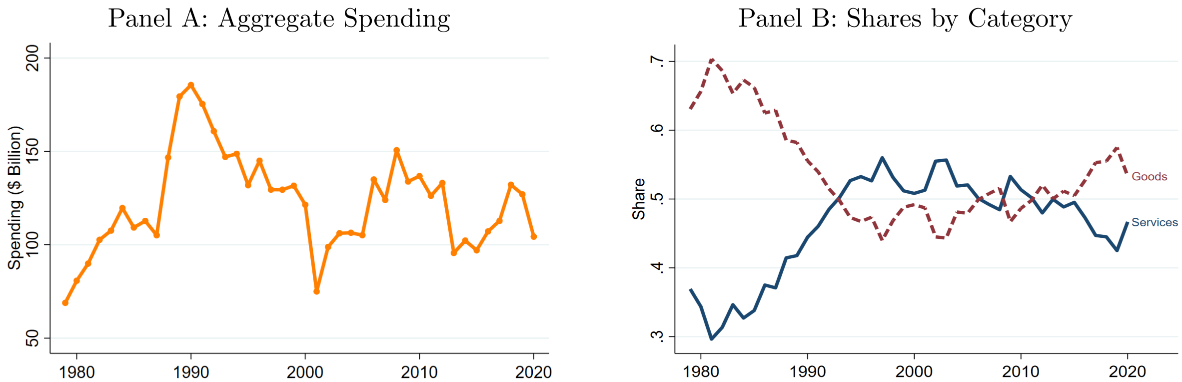
[16] As a robustness check, we have also replicated our results including 2020. Results remain unaffected.

of both the total number of contracts and the aggregate spending. These shares are high for both types of spending. For spending in goods, the share of contract and spending in the sample accounts for about 95% of the total spending in goods. For spending in services, we include 85% of the contracts worth 92% of the total spending in services.[17]

4 Descriptive Statistics

Figure 1 reports the spending over the period 1979 − 2019. Panel A shows the evolution of aggregate military spending. There are three sharp rises in the spending. The first increase is in the 1980s as a consequence of the Reagan military buildup, the second is in the early 2000s due to the Afghani and Iraqi wars, and the last is in the recent years due to the escalations in military buildups with Russia and China.

Figure 1: Military Spending



Notes: The national level statistics are calculated by aggregating the microdata on military procurement contracts available from NARA and USASpending. The face value in the DoD procurement contracts is deflated by the CPI. The classification of the spending into goods and services is based on the Federal product classification. The figures are computed using the 334 MSAs in the final sample.

Panel B reports the shares of military spending by category. Between the 1980s and the 1990s we document a reallocation of the federal government spending from goods to services. At the beginning of the 1980s, the spending in goods was more than 70% of the value of the military procurement contracts. From the beginning of the 1990s onwards, the spending has been split between goods and services with similar shares. The average shares for the two categories of spending are reported in the first row of Table 1.

The second row reports some basic descriptive statistics on the contract characteristics by category. The DoD subscribes most of its procurement contracts (91%) to purchase goods, and only a small share of 9% is directed to the acquisition of services. This set of results implies the average value of a contract providing services is about three times the average contract used to purchase goods.

The last two rows of Table 1 explore the distributional characteristics of spending within categories. The distribution of contracts for the purchase of goods has a significantly fatter right-tail than the distribution for spending in services. In the case of spending for goods, the contract at the top decile of the distribution

[17]Descriptive statistics and regression estimates refer to the military procurement contracts and MSAs in the final sample.

Table 1: Descriptive Statistics: Contracts by Components

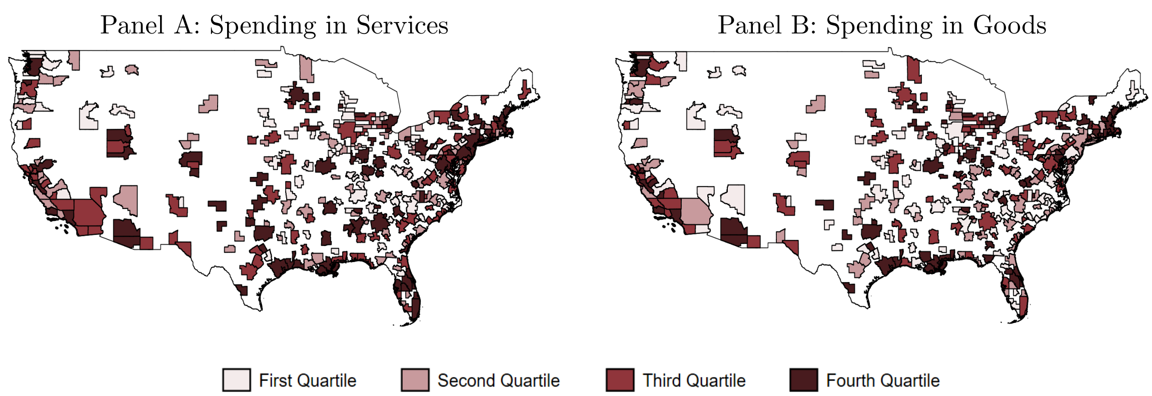
	All	Goods	Services
Share of Spending (%)	-	54	46
Share of Contracts (%)	-	91	9
Average Value (Thousand USD)	143	99	293
90%-to-50%-percentile Ratio	52	46	11
Share of Spending for Top 10% (%)	97	99	84

Note: The national level statistics are calculated by aggregating the microdata on military procurement contracts available from NARA and USASpending.gov. The face value in the DoD procurement contracts is deflated by the CPI. The classification of the spending into goods and services is based on the Federal product classification. The statistics are computed using the 332 MSAs in the final sample.

is over 40 times greater than the median contract. The 90%-to-50%-percentile ratio is smaller for spending in services, with the contract at the top decile to be 11 times greater than the median contract.

The last row of Table 1 shows the share of spending allocated to contracts in the top decile. The top 10% of contracts awarded to purchase goods accounts for virtually the total goods spending. This result implies that although the majority of contracts purchases goods, the largest bulk of the spending is allocated to a relatively small subset. The share of services spending allocated to the top decile is around 85% implying a more equal allocation of spending across contracts.

Figure 2: Military Spending - Geographic Distribution



Notes: The quartile to which a MSA belongs is assigned based on the average military spending in real terms that the MSA receives over the period $1979-2020$.The classification of the spending into goods and services is based on the Federal product classification. The figures are computed using the 332 MSAs in the final sample.

As the identification strategy exploits the cross-sectional variation in spending, Figure 2 explores the geographic heterogeneity in the allocation of spending across MSAs. The figure shows the quartile to which a MSA belongs based on the average value of the military spending that it has receive over the period $1979-2019$. There are two main results to highlight. First, military spending is unequally distributed across MSAs. The top 30 MSAs in terms of awarded spending accounts for more than 30% and 45% of the spending in goods and services, respectively. Most of the MSAs that receive the largest share of DoD spending are located along the two coastal regions, and the Midwest. Second, the geographic allocation of spending differs between the two categories. Only 35% of MSAs are in the same quartile in both categories. The heat-maps suggest most MSAs in the Midwest are in the top two quartiles of the spending in goods, and only a few of these MSAs are ranked as high in the distribution of services spending. This result mirrors the geographic economic structure of the US.[18]

These descriptive statistics provide evidence that although the DoD signs a large number of contracts, the largest share of government procurement spending is captured by a handful of contracts. Furthermore, these contracts are not equally distributed across the MSAs. These findings imply the distribution of spending at both contract and MSA-level is skewed to the right. Our findings are complementary to Cox et al. (2020). They show spending is granular and concentrated among a few firms. Their analysis is based on the contract recipients. The granularity in recipients does not necessary imply a granularity in firms actually performing the tasks of a contract. It could happen contract recipients allocate several tasks of a contract to different performers evenly located in the national territory. If that were the case, we would not observe a geographic concentration in the spending based on the place of performance. Our results show the opposite suggesting, in addition to a granularity in contract recipients, also a granularity in contract performers.[19] Similarly to Cox et al. (2020), we also show a substantial variation in the range of contract values, and we emphasize significant differences in the distributional features of contract size across categories of spending.

5 Main Results

5.1 Local Fiscal Multiplier by Category

Table 2 reports the estimates of the employment local fiscal multiplier at different horizons after a local shock in the aggregate spending.[20] The estimates suggest a positive and significant employment multiplier.[21] Our estimates for the aggregate spending are reasonably close to the effects estimated by other authors. Specifically, Auerbach et al. (2020) report the on-impact employment multiplier for local spending shocks at CBSA-level to be 0.186. This estimate is very close to ours (0.139). To the following horizons the discrepancy between the two sets of estimates increases, with ours being greater than theirs. This discrepancy is due to two sample differences. First, our period is from $1979 - 2019$, while theirs starts in the early 2000s. There

[18]The Midwest has the highest concentration of production occupations with a average employment share in production jobs almost 50% higher than the US average.

[19]We cannot test directly this granularity in contract performers because the data do not report the name of the firm that performs a contract. We only observe the place of performance.

[20]As output by MSA is not available before the 2000s, consistently with previous studies we report the employment multiplier rather than the output multiplier. Focusing on employment rather than output effects of spending is a common practice in the applied fiscal literature. For example, all studies that have exploited the variation in some components of the American Recovery and Reinvestment Act report the employment multiplier (Chodorow-Reich et al., 2012; Conley and Dupor, 2013; Dupor and Mehkari, 2016; Dupor and McCrory, 2018). Intuitively, a positive effect on the employment multiplier would imply a higher demand for workers, a higher production, and a positive output multiplier. The quantification of the exact size of the relationship between the two multipliers is less straightforward. Chodorow-Reich (2019) shows that for the United States the translation from an employment to output multiplier is to divide output per worker by the cost per job.

[21]We test the validity of the instruments, and we find the first-stage F-statistic are generally larger than 10.

is a large literature arguing the national fiscal multiplier has decreased over time (Blanchard and Perotti, 2002).[22] Thus, one should expect lower estimates by restricting the period from 2001. Second, we exclude micropolitan statistical areas. In smaller geographic areas, like micropolitan statistical areas, it is more likely that part of the spending spills into other geographic areas attenuating the estimates of local fiscal shocks.

Table 2: Employment Local Fiscal Multiplier - Aggregate Spending

	(1) t	(2) t+1	(3) t+2	(4) t+3	(5) t+4
Spending	0.139***	0.414***	0.453***	0.550**	0.499***
	(0.051)	(0.153)	(0.153)	(0.236)	(0.192)

Notes: The estimates are computed from equation (1). The instruments are computed as in equation (3). The balanced panel consists of $12,024$ observations, and includes 334 MSAs for the period $1979-2019$. The instrument is calculated over the same period. Standard errors are clustered by MSA. The symbols *, **, and *** represent the 10%, 5%, and 1% significance levels, respectively. All regressions include MSA and year fixed effects.

Now, we turn to the core results of the paper. We explore the different effects of shocks in two types of government spending —spending in goods and spending in services. Table B.1 in Appendix B reports the exact point estimates. Figure 3 plots the estimated effects. Panels A and B report the employment local multipliers for shocks to goods and services spending, respectively. We observe a marked and striking difference between the two patterns. On the one hand, the estimates of the employment multiplier after a shock in the services spending are positive, and statistically significant at the 1% significance level. The effects are sizeable. Indeed, a percent increase in services spending at local-level measured in personal income percentage points generates an increase in employment between 0.36% on-impact and 1.45% three periods after the shock. On the other hand, shocks in goods spending have small effects on the employment multiplier. The estimates range between 0.059% and 0.186%, and they are statistically non-significant.[23]

The remaining panels of Figure 3 explore the effects of the two types of shocks on other three outcomes variables—personal income, salary and wages, and dividends. These variables compose the household's budget constraint, and shed some light on the changes in the sources of income available for consumption. Any government shock has a direct effect. The increase in the demand for products leads firms hiring more workers. This mechanism explains part of the increase in employment we have documented. The large reaction to government shocks comes from the indirect effect.[24] As firms hire more workers (or similarly, workers work more hours), the labor income increases, and households can increase their consumption. The additional increase in demand for products increases further the labor demand, leading to a multiplier effect. Our results suggest this virtuous circle only activates after shocks to services spending.

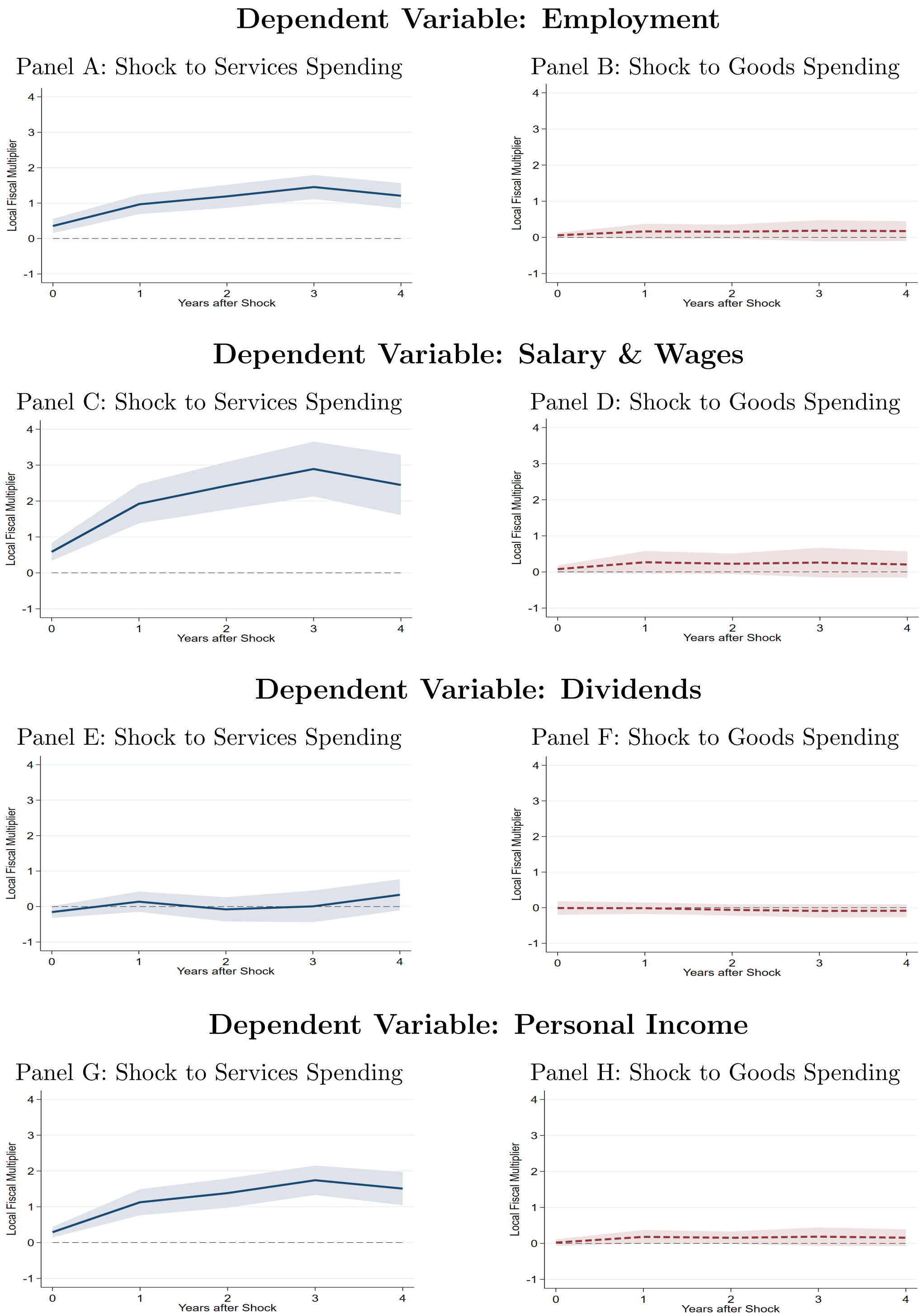
Panel C and Panel D of Figure 3 report the behavior of labor income after the two types of shocks.

[22] We are not aware of any studies that show this decline using local variation in government spending.

[23] We reject the null hypothesis that the two estimates are equal at least at the 5% significance level.

[24] Alonso (2017) and Bouakez et al. (2020) show the direct effect of government spending shocks is relatively small. Most of the aggregate impact comes from the indirect effect consisting in the responses of households to increases in their income.

Figure 3: Local Fiscal Multipliers - Spending by Category



Notes: All panels report the estimates from equation (2) for different outcome variables. The instruments are computed as in equation (4). The balanced panel consists of 12,024 observations, and includes 334 MSAs for the period 1979 – 2019. The instrument is calculated over the same period. Standard errors are clustered by MSA. The shaded areas represent the 90% confidence intervals. For graphical comparability, the y-axis are on the same scale.

In Panel C, one can notice a sharp and sizeable increase in wages and salary after a change in services spending. We quantify a 1% increase in services spending generates a rise in labor income up to 2.5% within four periods after the shock. Panel D shows a very different response of wages to goods spending shocks. Labor income only slightly increases after a goods spending shock.

As the final owners of firms are the households, changes in their available budget could also come from changes in the dividends paid by firms. Panel E and Panel F show similar picture for the two types of spending. Firms do not adjust the amount of dividends paid to households after spending shocks. The cumulative effect of fiscal spending on the household's personal income is reported in the two bottom panels of Figure 3. The responses of personal income closely track the responses to labor income, with the only difference the size of the effect is smaller as changes in personal income are given by a combination of changes in labor income and dividends.

5.2 Robustness

One concern is that the specification ignores the dynamics of spending and employment for the years in between the period of study. For example, when the change in spending is measured as the difference in spending between $t + k$ and $t - 1$, one ignores the changes in the years $t \in \{t-1, t+k\}$. To circumvent this problem, we follow Ramey and Zubairy (2018), and estimate a specification where the outcome and spending variables are defined as the cumulative change between $t - 1$ and $t + k$. In other words, the dependent variables are defined as

$$\sum_{h=0}^{k} \frac{v_{l,t+h} - v_{l,t-1}}{v_{l,t-1}},$$

and, similarly, the spending variables are computed as

$$\sum_{h=0}^{k} \frac{G^i_{l,t+h} - G^i_{l,t-1}}{Y_{l,t-1}} \quad \forall i \in \{s, g\}.$$

Results are presented in Table B.2 in Appendix B. All coefficients are quantitative similar to the ones presented in our main specification, i.e. services spending generates large employment multipliers, while the estimated effects of a shock to goods spending are small.

A second concern is related to the timing of the spending. We observe the start and end date of a contract, but we do not observe the dates at which the government disbursements actually occur. In the benchmark specification, we use the year in which a contract has been signed as the year in which the spending occurs. We follow Auerbach et al. (2020), and we construct a proxy for outlays by allocating the value of a contract equally over its duration. The results are reported in Table B.3 in Appendix B.[25] The employment multiplier is positive and significant only after the occurrence of shocks in the services spending.[26]

[25]The estimation period starts in 1989 because the completion date needed to compute the flow of spending is only available starting from that year.

[26]The only difference from the benchmark specification consists in the results for the dividends. Using the outlays rather the contract values, we document an effect of services spending on dividends.

5.3 Mechanism: The role of labor intensity

We argue the differences in the local fiscal multipliers generated by spending in goods and services are due to the transmission of the shock to the household's budget constraint. We show the labor intensity of the spending is an important determinant to quantify the effect of fiscal interventions. Intuitively, in labor-intensive industries a larger share of the government spending passes through the households. Public spending pushes up the demand for products, firms hire more workers to match the demand, labor income and personal income raise, enabling households to spend more in consumption and further pushing up demand for products. In non-labor-intensive industries this pass-through mechanism is alleviated because labor demand reacts by less after a government spending shock.

Our dataset does not contain any information on the amount of inputs used in the production. Thus, we use an alternative strategy to assess the labor intensity of each contract. We collect annual data from the BEA on value added and employees' compensation by industry.[27] We then compute labor intensity as the contribution of the employees' compensation to the value added. Finally, we assign the constructed measure of labor intensity to each contract based on the industry to which the contractor belongs. Table A.3 reports the classification of the available industries separated between labor- and non-labor-intensive. Our classification matches the common sense. Indeed, industries as healthcare, education, hospitality and food service are classified as labor-intensive, while manufacturing and retail as non-labor-intensive.

Imperfect competition may also be a confounder in our classification. Firms with more monopsony power may pay less their employees. As our industry classification is based on the employees' compensation, industries with firms that intensively use labor but pay their employees little due to their monopsony power would be misclassified. As we cannot disentangle labor intensity from monopsony power, our classification may be driven by the latter rather than the former. To shed some light about the importance of imperfect competition as a confounder, we explore the relationship between labor intensity and measures of concentration. As we do not have direct measures of monopsony power, we use employment concentration shares as proxies for the degree of imperfect competition.[28] Specifically, we regress labor intensity against the measures of concentration after having absorbed the contribution of years and industries fixed effects. The result are showed in Table B.4 in Appendix B. We document a weak negative relationship between employment concentration and labor intensity at industry-level. We interpret these findings as evidence that monopsony power is not an important confounder of our industry classification.

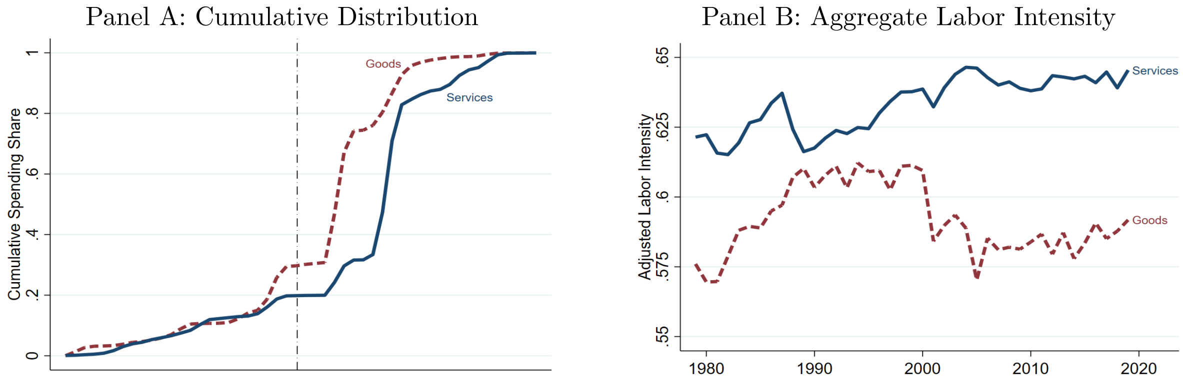
Figure 4 shows the distribution of spending by components and labor intensity. Panel A shows the cumulative distribution of contracts sorting them by labor intensity. The black dashed line marks the switch between labor- and non-labor-intensive industries. The largest share of government spending is directed to labor-intensive industries. The shares of spending allocated to contracts with contractors in labor-intensive industries significantly differ between spending in goods and services. The share of goods spending allocated

[27] The data contain 66 NAICS codes starting from 1997. Although valued added and its components are available from the BEA, we use an imputation strategy described in Appendix A.4 to assign industry to contracts awarded before 1997.

[28] The Census Bureau calculates employment concentration shares by industry for the years 2002 and 2017.

to non-labor-intensive industries is 50% higher than the share allocated from spending in services (30% vs. 20%). Furthermore, not only the cumulative distribution at the cutoff is higher for goods spending, but also the cumulative distribution of services spending is shifted to the right, implying a larger share of contracts with services spending is directed to labor-intensive industries.

Figure 4: Military Spending and Labor Intensity by Category



Notes: The classification of the spending into goods and services is based on the Federal product classification. The figures are computed using the 334 MSAs in the final sample. The dashed black line represents the separation between high and low labor-intensive industries based on the classification in Table A.3.

To capture both the distribution of contracts across industries and the industry-specific labor intensity, we construct a measure of aggregate labor intensity for each category of spending. This measure is the sum of the average industry-specific labor intensity weighted by the share of spending in a category directed to that industry.[29] Formally, the aggregate labor intensity measure is computed as follows:

$$AggLI_{i,t} = \sum_{j \in \mathcal{J}} \omega_{i,j,t} \overline{LI}_j \quad \forall i \in \{s, g\} \tag{5}$$

with the weight, $\omega_{i,j,t}$, is equal to the share of spending in a category i directed to an industry j over the total spending in category i:

$$\omega_{i,j,t} = \frac{G^i_{j,t}}{G^i_t} \quad \forall i \in \{s, g\}$$

Panel B of Figure 4 reports the results for the aggregate labor intensity. The aggregate labor intensity is higher for services spending than goods spending. Apart from the 1990s, in which the gap narrowed, the aggregate labor intensity measure is about 10% higher for services spending than goods spending.[30]

These results point out the composition of the spending in the two categories differs in terms of the labor intensity of the industries that receives the contracts, and that labor intensity could play an important role in explaining the differences in the findings documented in section 5.1. We assess the role of labor intensity

[29]We use the average industry-specific labor intensity because labor intensity data are only available from 1997. We replicate this figure also using time-varying industry-specific labor intensity for the period 1997 − 2019. Results are consistent with the ones showed in Panel B.

[30]We cannot assess how sizeable these differences are. We believe we need a structural model to quantify the general equilibrium effect of these differences. This analysis is left for future work.

in determining the differences in the local fiscal multipliers across categories of spending by considering separately the components of spending for each category that goes to labor- and non-labor-intensive industries. We formally test this hypothesis by implementing the following specification:

$$\frac{v_{l,t+k} - v_{l,t-1}}{v_{l,t-1}} = \sum_{i \in \{s,g\}} \sum_{p \in \{L,N\}} \gamma_{i,p}^{k} \frac{G_{l,t+k}^{i,p} - G_{l,t-1}^{i,p}}{Y_{l,t-1}} + \alpha_l^k + \delta_{t+k} + \varepsilon_{l,t+k} \tag{6}$$

with the independent variables instrumented by

$$Z_{l,t+h}^{i,p} = s_l^{i,p} \frac{G_{t+k}^{i,p} - G_{t-1}^{i,p}}{Y_{l,t-1}} \quad \forall i \in \{s,g\} \text{ and } \forall p \in \{L,N\}, \tag{7}$$

where $G^{i,p}$ is the spending in either goods or services directed to either labor- or non-labor-intensive industries. Intuitively, if labor intensity plays a central role in determining the fiscal multiplier, we should expect large and positive effects coming from spending in labor-intensive industries independently from the type of spending. The estimates for employment, and salary and wages in Table 3 provide support for the hypothesis.[31] Different effects by category of spending are linked to the amplification of shocks through the intensity of labor used in the production of the different products required by the types of spending.

Spending in both services and goods directed to labor-intensive industries generate positive and strongly significant effects on employment. The first panel of Table 3 shows the estimates for the employment multiplier for services spending ranges between 0.36% and 1.5%. Although the employment multiplier coming from goods spending is also positive and significant, the size of this effect is significantly smaller, with estimates between 0.07% and 0.54%.[32] The effect of spending in non-labor-intensive industries is negative for both types of spending, although only significant for goods spending. One potential explanation for the negative effect of spending in non-labor-intensive industries may be the crowding-out of private consumption. That would be the case when an increase in government demand increases prices rather than production.[33] It is worth to notice the estimates for goods spending in labor and non-labor-intensive industries are comparable in magnitude, but in opposite directions. Thus, the small multiplier effect for the aggregate goods spending is the result of these two opposite effects canceling out. The second panel of Table 3 reports the estimated effects for labor income. The results are consistent with the findings previously discussed. Spending directed to labor-intensive industries generates increases in labor income, while the effects from spending in non-labor-intensive industries are either negative or non-significant.

Finally, the size of the local fiscal multiplier also depends on the connection across industries though their input-output linkages. If an increase in demand occurs in an industry that relies on inputs from many other connected industries, then the effect of the increase in demand will propagate to a large number of

[31]As showed in section 5.1, changes in personal income are driven by changes in salary and wages rather than dividends. Thus, we only report the results for salary and wages. Results for dividends and personal income can be provided upon request.

[32]We test the null hypothesis that the two coefficients are the same, and we rejected this hypothesis at any horizon.

[33]There is no consensus whether government spending crowds-in or crowds-out private consumption. There is advocates for both sides, namely, a crowding-out effect (Barro, 1981) and a crowding-in effect (Mountford and Uhlig, 2009).

Table 3: Local Fiscal Multipliers - Spending by Category and Labor Intensity

	(1) t	(2) t+1	(3) t+2	(4) t+3	(5) t+4
	Dependent Variable: Employment				
Services Spending - High Labor Intensity	0.365**	0.978***	1.183***	1.554***	1.276***
	(0.141)	(0.172)	(0.257)	(0.302)	(0.307)
Services Spending - Low Labor Intensity	0.016	-0.164	-0.259	-0.839	-0.868
	(0.574)	(0.736)	(0.881)	(0.993)	(0.951)
Goods Spending - High Labor Intensity	0.073**	0.283*	0.302**	0.537*	0.476**
	(0.032)	(0.150)	(0.148)	(0.309)	(0.240)
Goods Spending - Low Labor Intensity	-0.062	-0.190*	-0.175*	-0.331**	-0.336**
	(0.074)	(0.113)	(0.090)	(0.148)	(0.150)
	Dependent Variable: Salary & Wages				
Services Spending - High Labor Intensity	0.504***	1.828***	2.278***	2.968***	2.483***
	(0.161)	(0.345)	(0.474)	(0.572)	(0.625)
Services Spending - Low Labor Intensity	0.752	0.958	1.207	0.098	0.063
	(0.883)	(1.196)	(1.415)	(1.482)	(1.392)
Goods Spending - High Labor Intensity	0.114**	0.495**	0.464**	0.795**	0.615**
	(0.047)	(0.207)	(0.189)	(0.391)	(0.286)
Goods Spending - Low Labor Intensity	-0.102	-0.334*	-0.259*	-0.474**	-0.461**
	(0.115)	(0.179)	(0.140)	(0.202)	(0.200)

Notes: Both panels report the estimates from equation (6). The instruments are computed as in equation (7). The balanced panel consists of 12,024 observations, and includes 334 MSAs for the period 1979 − 2019. The instrument is calculated over the same period. Standard errors are clustered by MSA. The symbols *, **, and *** represent the 10%, 5%, and 1% significance levels, respectively. All regressions include MSA and year fixed effects.

industries.[34] The difference in the size of the estimates between spending in goods and services in labor-intensive industries could be partially explained by their production network.

Figure 5 is informative about the input linkages across industries.[35] The x-axis reports the number of connected industries to a specif sector, and it is a measure of the extensive margin.[36] The y-axis reports the average direct requirement coming from the other industries, and it represents a measure of the intensive margin.[37] The red circles identify labor-intensive industries, while the blues triangles the non-labor-intensive industries. The difference between Panel A and B is in the size of the markers that represent the share in each type of spending allocated to a specific industry.[38] First of all, labor-intensive industries require inputs

[34]Bouakez et al. (2021) also shows network centrality is the primary amplification mechanism of a fiscal shock.

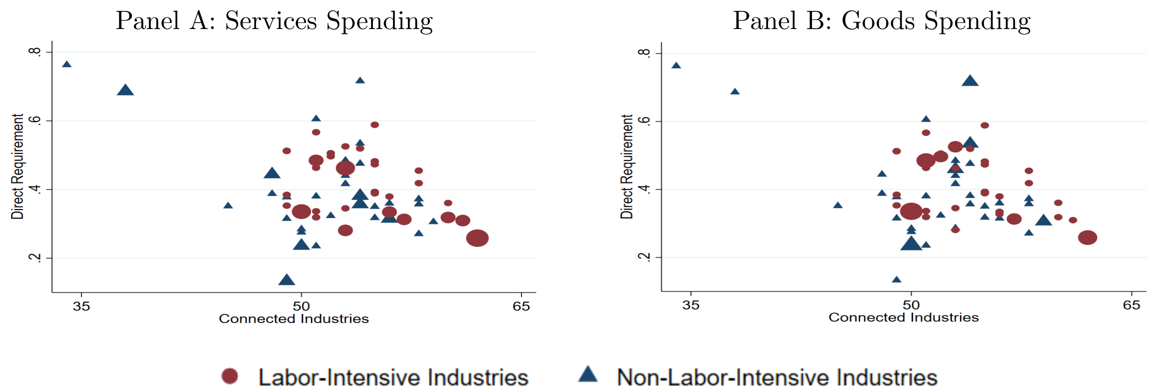
[35]We use the Direct Requirement tables of the Input-Output accounts to derive the input linkages across industries.

[36]An industry i is connected to an industry j if j receives a flow of inputs from i in each period between 1997 and 2019.

[37]A direct requirement consists of the value of input of commodities that an industry requires to produce a dollar of output. Direct requirements are expressed in terms of the industry's output value.

[38]For graphical purposes, we classify the spending shares into four groups: smaller or equal to 0.01, between 0.01 and 0.05, between 0.05 and 0.15, greater than 0.15.

Figure 5: Input-Output Linkages



Notes: The linkages across industries are computed using the Direct Requirement tables of the Input-Output accounts. The x-axis reports the number of connected industries to a specific sector defined as the industries that provide a stable flow of inputs in each period between 1997 and 2019. The y-axis reports the average direct requirement. The red circle identify high labor-intensive industries, while the blues triangles the low labor-intensive industries. The size of the circles represent the share of spending allocated to a specific industry, and they are computed using the contracts included in the final sample.

from a larger number of industries (more connected industries), and have a stronger dependence from their input providers (higher direct requirement).[39]

The comparison between Panel A and B shows the spending in goods and services is allocated differently across industries based on their production network, and these differences mostly come from the allocation of spending to labor-intensive industries. Panel A highlights a large share of services spending is directed to contractors in industries that receive inputs from many other industries (graphically, the largest red circles are towards the right). Panel A also emphasizes a large share of services spending is captured by industries with a direct requirement between 0.2 and 0.4. Overall, industries that receive services spending are connected to many other industries, but their input requirement from those industries is moderate.

Panel B indicates a large share of goods spending is directed to industries with relatively low input linkages. In terms of direct requirement, the picture for goods spending is very similar to the one for services spending. Industries that receive a large share of the goods spending require a moderate amount of inputs from their connected industries. The differences in the size of the estimates for goods and services spending directed to labor-intensive industries could be explained by the input-output linkages. A government spending directed to a more central industry in the production network generates an amplification of the shock through responses in a large number of industries linked via the input-output relationships. Our suggestive evidence also points out the extensive margin (the number of connected industries) may play a more important role than the intensive margin (the share of input dependence).

[39] On average, labor-intensive industries receive direct inputs for production from 54 different industries compared to 51 for non-labor-intensive industries. The average strength of the input dependence is slightly higher for labor-intensive industries, with a direct requirement of 0.41 versus 0.39 in non-labor-intensive industries.

6 Alternative Mechanisms

This section explores alternative mechanisms or threats to identification that could explain the differences in the previous estimates by category of spending. Overall, these tests suggest that none of these alternative mechanisms seem to play, if any, a marginal role in explaining these differences.

6.1 Tradeability

Goods are more tradeable than services.[40] If the production of goods occurs in the neighboring locations, one could observe small effects of goods spending on the local fiscal multiplier. If that were the case, differences in the fiscal multipliers would reflect geographic spillovers rather than the actual nature of the spending, and these spillovers would be more sizable for more tradeable products as goods.

This explanation could be relevant in our context due to the well-documented issues in correctly allocating government spending to localities. Our data only contain records on prime contracts and they do not reflect the amount of subcontracting for basic and intermediate materials and components. A contract is assigned to its place of performance defined as the place where the product is assembled or processed. If the intermediate steps of the production are done by sub-contractors outside the location of interest, we would geographically misallocate part of the spending. In addition to that, the definition of the place of performance slightly varies across categories of products.[41] The measurement errors in allocating contracts to MSAs could affect the previous findings. In this sub-section, we show, even accounting for these factors, previous findings do not change. These results imply the heterogeneous reactions of employment to different types of spending does not depend of products tradeability or geographic spillovers.

We undertake two strategies to quantify the importance of the geographic allocation of contracts. The literature has extensively argued the geographic misallocation of contracts becomes less important as one moves to larger geographic aggregation. Specifically, it becomes a minor issue at state-level. Isard (1962) argues geographic disaggregation at state-level does not contain significant measurement errors. Nakamura and Steinsson (2014) use shipment data to the government from defense industries, reported by the U.S. Census Bureau from 1963 to 1983, and verify that, on average, the relationship between the prime contracts allocated to a state and the shipments from that state is one-for-one. These two studies imply that on average, all contracts allocated to a state are also performed in that state.

Our first test consists in comparing the estimates from the regressions using MSA-level data with the ones using state-level data. If the concern with tradeability and geographic allocation of spending to MSAs are not important drivers of our estimates, one should expect the state-level analysis leading to the same conclusions. The state-level estimates are in Table B.5 in Appendix B. We do not find any effect of goods

[40] The average service industry is less tradeable than the average manufacturing industry. However service tradeability has been growing over time.

[41] The location of the majority of manufacturing contracts reflect the location of the plant where the product is finally assembled or processed. The location of construction contracts corresponds to the location were the construction is performed. The location for contracts involving purchases from wholesale or other distribution firms reflects the location of the contractor's place of business. Finally, for service contracts, the location is the place where the service is performed, with the exception of transportation and communication services that report the location of the contracting firm.

spending on employment, while the effects for services spending remain sizable and significant. As output data are available at state-level sine 1979, we also report the output multiplier.[42] We find strong positive effects on output after shocks in the services spending, and no effects due to changes in the goods spending. Similarly to Demyanyk et al. (2019) and Auerbach et al. (2020), the effects are larger by using the state aggregation. The discrepancy between MSA- and state-level estimates can be attributed to within-state subcontracting. In smaller geographic areas, it is more likely that part of the spending to spill into/from other areas. The potential measurement error in subcontracting outside a MSA attenuates the estimates of the local fiscal multipliers, implying the MSA-level estimates to be lower bounds.

Still, MSAs along the state borders have strong economic interactions with other MSAs outside the state. In these cases, government spending allocated to locality l could be used for production in neighboring locations outside the state borders. Our previous test would not capture these cross-state "outflows." Our second test, instead, captures these interactions. We implement the following specification to investigate whether military spending shocks in location l have some positive effect in the neighboring locations:

$$\frac{\tilde{Y}_{l,t+k} - \tilde{Y}_{l,t-1}}{\tilde{Y}_{l,t-1}} = \tilde{\beta}_g^k \frac{G^g_{l,t+k} - G^g_{l,t-1}}{\tilde{Y}_{l,t-1}} + \tilde{\beta}_s^k \frac{G^s_{l,t+k} - G^s_{l,t-1}}{\tilde{Y}_{l,t-1}} + \tilde{\beta}_{rd}^k + \alpha_l^k + \delta_{t+k} + \varepsilon_{l,t+k}, \tag{8}$$

where $\tilde{Y}_{l,t+k}$ is the outcome variable for the neighboring locations of l. We define neighboring locations to l MSAs whose center is located within a 100 miles radius distance from the center of l.[43] Similarly as before, we instrument the main regressors with a shift-share instrument defined as

$$Z^g_{l,t+h} = s^g_l \frac{G^g_{t+k} - G^g_{t-1}}{\tilde{Y}_{l,t-1}}; \quad Z^s_{l,t+h} = s^s_l \frac{G^s_{t+k} - G^s_{t-1}}{\tilde{Y}_{l,t-1}}. \tag{9}$$

Results for employment and labor income are reported in Table B.6 in Appendix B.[44] The findings show the higher tradeability of goods cannot explain the differences in the multipliers.[45] On the one hand, the estimates point out there are no spillover effects of a shock in goods spending in locality l on the neighboring localities. On the other hand, a spending shock in services has some positive and significant "outflow" effects on neighboring locations. These results reinforce the previous conclusion that the small effects of goods spending cannot be explained by the fact that goods spending generates positive multiplier effects in neighboring locations rather than in the location that receives the contracts. The presence of some geographic spillover effects for the services spending also suggests the estimates we reported in Table B.1 are downward biased, implying they could be considered as lower bounds of the effects of government spending.

[42] Government spending is usually normalized by output. For consistency with the estimates at MSA-level, we normalize it by personal income. Thus, the magnitude of our estimates is not directly comparable with the ones reported by other studies.

[43] While Auerbach et al. (2020) consider one neighbor location within the distance with a size similar to location l, we include all neighboring MSAs within the distance.

[44] As showed in section 5.1, changes in personal income are driven by changes in salary and wages rather than dividends. Thus, we only report the results for salary and wages.

[45] We exclude 149 MSAs that do not have any neighboring locations within a radius of 100 miles. We also tested the effect of increasing the distance on the results. The main findings remain unchanged.

6.2 Productivity Gains

Government spending may improve productivity. Increases in spending may reduce the uncertainty about future profits, ease credit constraints leading to higher firms' turnover rates, or generate a faster growth of incumbents. These channels, typical in growth models with firm heterogeneity, may generate productivity improvements.[46] If different types of spending generate different firm dynamics, the differences in the fiscal multipliers could be partially explain by differences in productivity gains.

Government spending may lower entry costs, ease credit constraints, incentivize firm entry, and improve the allocation of production factors. As goods producers face usually larger entry costs than services producers, spending directed to goods producers may be not sufficient to lower the entry costs. In this scenario, one should observe the mass of firms entering the market after a shock to services spending to be higher than after a shock to goods spending. The top panel of Table B.7 reports the impact of the different types of spending on the establishment entry rate.[47]. The establishment entry rate increases a few periods after the occurrence of a shock to the services spending. Consistently with the higher entry costs hypothesis, the effects are smaller after shocks to goods spending. To dig further into this result, we explore the impacts by the size of the entrant firms. If government spending relaxes the credit constraint and lowers the entry cost, one should observe larger effects for small-size firms. Our results, indeed, show the increase in entry rates is driven by the entry of small-size establishments. As above, these effects are stronger after shocks to services spending than goods spending.

Government demand may also change firms' incentives of exiting the market, and keeps in business low-productive firms that would have otherwise left the market. If low productive firms remain in the market, the production factors would not be allocated to the most productive firms that largely contribute to the productivity growth, causing losses in aggregate productivity. The last panel of Table B.7 explores the effect of types of government spending on the establishment exit rates. The estimates suggest a decline in the exit rates of establishments after shocks to the services spending. Shocks to the goods spending have marginal effects on the establishment exit rates.

The last source of productivity growth comes from the improvement of incumbents' technology via innovation activities. We proxy innovation with patenting activities.[48] The estimates for both types of spending reported in Table B.8 show a decline in both the number of patents firms applied for, and the quality of the granted patents in terms of forward citations.[49] Although the effects after both types of spending are negative, the declines are larger after shocks to the services spending than to the goods spending.

The empirical estimates suggest shocks to services spending increase the establishment entry rates, de-

[46] Recent empirical studies have provided insights on the relationship among fiscal policy, business dynamism, and growth. Ferraz et al. (2015) show firms who win government procurement contracts grow more than their competitors. Lewis and Winkler (2017) argues that net firm entry rises after an expansion in the U.S. government spending. Slavtchev and Wiederhold (2016) show that technological content of government spending matters for R&D investment. Juarros (2020) argues the share of small firms matters for the transmission of fiscal shocks.

[47] We collect data at MSA-level on establishments entry and exit rates from the Business Dynamics Statistics Datasets.

[48] Appendix A.5 describes the data collection and preparation.

[49] Atanassov and Nanda (2018) show government spending crowds out the innovation effort of private firms.

crease the exit rates, and hurt the innovation activities of incumbent firms. Shocks to goods spending positively affect the entry rates, but these effects are small. Increases in the entry rates may generate positive multiplier effects due to young and innovative firms entering the market. Declines in exit rates may affect the allocation of factors across firms and the aggregate productivity by keeping alive low-productive firms. The worsening of innovative activities of incumbents may lower the economy-wide productivity. These channels impact the fiscal multiplier in different directions, and it is unclear which one dominates. Some of the differences in the fiscal multipliers by category of spending could be attributed to the productivity gains due to changes in firm behaviors. Still, as these mechanisms operate in opposite directions, it is unclear how sizable this contribution is.[50]

7 Conclusions

The Great Recession renewed the interest on the effectiveness of fiscal policy as a counter-cyclical policy tool. Most of the studies that estimate local fiscal multipliers find positive effects of the fiscal spending. However, there is substantial heterogeneity in the estimates. Theory suggests local economic characteristics and the composition of government purchases matter for the size of the fiscal multiplier. While a growing empirical literature shows the amplification role of local economic characteristics, much less is known about the role of the composition of the government spending. This paper aims to contribute in that direction.

We show purchases of services generate positive multiplier effects compared to the purchase of goods. We find that the difference in the response of employment to the type of spending is associated with the intensity of labor used to produce the products demanded by the government. We also highlight the role of the input-output network in amplifying a demand shock. We investigate the importance of other mechanisms such as tradeability and productivity improvements, in explaining these differences. We find, if any, little support for these alternative explanations.

Our findings suggest there is room for governments to redesign their fiscal spending and obtain higher multipliers by reallocating dollars from goods towards services. In this respect, our paper reinforces the idea that fiscal authorities should design policy interventions by not only choosing how much to spend, but also in what to spend. Furthermore, as the composition of the spending matters, the declining effectiveness of fiscal policy in the recent decades could be due to changes in the composition of the spending or an economy-wide decline in the importance of labor in production. Quantifying the importance of these structural changes in determining the national fiscal multiplier remains an open question and an avenue for future research.

[50] To assess the general equilibrium effect, we need a structural model analysis. We leave this part of the analysis as a future avenue of research.

References

Adao, R., Kolesár, M., and Morales, E. (2019). Shift-share designs: Theory and inference. *The Quarterly Journal of Economics*, 134(4):1949–2010.

Alonso, C. (2017). Cutting back on labor intensive goods? implications for fiscal stimulus. Mimeo.

Atanassov, J. and Nanda, V. (2018). Crowding-out innovation. In *Proceedings. Annual Conference on Taxation and Minutes of the Annual Meeting of the National Tax Association*, volume 111, pages 1–64.

Auerbach, A., Gorodnichenko, Y., and Murphy, D. (2020). Local Fiscal Multipliers and Fiscal Spillovers in the USA. *IMF Economic Review*, 68(1):195–229.

Barnichon, R., Debortoli, D., and Matthes, C. (2020). Understanding the Size of the Government Spending Multiplier: It's in the Sign. Working Papers 1145, Barcelona Graduate School of Economics.

Barro, R. (1981). Output Effects of Government Purchases. *Journal of Political Economy*, 89(6):1086–1121.

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

Basso, H. S. and Rachedi, O. (2021). The Young, the Old, and the Government: Demographics and Fiscal Multipliers. *American Economic Journal: Macroeconomics*, 13(4):110–141.

Bilbiie, F. O., Meier, A., and Müller, G. J. (2008). What Accounts for the Changes in U.S. Fiscal Policy Transmission? *Journal of Money, Credit and Banking*, 40(7):1439–1470.

Blanchard, O. 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.

Boehm, C. E. (2020). Government consumption and investment: Does the composition of purchases affect the multiplier? *Journal of Monetary Economics*, 115(C):80–93.

Born, B., Juessen, F., and Müller, G. J. (2013). Exchange rate regimes and fiscal multipliers. *Journal of Economic Dynamics and Control*, 37(2):446–465.

Borusyak, K., Hull, P., and Jaravel, X. (2022). Quasi-experimental shift-share research designs. *The Review of Economic Studies*, 89(1):181–213.

Bouakez, H., Rachedi, O., and Santoro, E. (2020). The sectoral origins of the spending multiplier. Mimeo.

Bouakez, H., Rachedi, O., and Santoro, E. (2021). The government spending multiplier in a multi-sector economy. *American Economic Journal: Macroeconomics*.

Buchheim, L., Watzinger, M., and Wilhelm, M. (2020). Job creation in tight and slack labor markets. *Journal of Monetary Economics*, 114(C):126–143.

Chodorow-Reich, G. (2019). Geographic cross-sectional fiscal spending multipliers: What have we learned? *American Economic Journal: Economic Policy*, 11(2):1–34.

Chodorow-Reich, G., Feiveson, L., Liscow, Z., and Woolston, W. G. (2012). Does State Fiscal Relief during Recessions Increase Employment? Evidence from the American Recovery and Reinvestment Act. *American Economic Journal: Economic Policy*, 4(3):118–145.

Cole, H. L. and Ohanian, L. E. (2004). New Deal Policies and the Persistence of the Great Depression: A General Equilibrium Analysis. *Journal of Political Economy*, 112(4):779–816.

Conley, T. G. and Dupor, B. (2013). The American Recovery and Reinvestment Act: Solely a government jobs program? *Journal of Monetary Economics*, 60(5):535–549.

Corbi, R., Papaioannou, E., and Surico, P. (2019). Regional transfer multipliers. *The Review of Economic Studies*, 86(5):1901–1934.

Cox, L., Müller, G., Pastén, E., Schoenle, R., and Weber, M. (2020). Big G. NBER Working Papers 27034, National Bureau of Economic Research, Inc.

Demyanyk, Y., Loutskina, E., and Murphy, D. (2019). Fiscal stimulus and consumer debt. *Review of Economics and Statistics*, 101(4):728–741.

Dolls, M., Fuest, C., and Peichl, A. (2012). Automatic stabilizers and economic crisis: US vs. Europe. *Journal of Public Economics*, 96(3):279–294.

Dupor, W. and Guerrero, R. (2017). Local and aggregate fiscal policy multipliers. *Journal of Monetary Economics*, 92(C):16–30.

Dupor, W., Karabarbounis, M., Kudlyak, M., and Mehkari, M. S. (2021). Regional Consumption Responses and the Aggregate Fiscal Multiplier. CEPR Discussion Papers 16189, C.E.P.R. Discussion Papers.

Dupor, W. and McCrory, P. B. (2018). A Cup Runneth Over: Fiscal Policy Spillovers from the 2009 Recovery Act. *Economic Journal*, 128(611):1476–1508.

Dupor, W. and Mehkari, M. S. (2016). The 2009 Recovery Act: Stimulus at the extensive and intensive labor margins. *European Economic Review*, 85(C):208–228.

Ferraz, C., Finan, F., and Szerman, D. (2015). Procuring Firm Growth: The Effects of Government Purchases on Firm Dynamics. NBER Working Papers 21219, National Bureau of Economic Research, Inc.

Flynn, J. P., Patterson, C., and Sturm, J. (2021). Fiscal policy in a networked economy. NBER Working Papers 29619, National Bureau of Economic Research, Inc.

Galeano, L., Izquierdo, A., Puig, J. P., Vegh, C. A., and Vuletin, G. (2021). Can Automatic Government Spending Be Procyclical? NBER Working Papers 28521, National Bureau of Economic Research, Inc.

Goldsmith-Pinkham, P., Sorkin, I., and Swift, H. (2020). Bartik instruments: What, when, why, and how. *American Economic Review*, 110(8):2586–2624.

Gorodnichenko, Y., Mendoza, E. G., and Tesar, L. L. (2012). The Finnish Great Depression: From Russia with Love. *American Economic Review*, 102(4):1619–1644.

Hagedorn, M., Manovskii, I., and Mitman, K. (2019). The Fiscal Multiplier. NBER Working Papers 25571, National Bureau of Economic Research, Inc.

Ilzetzki, E., Mendoza, E. G., and Végh, C. A. (2013). How big (small?) are fiscal multipliers? *Journal of Monetary Economics*, 60(2):239–254.

Isard, W. (1962). *Awards of Prime Military Contracts by County, State and Metropolitan Area of the United States, Fiscal Year 1960*. Regional Science Research Institute.

Juarros, P. (2020). Fiscal stimulus, credit frictions and the amplification effects of small firms. Mimeo.

Lewis, V. and Winkler, R. (2017). Government Spending, Entry, And The Consumption Crowding-In Puzzle. *International Economic Review*, 58(3):943–972.

Mountford, A. and Uhlig, H. (2009). What are the effects of fiscal policy shocks? *Journal of Applied Econometrics*, 24(6):960–992.

Nakamura, E. and Steinsson, J. (2014). Fiscal Stimulus in a Monetary Union: Evidence from US Regions. *American Economic Review*, 104(3):753–792.

Ramey, V. (2020). The Macroeconomic Consequences of Infrastructure Investment. In *Economic Analysis and Infrastructure Investment*, NBER Chapters: 219-268. National Bureau of Economic Research, Inc.

Ramey, V. A. and Zubairy, S. (2018). Government spending multipliers in good times and in bad: evidence from us historical data. *Journal of Political Economy*, 126(2):850–901.

Riera-Crichton, D., Vegh, C. A., and Vuletin, G. (2015). Procyclical and countercyclical fiscal multipliers: Evidence from OECD countries. *Journal of International Money and Finance*, 52(C):15–31.

Slavtchev, V. and Wiederhold, S. (2016). Does the Technological Content of Government Demand Matter for Private R&D? Evidence from US States. *American Economic Journal: Macroeconomics*, 8(2):45–84.

Suárez-Serrato, J. C. and Wingender, P. (2016). Estimating Local Fiscal Multipliers. NBER Working Papers 22425, National Bureau of Economic Research, Inc.

Vom Lehn, C. and Winberry, T. (2022). The investment network, sectoral comovement, and the changing us business cycle. *The Quarterly Journal of Economics*, 137(1):387–433.

Wilson, D. J. (2012). Fiscal spending jobs multipliers: Evidence from the 2009 American Recovery and Reinvestment Act. *American Economic Journal: Economic Policy*, 4(3):251–82.

Woodford, M. (2011). Simple Analytics of the Government Expenditure Multiplier. *American Economic Journal: Macroeconomics*, 3(1):1–35.

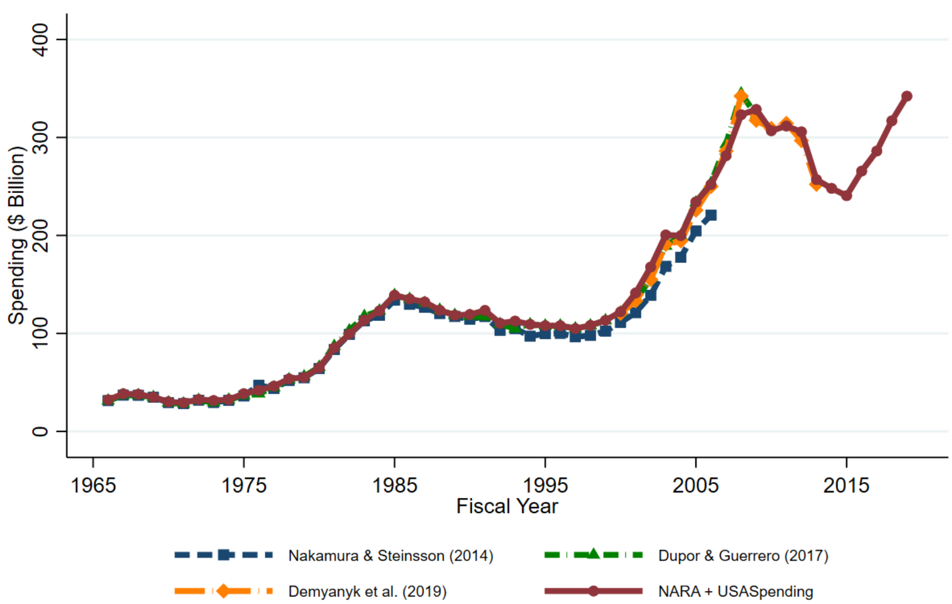
Appendix

A. Data

A.1. Military Spending Validation

This section discusses the quality of the DoD military procurement contract microdata, and we compare them with data previously used in the literature. Figure A.1 compares the aggregate military spending in nominal terms at national level by fiscal years. We compare the aggregate military procurement spending derived from our microdata with the ones calculated by Nakamura and Steinsson (2014), Dupor and Guerrero (2017), and Demyanyk et al. (2019). Overall, we conclude our data match well the trend and the level of aggregate spending used in previous works.

Figure A.1: Military Spending Comparisons



Note: The period for the comparison matches the research design of the comparison studies. Data for the comparison studies have been downloaded from their journal's data repository.

We also compare the spending at different geographically-disaggregated areas. To this end, we regress

$$Spend_{l,t}^{ours} = \beta Spend_{l,t}^{comp} + \alpha_l + \delta_t + \varepsilon_{l,t}$$

where $Spend_{l,t}^{ours}$ represents the military spending from our data in locality l at time t; $Spend_{l,t}^{comp}$ is the spending for one of the comparison datasets; and α_l and δ_t are locality and time fixed effects, respectively.

Table A.1 reports the results of our comparison. In column 1, we disaggregate the spending at state-level and compare our calculations with the data used by Nakamura and Steinsson (2014). Column 2 shows the comparison between our state-level aggregation and the ones constructed by Dupor and Guerrero (2017). Finally, in column 3, we show the CBSA-level comparison between our data and Demyanyk et al. (2019).

Table A.1: Military Spending Comparisons by Geography

	(1) Nakamura and Steinsson (2014)	(2) Dupor and Guerrero (2017)	(3) Demyanyk et al. (2019)
β	1.12	0.94	0.98
95% C. I. for β	(1.02 - 1.21)	(0.89 - 0.99)	(0.89 - 1.07)
Observations	2,050	2,200	10,636
Geographic Unit	State	State	CBSA
Number Localities	50	50	862
Period	$1966-2006$	$1966-2009$	$2000-2012$
Locality FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
$Within-R^2$	0.97	0.96	0.79

Notes: The values in the brackets report the 95% confidence interval. Standard errors are clustered at the geographic unit level. The period of analysis and the geographic aggregation are chosen to match the research design of the comparison studies. Data for the comparison studies have been downloaded from their journal's data repository.

We focus on two tests to evaluate the quality of the geographic distribution of our data compared to previous studies. First, if there were a one-to-one relationship between the geographic allocation between our data and previously used data, β should be equal or close to one. Second, if there were a strong similarity between our data and the others, the within-R^2 should be high. The results show that the value of one is either included in the 95% confidence interval or it is close to the either upper or lower bound. The largest discrepancy is between our data and the ones from Nakamura and Steinsson (2014). This discrepancy, as showed in Figure A.1, comes from the years between 2000 and 2006. The data collected by Nakamura and Steinsson (2014) underestimate the aggregate spending, while ours are similar to the other sources of comparison. The within-R^2 are over 0.95 for the comparison with Nakamura and Steinsson (2014) and Dupor and Guerrero (2017), and it is a little bit lower in the comparison with Demyanyk et al. (2019). As our data are constructed until 2006 using NARA and USASpending after that year and Demyanyk et al. (2019) use only USASpending starting from 2001, these tests also confirm the comparability between the information provided by NARA and USASpending. Overall, these results suggest our data are highly comparable with the ones used in previous studies.

A.2. Product Codes Classification

Our dataset includes 3239 distinct 4-digit product codes of which 2547 are Product Services Codes (PSCs) and identify spending in services, and 692 are the Federal Supply Codes (FSCs) and identify spending in goods. Product Services Codes are grouped into 24 sub-groups, while FSCs are grouped into 85 groups. Table A.2 reports the list of macro-groups for the two categories of spending.

Table A.2: List of Product Codes by Category of Spending

Product Service Codes (PSCs)	Federal Supply Codes (FSCs)
A - Research and Development	10 - Weapons
B - Special Studies/analysis, Not R&d	11 - Nuclear Ordnance
C - Architect/engineer Services	12 - Fire Control Equipment
D - IT and Telecommunication - Information Technology and Telecommunications	13 - Ammunition and Explosives
E - Purchase Of Structures/facilities	14 - Guided Missles
F - Natural Resources Management	15 - Aerospace Craft and Structural Components
G - Social Services	16 - Aerospace Craft Components and Accessories
H - Quality Control, Test, Inspection	17 - Aerospace Craft Launching, Landing, Ground Handling and Servicing Equipment
J - Maintenance, Repair, Rebuild Equipment	19 - Ships, Small Craft, Pontoon, Docks
K - Modification Of Equipment	20 - Ship and Marine Equipment
L - Technical Representative Services.	22 - Railway Equipment
M - Operation Of Govt Owned Facility	23 - Motor Vehicles, Cycles, Trailers
N - Installation Of Equipment	24 - Tractors
P - Salvage Services	25 - Vehicular Equipment Components
Q - Medical Services	26 - Tires and Tubes
R - Support Services (prof, Admin, Management)	28 - Engines and Turbines and Component
S - Utilities and Housekeeping	29 - Engine Accessories
T - Photo, Map, Print, Publication	30 - Mechanical Power Transmission Equipment
U - Education and Training	31 - Bearings
V - Transport, Travel, Relocation	32 - Woodworking Machinery and Equipment
W - Lease/rent Equipment	34 - Metalworking Machinery
X - Lease/rent Facilities	35 - Service and Trade v
Y - Construct Of Structures/facilities	36 - Special Industry Machinery
Z - Maintenance, Repair, Alter Real Property	37 - Agricultural Machinery and Equipment
	38 - Construct/mine/excavate/highway Equipment
	39 - Materials Handling Equipment
	40 - Rope, Cable, Chain, Fittings
	41 - Refrigeration, Air Condition/circulation Equipment
	42 - Fire/rescue/safety; Environment Protect
	43 - Pumps and Compressors
	44 - Furnace/steam/drying; Nuclear Reactor
	45 - Plumbing, Heating, Waste Disposal
	46 - Water Purification/sewage Treatment
	47 - Pipe, Tubing, Hose, and Fittings
	48 - Valves
	49 - Maintenance/repair Shop Equipment
	51 - Hand Tools
	52 - Measuring Tools
	53 - Hardware and Abrasives
	54 - Prefab Structures/scaffolding
	55 - Lumber, Millwork, Plywood, Veneer
	56 - Construction and Building Material
	58 - Comm/detect/coherent Radiation
	59 - Electrical/electronic Equipment Components
	61 - Electric Wire, Power Distribution Equipment
	62 - Lighting Fixtures, Lamps
	63 - Alarm, Signal, Security Detection
	65 - Medical/dental/veterinary Equipment/supply
	66 - Instruments and Laboratory Equipment
	67 - Photographic Equipment
	68 - Chemicals and Chemical Products
	69 - Training Aids and Devices
	7A - IT and Telecommunication - Applications
	7B - IT and Telecommunication - Compute
	7C - IT and Telecommunication - Data Center
	7D - IT and Telecommunication - Delivery
	7E - IT and Telecommunication - End User
	7F - IT and Telecommunication - IT Management
	7G - IT and Telecommunication - Network
	7H - IT and Telecommunication - Platform
	7J - IT and Telecommunication - Security and Compliance
	7K - IT and Telecommunication - Storage
	71 - Furniture
	72 - Household/commercial Furnish/appliance
	73 - Food Preparation/serving Equipment
	74 - Office Mach/text Process/visib Rec
	75 - Office Supplies and Devices
	76 - Books, Maps, Other Publications
	77 - Musical Inst/phonograph/home Radio
	78 - Recreational/athletic Equipment
	79 - Cleaning Equipment and Supplies
	80 - Brushes, Paints, Sealers, Adhesives
	81 - Containers/packaging/packing Suppl
	83 - Textile/leather/fur; Tent; Flag
	84 - Clothing, Individual Equipment, Insigna, and Jewelry
	85 - Toiletries
	87 - Agricultural Supplies
	88 - Live Animals
	89 - Subsistence
	91 - Fuels, Lubricants, Oils, Waxes
	93 - Nonmetallic Fabricated Materials
	94 - Nonmetallic Crude Materials
	95 - Metal Bars, Sheets, Shapes
	96 - Ores, Minerals and Primary Products
	99 - Miscellaneous

A.3. Industry Classification by Labor Intensity

We implement the following steps to assign a contract to an industry with high- or low-labor intensity. First, we collect from the BEA data on value added and its decomposition —compensation of employees, taxes on production and imports less subsidies, and gross operating surplus—for 66 macro-industries starting from 1997. Next, we compute the industry's average labor share as the average of the ratios between the compensation of employees and the value added. We then calculate the median average labor intensity, and we classify high-labor intensive industries as the industries with industry averages above the median, low-labor intensive industries otherwise. Finally, we use the first two, three, or four digits of the contractor's 6-digit NAICs code to assign a contract to a high- or low-labor intensive industry.

Table A.3: Industry Classification by Labor Intensity

Low-Labor Intensive Industries	High-Labor Intensive Industries
111-112 - Farms	213 - Support activities for mining
113-115 - Forestry, fishing, and related activities	23 - Construction
211 - Oil and gas extraction	313-314 - Textile mills and textile product mills
212 - Mining, except oil and gas	315, 316 - Apparel and leather and allied products
22 - Utilities	321 - Wood products
311-312 - Food and beverage and tobacco products	323 - Printing and related support activities
322 - Paper products	326 - Plastics and rubber products
324 - Petroleum and coal products	332 - Fabricated metal products
325 - Chemical products	333 - Machinery
327 - Nonmetallic mineral products	35 - Electrical equipment, appliances, and components
331 - Primary metals	3364, 3369 - Other transportation equipment
334 - Computer and electronic products	337 - Furniture and related products
3361-3363 - Motor vehicles, bodies and trailers, and parts	339 - Miscellaneous manufacturing
42 - Wholesale trade	445 - Food and beverage stores
441 - Motor vehicle and parts dealers	452 - General merchandise stores
442-444, 446-448, 451, 453-454 - Other retail	482 - Rail transportation
481 - Air transportation	484 - Truck transportation
483 - Water transportation	487-488, 492 - Other transportation and support activities
485 - Transit and ground passenger transportation	491 - Federal Government Enterprises
486 - Pipeline transportation	493 - Warehousing and storage
511 - Publishing industries (includes software)	523 - Securities, commodity contracts, and investment
512 - Motion picture and sound recording industries	5412-5414, 5416-5419 - Miscellaneous professional, scientific, and technical services
515-517 - Broadcasting and telecommunications	5415 - Computer systems design and related services
514, 518-519 - Information and data processing services	55 - Management of companies and enterprises
521-522 - Federal Reserve banks, credit intermediation, and related activities	561 - Administrative and support services
524 - Insurance carriers and related activities	611 - Educational services
525 - Funds, trusts, and other financial vehicles	621 - Ambulatory health care services
531 - Real Estate	622 - Hospitals
532-533 - Rental and leasing services and lessors of intangible assets	623 - Nursing and residential care facilities
5411 - Legal services	624 - Social assistance
562 - Waste management and remediation services	713 - Amusements, gambling, and recreation industries
711-712 - Performing arts, spectator sports, museums, and related activities	722 - Food services and drinking places
721 - Accommodation	81 - Other services, except government

Table A.3 lists the industries by labor intensity. Out of the 66 industries, half are classified as low-labor intensive, and half as high-labor intensive. The contracts in our final sample are allocated to all 66 BEA's industry groups. The DoD spending in unequally split between high- and low-labor intensive industries. Indeed, about 85% of spending is directed to firms operating in high-labor intensive industries.

A.4. Imputation Strategy

In order to precisely estimate the effects of local government spending shocks on the outcome variables, the availability of the longest time series possible is relevant. In our data, we face two challenges. First, information on the industry which a contractor belongs to is available from 1988. We use the product codes to assign the industry to contracts for which this information is not reported. Intuitively, we assign to a contract with no information about the industry of the contractor the dominant industry in the production of the product code associated with the contract. We follow several sequential steps to infer the largest number of missing information. We first assign the BEA's industry that receives the largest share of spending for a specific product code to the contract with missing industry. As the BEA industry classification includes several narrower industries, if there are two or more BEA industries having the same share, we assign the finer industry with the highest share of contracts. Finally, in the cases in which the previous steps do not successfully impute the missing industry code, if the two or more potential industry choices are all classified as either low or high labor-intensive, we don't impute the industry but we classify that contract as either high or low labor-intensive. For the scope of our analysis, we are interested in correctly allocating contacts to one of these groups. Thus, if all potential industries are in the same group, we can safely assign that contract to that group. After the previous steps, if the imputation is not successful, we classify the product code as not imputed.

Industry information is missing for about 5% of the contracts awarded after 1979. These contracts account for 17% of the total spending after 1979. Using the imputation strategy, we fail to impute the industry for less than 1% of the contracts, corresponding also to less than 1% of spending. Out of the imputed contracts, the vast majority, 16% of the total spending, is imputed using the first step of our imputation strategy. The following steps of the imputation procedure only lead to minor gains in imputing the missing industry codes.

The second challenge consists in classifying industries before 1997 between more and less labor-intensive as data on value added and its components are reported only from 1997. We simply use the classification based on data from 1997 also to prior years. For example, if an industry is classified as low-labor intensive using the data for $1997-2020$, we classify that industry as low-labor intensive also for the years before 1997.

Our imputation strategy relies on two assumptions. First, the share of spending allocated to a product code is a good predictor of the industry that produce. To address this concern, we compute for each product code the share of the spending covered by the industry that we assign as imputed industry for that product code. Results show that on average the share of spending in a product code covered by the industry used for the imputation is over 70%. Furthermore, for over 75% of product codes the industry we elect as imputed industry covers over 50% of the spending in that product code. Overall, these tests suggest the spending for product codes is concentrated, with an industry receiving a large share of this spending.

Second, the classification of industries between high and low labor-intensive is stable over time. We test the stability by re-classify industries between high and low labor-intensive year by year, and comparing this new classification with the benchmark used in the paper. We find that the classification based on the annual labor intensity measures matches for more than 91% of the cases the classification based on the full sample, implying the allocation of industries between the two groups does not significantly vary over time. The mismatches are distributed across about one-third of the industries implying they are not dominated by a few industries. Furthermore, apart from a couple of industries, the mismatches do not follow a trend that could signal a change in the industry technology. Overall, these tests reinforce the idea that the classification of industries between high and low labor-intensive is quite stable over time.

A.5. Patent Data

We collect patent data from PatentsView at the end of 2019. The PatentsView database contains the universe of granted patents from the US Patent and Trademark Office (USPTO) starting from 1976 until 2019. These data contain patent-level information including application dates, the type of patent, the inventors' names, the latitude and longitude of their addresses, and the citations' network consisting of the number of citations made to and received from other patents.

We restrict our analysis to utility patents that cover the creation of a new or improved product, process, or machine. Utility patents are also commonly known as "patents for invention," and they account for about 98% of the universe of patents granted by the USPTO. We also restrict to patents with application year starting from 1976. We do not observe the exact date in which an innovation occurs. As common in the literature, we identify the year when an innovation occurs as the application year of a patent, which is the year when the provisional application is considered complete by the USPTO, and a filing date is set.

Patent data suffer from three major issues that affect the over-time comparison of patent statistics: 1) the changes in the propensity to cite; 2) the lifespan of a patent; and 3) the truncation bias. We address both issues and adjust the microdata by following different strategies proposed in the literature.

The propensity to cite bias is generated by the changing patterns of patenting and citing over time. Starting from the 1980s, patenting and citing activity in the US have experience a dramatic acceleration. This increase in the use of patents is view as a response to the increase in the patent protection provided by the legislator, rather than an endogenous rise in the amount of innovation. An over-time comparison of patent activities without a proper adjustment to correct these trends could generate misleading conclusions. Similarly, a shift to the left of the citation-lag distribution, implying that citations are coming sooner than they used to in the previous decades, may also lead to the same misreading of the results. We implement a standard "quasi-structural" approach to correct for the propensity-to-cite bias.

The second critical time effect is linked to the lifespan of a patent. Older patents have longer time to accumulate citations than more recent patents. As common in the literature, we measure the quality-adjusted innovation rates within a fixed window of five years after the grant year. In this way, we make the citing activities comparable between patents that have different lifespans.

Finally, the truncation bias also mechanically affects the the number of citations that a patent receives. Patent records are released at the grant dates when the review process is completed. As a result, the truncation bias causes patents in the last years of the sample to be mechanically less cited, independently of their innovativeness. The review process takes essentially 8 years to be fully completed.[51] Thus, we restrict the analysis to patents that applied by the end of 2006. We undertake this very cautious approach because we also want to ensure that citing patents in the 5-year window have completed their review process.

We associate a patent with a MSA by using the latitude and longitude of the address of the patent inventors. We geolocate the latitude and longitude into MSAs by using the 2010 TIGER/Line Shapefile constructed by the US Census Bureau. When patents have multiple inventors whose addresses are in different MSAs, we equally split those patents across MSAS according to the number of inventors. In the geographical aggregation, we restrict our sample to patents whose at least one inventor is located in the US mainland. The final is constructed by including only patents for which the DoD has no economic interest. This restriction serves to better capture private innovation efforts. Finally, we also remove all MSAs with incomplete histories in granted patents or citations.

We construct two measures of innovation quality: the number of patents, and the patents weighted by the number of adjusted forward citations within a 5-year windows after the grant year. The first statistics measures the number of patents filed at the USPTO, while the second measures the quality of the stock of patents in a specific year.

[51] The USPTO reports: "*As of 12/31/2012, utility patent data, as distributed by year of application, are approximately 95% complete for utility patent applications filed in 2004, 89% complete for applications filed in 2005, 80% complete for applications filed in 2006, 67% complete for applications filed in 2007, 49% complete for applications filed in 2008, 36% complete for applications filed in 2009, and 19% complete for applications filed in 2010.*"

B. Empirical Results

Table B.1: Local Fiscal Multipliers - Spending by Components

	(1) t	(2) t+1	(3) t+2	(4) t+3	(5) t+4
	Dependent Variable: Employment				
Services Spending	0.355***	0.968***	1.191***	1.454***	1.208***
	(0.123)	(0.167)	(0.199)	(0.206)	(0.217)
Goods Spending	0.059	0.165	0.157	0.186	0.174
	(0.041)	(0.128)	(0.121)	(0.177)	(0.167)
Test $\beta_g^k = \beta_s^k$	[.0333]	[.0003]	[.0000]	[.0000]	[.0001]
	Dependent Variable: Salary & Wages				
Services Spending	0.585***	1.924***	2.424***	2.892***	2.448***
	(0.151)	(0.331)	(0.403)	(0.464)	(0.510)
Goods Spending	0.078	0.269	0.227	0.260	0.207
	(0.064)	(0.189)	(0.173)	(0.249)	(0.220)
Test $\beta_g^k = \beta_s^k$	[.0048]	[.0001]	[.0000]	[.0000]	[.0001]
	Dependent Variable: Dividends				
Services Spending	-0.155	0.137	-0.079	0.005	0.331
	(0.101)	(0.174)	(0.208)	(0.271)	(0.267)
Goods Spending	-0.010	-0.013	-0.062	-0.090	-0.086
	(0.116)	(0.099)	(0.096)	(0.113)	(0.110)
Test $\beta_g^k = \beta_s^k$	[.3745]	[.4153]	[.9377]	[.7248]	[.1281]
	Dependent Variable: Personal Income				
Services Spending	0.291***	1.128***	1.383***	1.742***	1.508***
	(0.090)	(0.221)	(0.249)	(0.250)	(0.282)
Goods Spending	0.023	0.181	0.156	0.187	0.159
	(0.056)	(0.116)	(0.107)	(0.155)	(0.141)
Test $\beta_g^k = \beta_s^k$	[.0176]	[.0004]	[.0000]	[.0000]	[.0000]

Notes: All panels report the estimates from equation (2). The instruments are computed as in equation (4). The balanced panel consists of 12,024 observations, and includes 334 MSAs for the period 1979 − 2019. The instrument is calculated over the same period. Standard errors are clustered by MSA. The symbols *, **, and *** represent the 10%, 5%, and 1% significance levels, respectively. All regressions include MSA and year fixed effects.

Table B.2: Local Fiscal Multipliers - Cumulative Spending by Components

	(1) t	(2) t+1	(3) t+2	(4) t+3	(5) t+4
	Dependent Variable: Employment				
Services Spending	0.355***	0.572***	0.561***	0.547***	0.449***
	(0.123)	(0.113)	(0.093)	(0.072)	(0.060)
Goods Spending	0.059	0.099	0.092	0.083	0.073
	(0.041)	(0.071)	(0.059)	(0.061)	(0.053)
Test $\beta_g^k = \beta_s^k$	[.0333]	[.001]	[.0001]	[.0000]	[.0001]
	Dependent Variable: Salary & Wages				
Services Spending	0.585***	1.086***	1.100***	1.079***	0.900***
	(0.151)	(0.198)	(0.180)	(0.156)	(0.143)
Goods Spending	0.078	0.155	0.133	0.114	0.091
	(0.064)	(0.106)	(0.084)	(0.086)	(0.072)
Test $\beta_g^k = \beta_s^k$	[.0048]	[.0002]	[.0000]	[.0000]	[.0001]
	Dependent Variable: Dividends				
Services Spending	-0.155	0.109	0.098	0.139	0.191**
	(0.101)	(0.098)	(0.087)	(0.093)	(0.087)
Goods Spending	-0.010	-0.007	-0.031	-0.029	-0.018
	(0.116)	(0.061)	(0.048)	(0.040)	(0.034)
Test $\beta_g^k = \beta_s^k$	[.3745]	[.2825]	[.1672]	[.0749]	[.0182]
	Dependent Variable: Personal Income				
Services Spending	0.291***	0.639***	0.644***	0.657***	0.554***
	(0.090)	(0.135)	(0.114)	(0.095)	(0.084)
Goods Spending	0.023	0.099	0.086	0.076	0.063
	(0.056)	(0.066)	(0.054)	(0.055)	(0.048)
Test $\beta_g^k = \beta_s^k$	[.0176]	[.0008]	[.0000]	[.0000]	[.0000]

Notes: The balanced panel consists of $12,024$ observations, and includes 334 MSAs for the period $1979-2019$. The instrument is calculated over the same period. Standard errors are clustered by MSA. The symbols *, **, and *** represent the 10%, 5%, and 1% significance levels, respectively. All regressions include MSA and year fixed effects.

Table B.3: Local Fiscal Multipliers - Outlays by Components

	(1) t	(2) t+1	(3) t+2	(4) t+3	(5) t+4
	Dependent Variable: Employment				
Services Spending	0.283***	0.334***	0.413*	0.405	0.423*
	(0.089)	(0.125)	(0.227)	(0.261)	(0.249)
Goods Spending	-0.108	-0.122	-0.115	-0.086	-0.104
	(0.124)	(0.127)	(0.116)	(0.094)	(0.103)
Test $\beta_g^k = \beta_s^k$	[.0189]	[.0176]	[.0549]	[.0931]	[.0662]
	Dependent Variable: Salary & Wages				
Services Spending	0.865***	1.016**	1.219**	1.240**	1.333**
	(0.292)	(0.410)	(0.564)	(0.615)	(0.607)
Goods Spending	-0.263*	-0.310*	-0.315*	-0.261*	-0.302*
	(0.151)	(0.161)	(0.171)	(0.154)	(0.166)
Test $\beta_g^k = \beta_s^k$	[.0013]	[.0057]	[.0165]	[.0257]	[.0149]
	Dependent Variable: Dividends				
Services Spending	0.847***	1.055***	1.221***	1.185***	1.194***
	(0.208)	(0.262)	(0.353)	(0.367)	(0.346)
Goods Spending	-0.055	-0.257	-0.268	-0.184	-0.194
	(0.396)	(0.222)	(0.200)	(0.194)	(0.214)
Test $\beta_g^k = \beta_s^k$	[.0672]	[.0003]	[.0003]	[.0008]	[.0004]
	Dependent Variable: Personal Income				
Services Spending	0.865***	1.016**	1.219**	1.240**	1.333**
	(0.292)	(0.410)	(0.564)	(0.615)	(0.607)
Goods Spending	-0.263*	-0.310*	-0.315*	-0.261*	-0.302*
	(0.151)	(0.161)	(0.171)	(0.154)	(0.166)
Test $\beta_g^k = \beta_s^k$	[.0013]	[.0057]	[.0165]	[.0257]	[.0149]

Notes: All panels report the estimates from equation (2). The instruments are computed as in equation (4). The balanced panel consists of 8,684 observations, and includes 334 MSAs for the period 1989 − 2019. The instrument is calculated over the same period. Standard errors are clustered by MSA. The symbols *, **, and *** represent the 10%, 5%, and 1% significance levels, respectively. All regressions include MSA and year fixed effects.

Table B.4: Labor Intensity and Employment Concentration Ratios

	(1) 4 largest firms	(2) 8 largest firms	(3) 20 largest firms	(4) 50 largest firms
Employment Concentration Ratio	-0.001*	-0.000	-0.001	-0.001
	(0.000)	(0.001)	(0.001)	(0.001)
Observations	198	204	204	204
Number 4-digits Industries	99	102	102	102
Industry Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes

Notes: The balanced panel includes the years 2002 and 2017. Robust errors are used. The symbols *, **, and *** represent the 10%, 5%, and 1% significance levels, respectively.

Table B.5: Local Fiscal Multiplier - Spending by Category and State Aggregation

	(1) t	(2) t+1	(3) t+2	(4) t+3	(5) t+4
	Dependent Variable: Employment				
Services Spending	0.845	2.736***	2.989**	3.430***	2.846***
	(0.709)	(0.929)	(1.221)	(0.781)	(0.799)
Goods Spending	0.240	0.624	0.650	1.150	1.295*
	(0.304)	(0.501)	(0.553)	(0.715)	(0.719)
	Dependent Variable: GDP				
Services Spending	0.921	4.882***	6.525**	7.897***	7.154***
	(0.665)	(1.258)	(2.582)	(2.137)	(2.241)
Goods Spending	-0.113	-0.160	-0.164	-0.110	-0.188
	(0.583)	(0.941)	(1.073)	(1.222)	(1.234)

Notes: All panels report the estimates from equation (2). The instruments are computed as in equation (4). The balanced panel consists of $1,800$ observations, and includes 50 states, excluding DC, for the period $1979 - 2019$. The instrument is calculated over the same period. Standard errors are clustered by state. The symbols *, **, and *** represent the 10%, 5%, and 1% significance levels, respectively. All regressions include state and year fixed effects.

Table B.6: Neighboring Locations - Spending by Components

	(1) t	(2) t+1	(3) t+2	(4) t+3	(5) t+4
	Dependent Variable: Employment				
Services Spending	0.109	0.262***	0.386**	0.573*	0.525*
	(0.084)	(0.093)	(0.170)	(0.298)	(0.303)
Goods Spending	0.070	-0.008	-0.217	-0.366	-0.277
	(0.164)	(0.139)	(0.332)	(0.509)	(0.479)
	Dependent Variable: Salary & Wages				
Services Spending	0.226*	0.492***	0.710***	1.006**	0.951**
	(0.133)	(0.132)	(0.267)	(0.392)	(0.375)
Goods Spending	-0.070	-0.120	-0.476	-0.709	-0.575
	(0.373)	(0.197)	(0.471)	(0.643)	(0.610)

Notes: All panels report the estimates from equation (8). The instruments are computed as in equation (9). The balanced panel consists of $6,948$ observations, and includes 195 MSAs for the period $1979 - 2019$. The instrument is calculated over the same period. Standard errors are clustered by MSA. A neighboring location is defined as a location within a radius of 50 miles from the comparison location. The symbols *, **, and *** represent the 10%, 5%, and 1% significance levels, respectively. All regressions include MSA and year fixed effects.

Table B.7: Entry and Exit Rates

	(1) t	(2) t+1	(3) t+2	(4) t+3	(5) t+4
	Dependent Variable: Entry Rate				
Services Spending	-0.171	0.112	0.390	0.637**	0.601**
	(0.399)	(0.282)	(0.256)	(0.292)	(0.237)
Goods Spending	-0.069	0.216	0.151*	0.163	0.180
	(0.116)	(0.135)	(0.084)	(0.108)	(0.125)
	Dependent Variable: Entry Rate Size 1-19				
Services Spending	0.175	0.555*	0.593**	0.900**	0.775***
	(0.345)	(0.318)	(0.271)	(0.367)	(0.262)
Goods Spending	-0.155	0.199	0.133	0.185*	0.201*
	(0.157)	(0.122)	(0.091)	(0.111)	(0.117)
	Dependent Variable: Entry Rate Size 20-499				
Services Spending	1.649	-1.420*	1.769	0.858	1.286*
	(1.831)	(0.826)	(1.684)	(0.693)	(0.753)
Goods Spending	0.944	0.641	0.403	0.186	0.306
	(0.743)	(0.396)	(0.256)	(0.254)	(0.298)
	Dependent Variable: Entry Rate Size 500+				
Services Spending	-3.043	-0.639	-0.839	-0.114	0.149
	(2.709)	(0.984)	(1.125)	(0.930)	(0.683)
Goods Spending	-0.513	0.642	0.435	0.718	0.788
	(0.555)	(0.614)	(0.300)	(0.541)	(0.540)
	Dependent Variable: Exit Rate				
Services Spending	-0.587	-1.196**	-1.105***	-1.109***	-0.782***
	(0.560)	(0.543)	(0.317)	(0.295)	(0.256)
Goods Spending	0.143	-0.243	-0.064	-0.199	-0.229
	(0.239)	(0.192)	(0.188)	(0.196)	(0.197)

Notes: All panels report the estimates from equation (2). The instruments are computed as in equation (4). The balanced panel consists of 12,024 observations, and includes 334 MSAs for the period 1979 − 2019. The instrument is calculated over the same period. Standard errors are clustered by MSA. The symbols *, **, and *** represent the 10%, 5%, and 1% significance levels, respectively. All regressions include MSA and year fixed effects.

Table B.8: Innovation Activities

	(1) t	(2) t+1	(3) t+2	(4) t+3	(5) t+4
	Dependent Variable: Number of Patents				
Services Spending	-3.294	-0.984	-3.072	-3.100	-5.304**
	(2.458)	(1.648)	(2.059)	(2.230)	(2.690)
Goods Spending	-1.762**	-0.438	0.171	0.598	1.066
	(0.851)	(0.629)	(0.506)	(0.731)	(0.878)
	Dependent Variable: Number of Citations				
Services Spending	-9.660	-3.428	-9.563*	-9.037*	-14.701**
	(6.328)	(4.291)	(5.369)	(5.200)	(6.163)
Goods Spending	-1.915	-1.042	-0.850	-0.186	2.103
	(4.660)	(2.438)	(2.759)	(4.524)	(3.112)

Notes: All panels report the estimates from equation (2). The instruments are computed as in equation (4). The balanced panel consists of $6,716$ observations, and includes 292 MSAs for the period $1979-2006$. The instrument is calculated over the same period. Standard errors are clustered by MSA. The symbols *, **, and *** represent the 10%, 5%, and 1% significance levels, respectively. All regressions include MSA and year fixed effects.