Heterogeneous Spending, Heterogeneous Multipliers*

Umberto Muratori†

Pedro Juarros‡

Daniel Valderrama§

January 30, 2023

Abstract

Do local fiscal multipliers depend on what the government purchases? We find that government purchases of services have larger effects on employment than spending on goods. Industries producing services are more labor-intensive than industries producing goods. This heterogeneity in labor intensity is an important mechanism behind these results. Spending directed toward labor-intensive industries generates stronger increases in income and consumption relative to spending toward non-labor-intensive industries.

JEL Codes: E10, E62, H57, R12

Keywords: Heterogeneous Local Fiscal Multipliers, Services and Goods Spending, Labor Intensity

*We are grateful to Laurent Bouton, Russell Cooper, Mark Huggett, Toshihiko Mukoyama, and Martin Ravallion for their helpful comments. We also thank seminar participants at the Univerisity of Florence and the University of Siena and conference participants at the Max-Weber Programme June Conference. We thank Digna Isabel Andrade Portilla for her excellent research assistance. We acknowledge the generous financial support of 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 literature on the government purchase multiplier commonly treats government spending as a homogeneous good. Very little is known about the heterogeneity in the fiscal multipliers coming from the heterogeneity in spending (Chodorow-Reich, 2019), mainly due to the lack of granular data on government purchases. The 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 Department of Defense (DOD) purchases (Nakamura and Steinsson, 2014; Dupor and Guerrero, 2017; Demyanyk et al., 2019; Auerbach et al., 2020) across regions in the US. These studies estimate local fiscal multipliers using aggregate spending and do not capture the variation in the estimates associated with different types of spending.

Our paper fills this gap by answering two questions. Do different types of government purchases generate different multipliers? What are the mechanisms driving this heterogeneity? We assemble a unique dataset containing over 40 years of military contract-level US procurement spending at an annual frequency.[1] Our data report information about the products that contracts require recipients to produce and the location where a contract is performed. We leverage this new information to classify fiscal spending into two categories: purchases of goods and purchases of services. The geolocation of contracts at the metropolitan statistical area (MSAs) in the US allows quantifying the causal effects of types of public purchases on employment using the cross-sectional and time variation (Nakamura and Steinsson, 2014).

The breakdown into goods versus services is a natural distinction to explore. The government's core decision is the budget allocation between these two types of products, for which it demands input factors to provide public goods and services to its clients. Our paper shows that changing the budget allocation between goods and services can improve the effectiveness of fiscal packages. Goods and services differ along several dimensions, such as labor intensity, tradability, and productivity. Their comparison sheds light on the relevance of specific channels in generating positive aggregate responses. Exploring these mechanisms leads to a better understanding of the effectiveness and design of fiscal stimuli.[2]

The first novel result of our paper consists of the switch of government defense spending from goods to services starting in the early 1990s. Before that date, more than 65% of procurement defense spending was used to purchase goods. In recent decades, the US federal government has redesigned its budget allocation, purchasing approximately equal shares of goods and services.

[1]For the analysis of this project, we rely on defense spending. Defense spending composes the largest fraction of federal procurement spending; it accounts for about 50% of transactions by count and more than 60% by value. Cox et al. (2020) show that the characteristics of defense spending are similar to those of the spending made by other agencies. We are unlikely to miss important features by restricting our attention to defense spending. Furthermore, due to political accountability, the US federal government has disclosed defense procurement contracts for longer than contracts signed by other federal agencies, allowing us to leverage a long time series for the analysis. Finally, as argued by Nakamura and Steinsson (2014), variation in spending awarded by the Department of Defense (DoD) is mainly driven by geopolitical events, and it is less sensitive to local economic conditions than the spending awarded by other federal agencies, making our identification strategy stronger.

[2]Ramey and Shapiro (1998) show that changes in the composition of government spending across sectors generate negative responses in the aggregate variables due to costly reallocation of the production factors. Cox et al. (2020) highlight that the degree of sectoral price stickiness impacts the fiscal transmission mechanism.

We document significant differences in the employment of local fiscal multipliers between shocks to spending in goods or services. A one-percentage-point increase in government services spending (normalized by personal income) generates an on-impact employment response of 0.36 percent. The estimates increase as time passes and become 1.19 two years after the shock.[3] 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 at impact and 0.19 percent at the two-year horizon, respectively; they are between one-fourth and one-tenth smaller than the estimates for services spending. Unlike goods spending, services spending also leads to a significant increase in the household's labor income. To grasp a better sense of the size of these estimates, we calculate the newly-created jobs for $100,000 spent in procurement contracts. The number of new jobs created substantially differs between the two types of spending. $100,000 in procurement contracts that acquire services create 0.81 new jobs at impact, and 2.71 after two years. By contrast, the purchase of goods generates 0.13 new jobs at impact and 0.36 after two years. Chodorow-Reich (2019) reports a range from 0.76 to 3.80 of new jobs created at the two-year horizon. The number of new jobs created by services spending is close to the upper bound of that range, while the new jobs generated by goods spending is slightly below the lower bound. Our estimates are in line with the literature and highlight the role of heterogeneous spending on the size of the local fiscal multiplier.

We also investigate whether the employment responses occur in the industries that receive the procurement spending or the effects spill over to industries that do not receive the spending. We find that both types of spending generate positive and significant direct effects on the industry that receives most of that spending.[4] Still, indirect effects consisting of positive employment responses that spill to other sectors occur only after service spending shocks. We find that labor income increases only after services spending shocks, generating multiplier effects that propagate to all industries.[5] These results highlight the importance of the transmission of fiscal spending through changes in labor income.

In the second part of the paper, we separately test three mechanisms that may explain the differences in the estimates between services and goods spending. First, the production of services is more labor-intensive than the production of goods. Government dollars spent on acquiring services may increase the demand for workers, labor income, and private consumption by more than the dollars spent in producing goods. This pass-through mechanism may generate a virtuous circle leading to larger multiplier effects after services spending shocks. Second, goods are more tradeable than services. If the production of goods spills into neighboring locations, small fiscal multiplier effects after goods spending shocks may be due to geographic spillovers. Third, government spending may generate demand-driven growth by facilitating firm turnover

[3] Two-year fiscal multipliers are the benchmark estimates in the literature for three reasons. First, two years is 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.

[4] About 80% of the goods spending is allocated to contractors belonging to the manufacturing sector, and 55% of the services spending is directed to contractors in the services industry.

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

or improving innovation. If firms producing services differ from those producing goods, a government shock may affect their productivity gains differently and generate heterogeneous multiplier effects.

Labor intensity is an important factor in driving the differences by spending category. Industries that produce services 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 of 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.

By contrast, we do not find strong support for tradeability and productivity gains as the main drivers of the differences in the local multipliers by type of spending. We implement two strategies to test the contribution of tradeability and geographic spillovers in explaining the differences in the estimates between services and goods spending. The first test consists in replicating the analysis at a higher geographic aggregation —at state level—in which the geographic misallocation of procurement contracts is less relevant. The second test quantifies the "outflow" effects on neighboring locations. Both tests indicate that geographic spillovers do not explain the differences in the multipliers by category of spending.

At the same time, we explore if the two types of spending shocks have different impacts on firms' entry rates, exit rates, or innovation outcomes. Our tests suggest that shocks to services spending increase firms' entry rates, decrease exit rates, and hurt innovation activities. Shocks to goods spending only positively affect entry rates, but these effects are small. The results suggest the effects of government spending on productivity improvements are, if any, relatively small.

We contribute to the literature by documenting the differences in fiscal multipliers by the type of spending, whereas previous research has emphasized 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 in 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 output responses 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 shifted 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 helps to identify important determinants of the fiscal multiplier. Our findings provide complementary insights to those already documented 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 the 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), the 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), the degree of monetary policy accommodation (Woodford, 2011), firm-size distribution (Juarros, 2020), and the 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.[6] Our findings show that the composition of the basket of products purchased by the government and their labor shares impact the effectiveness of the fiscal policy. The paper provides new insights into the design of fiscal spending and contributes to the policy debate on making fiscal stimuli more effective.

2 Data

2.1 Military Spending

We assemble new data on military procurement contracts awarded by the US Department of Defense (DoD). We harmonize military procurement contracts data from two sources: the 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 account for approximately 96% of contracts awarded by the DoD.[8]

Our data have unique advantages compared to previous studies (Nakamura and Steinsson, 2014; Dupor and Guerrero, 2017; Auerbach et al., 2020). The data contain detailed information, including the contract identification number, 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, which occurs when the contract is either awarded or modified.[9] The modifications of existing contracts could consist of 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 terms. For comparability over time, we convert the nominal transaction value into real values using the US Bureau of Labor Statistics' Consumer Price Index.

The Product Service Codes and the Federal Supply Codes play a crucial role in classifying fiscal spending

[6] Studies on the infrastructure multiplier 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.

[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 that used in previous studies (Nakamura and Steinsson, 2014; Dupor and Guerrero, 2017; Demyanyk et al., 2019). 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 military spending and other economic variables. Thus, we use the calendar year as a 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.

as 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." 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 allows us to geolocate contracts in narrow geographic areas. The available information differs between our two sources of data. In the USASpending.gov data, we know the city and the zip code of the performing firm.[11] We use these two pieces of information to identify the county where a firm operates. In NARA, the county where a firm performs the contract 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.

2.2 Economic Outcomes

We use two outcome variables to proxy changes in economic activity after a local spending shock. The main dependent variable is the number of employed persons. The employment figures are collected from the Bureau of Labor Statistics Quarterly Census of Employment and Wages (QCEW).

The second outcome variable consists of salary and wages and proxies for the labor income. These data are collected annually from the Bureau of Economic Analysis at MSA level. Salary and wages are reported in nominal terms. We convert them to real terms for consistency to absorb price variations.

Finally, we collect personal income from the Bureau of Economic Analysis. Personal income measures households' income from salary and wages, Social Security and other government programs, dividends and interest, business ownership, and other sources to normalize government spending. As in the case of salary and wages, we convert in real terms for consistency with the other monetary variables.[12]

2.3 Sample Construction

We construct the final sample by applying two sets of filters. From the universe of DoD procurement contracts, we drop those 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,[13] or the industry in

[11] 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 we adjust a marginal share of contracts differently from Demyanyk et al. (2019). That is 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.

[12] We also replicate the paper's results using personal income as the dependent variable. These results are not reported in the paper, but they are available upon request.

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

which the contractor operates.[14] We apply these filters because the absence of any of these variables prevents 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.[15]

The second set of filters is applied to the data after aggregating 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 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 the locality aggregated level rather than per capita.

The final sample of procurement contracts is 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% of the total number of contracts and the aggregate spending. These shares are high for both types of spending. For spending on goods, the share of contracts and spending in the sample accounts for about 95% of the total contracts and spending on goods. For spending on services, our sample includes 85% of the contracts worth 92% of the total spending on services.[16]

3 Descriptive Statistics

Figure 1 reports spending over $1979-2019$. Panel A shows the evolution of aggregate military spending. There are three sharp rises in 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 recent years driven by the escalations in military buildups and competition with Russia and China.

Panel B reports the shares of military spending by category.[17] The novel division of spending across the two categories highlights a reallocation of federal government spending from goods to services between the 1980s and the 1990s. As the product classification has been harmonized over the sample period, changes in the shares are not due to changes in the product definition. The reallocation of the spending seems to reflect the structural transformation the US economy experienced in the late 1980s. Until the beginning of the 1980s, the US economy had a large fraction of goods production, and the spending on goods was more than 70% of the value of the procurement contracts. Starting with the 1990s, the US economy has become a service-producer economy, and the spending has been split between goods and services with similar shares.

An interesting comparison is between the composition of government spending and private consumption expenditures. Figure B.1 in Appendix B shows the shares of private expenditures in goods and services. Similar to the DoD, private consumers have also increased the share of their expenditures allocated to services,

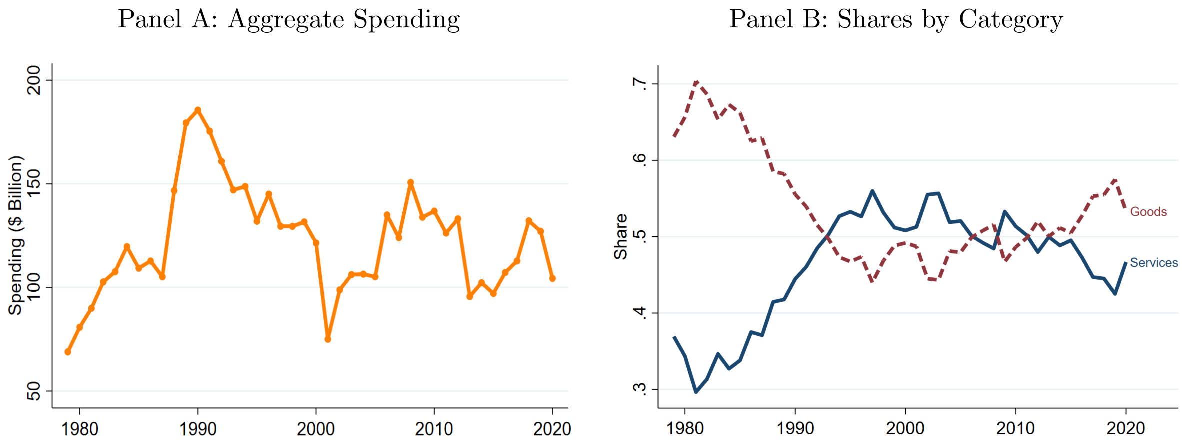
[14] For a small share of contracts, we impute the industry in which the contractor operates by using the product codes. We exclude contracts for which the industry is missing after the imputation. Appendix A.4 describes the imputation procedure.

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

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

[17] The first row of Table 1 reports the average shares.

Figure 1: Procurement 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.

showing an upward trend in the share of services consumption. Private consumers started to reallocate their spending from goods to service purchases at the beginning of the 1970s. The fraction of services expenditures is higher than the shares reported in Figure 1. This finding highlights that government spending is more goods-driven than private consumption, and the reallocation between categories of spending takes longer. This finding adds to the results of Cox et al. (2020) that emphasize the existence of a sectoral bias in government spending. In other words, the share of government spending in each sector differs substantially from the share consumers spend on the goods and services of that sector. We also find the importance of goods and services purchases differs substantially between government spending and private consumption.

The second row of Table 1 reports descriptive statistics on contract characteristics by category. The DoD awarded most of its procurement contracts (91%) to purchase goods, and only a small share (9%) is directed to the acquisition of services. These results imply that 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 on services. In the case of spending for goods, contracts at the top decile of the distribution are over 40 times greater than the median contract. The 90%-to-50%-percentile ratio is smaller for spending on services, with the contracts at the top decile 11 times greater than the median contract.

The last row of Table 1 shows the share of spending allocated to contracts in the top percentile. Amongst contracts to procure goods, the top 1% of contracts accounts for 86% of the 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 percentile is around half of the services spending, implying a more equal allocation of spending across contracts. Our findings are

Table 1: Descriptive Statistics

	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 1% (%)	72	86	50

Note: The national level statistics are calculated by aggregating the microdata on military procurement contracts available from NARA and USASpending.gov. DoD procurement contracts are 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 334 MSAs in the final sample.

complementary to Cox et al. (2020), which show that spending is granular and concentrated among a few firms. Their analysis is based on the contract recipients. The granularity in recipients does not necessarily imply that the actual performance of the tasks is granular. It could be the case that recipients allocate several tasks of a contract to different performers more evenly located across the country. 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 and highlight that, in addition to granularity in contract recipients, there exists a granularity in contract performance.[18] Similarly to Cox et al. (2020), we also document substantial variation in the range of contract values, and we emphasize significant differences in the distributional features of contract size across categories of spending.

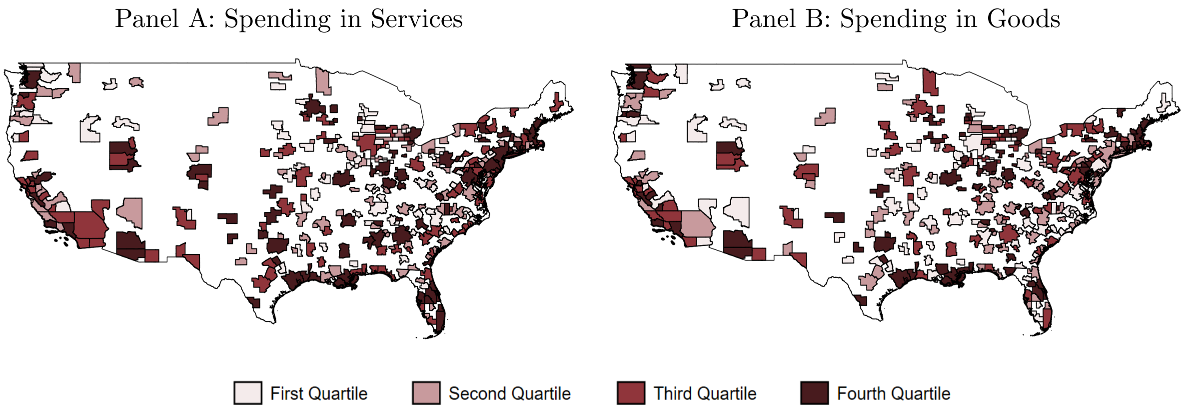
As the identification strategy exploits the cross-sectional variation in spending, Figure 2 explores the geographic heterogeneity in the allocation of spending across MSAs. The figure shows the quartile to which an MSA belongs based on the average value of the military spending that it received over 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 account for more than 30% and 45% of spending on goods and services, respectively. Most 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 for both categories. The heat-maps suggest that most MSAs in the Midwest are in the top two quartiles of spending on 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.[19]

These descriptive statistics provide evidence that although the DoD signs a large number of contracts,

[18] We cannot directly test 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.

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

Figure 2: Geographic Distribution of the Procurement Spending



Notes: The quartile to which an MSA belongs is assigned based on the average military spending in real terms that the MSA receives over 1979 − 2020.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 largest share of government procurement spending is captured by a handful of contracts. Furthermore, these contracts are not equally distributed across MSAs. These findings imply the distribution of spending at both contract and MSA-level is skewed to the right.

4 Empirical Strategy

4.1 Specification

Our empirical strategy builds on the works of Nakamura and Steinsson (2014), Dupor and Guerrero (2017), and Auerbach et al. (2020). We exploit variations in military procurement spending across time and localities to estimate local fiscal multipliers.[20] 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 $k = \{0, \ldots, 4\}$. 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 in location l, $Y_{l,t-1}$. 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.[21]

The main coefficient of interest is β^k which quantifies the multipliers in a window of k years. Since we normalize spending by personal income, in the case of the employment multiplier, the interpretation of β^k is the change in the employment growth rate produced by a one-percentage-point increase in government

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

[21] 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 the state level. Our results (available upon request) remain unaffected by adopting their clustering decision.

spending (normalized by personal income) at the local level.[22]

We augment the benchmark specification to examine the heterogeneous effects of the different government spending 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.

4.2 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, personal income, or employment. For example, private firms in locations with higher military contracts may have exerted more effort through lobbying to win those contracts. Our estimate may be biased if such firm-specific shocks affect the outcome variables.

We follow the instrumental variable research design proposed by Nakamura and Steinsson (2014) to obtain causal estimates of the fiscal multipliers. 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 that certain localities have in 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 captured by location l, respectively.

[22]Our empirical strategy provides estimates for the local fiscal multiplier. Estimates of the local fiscal multiplier cannot be straightforwardly translated into a national multiplier. This is the case for two reasons. First, spillover effects could 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.

4.3 Identification assumptions and threats

Although this instrumental variables research design has become the gold standard in the literature that studies cross-sectional fiscal multipliers, it is not free of identification threats that may bias our coefficient estimates. 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 induced by 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 place of performance are used in the intermediate steps of production. In that case, we would geographically misallocate part of the outsourced military spending. This measurement error will create an attenuation bias in our estimates of the local fiscal multiplier. For the purpose of our analysis, a more serious threat in interpreting the relative size of the product-spending multipliers would arise if the measurement error is systematically correlated with a specific spending category. As physical goods are more tradeable than services, the measurement error could be correlated with spending on goods. We rule out the importance of this measurement error in Section 6.2.

The second threat to identification is the potential presence of geographical spillovers. This would imply a violation of the stable unit treatment value assumption. More concretely, this violation would arise if defense spending in a specific MSA affects neighboring localities that 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 increases in the neighboring localities due to input-output linkages, the spillovers are positive. In this case, our estimates would suffer from a downward bias. When the increase in spending affects factor prices and the allocation of production factors across MSAs,[23] the spillovers could be negative, and our estimates would suffer from an upward bias. In Section 6.2, we provide evidence that our estimates are biased downward, and they can be considered as lower bounds.

The third concern relates to recent developments in the econometric literature on the drawbacks of shift-share instruments (Goldsmith-Pinkham et al., 2020; Borusyak et al., 2022; Adao et al., 2019). Our empirical strategy uses 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, which 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

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

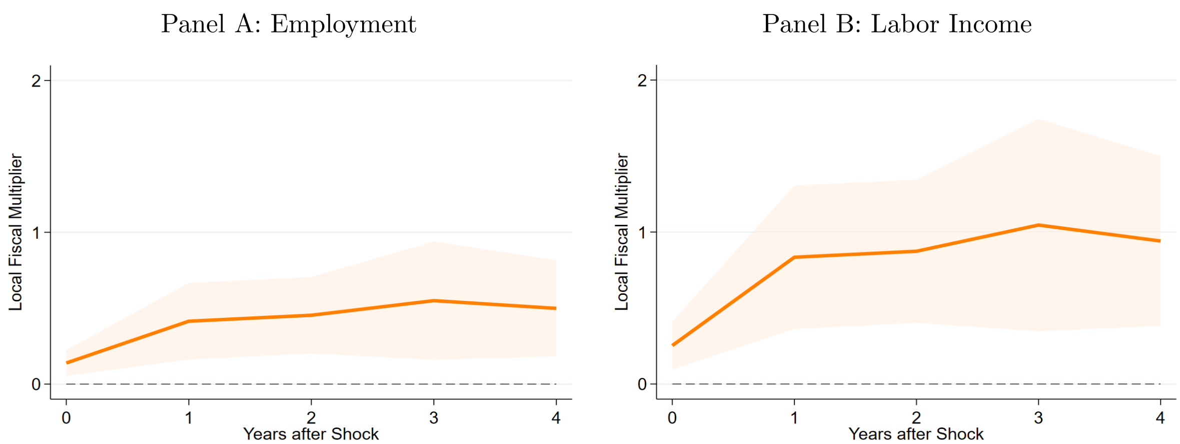
procurement contracts experience a different economic trend than MSAs with a low comparative advantage. As both outcomes and instruments are measured in changes, our research design purges the instrument from the potential correlation between the shares and the unobserved local economic trends. Furthermore, 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.

5 Main Results

5.1 Local Fiscal Multiplier by Category

Figure 3 plots the estimates of the employment and labor income responses at different horizons after a local shock in aggregate spending.[24] Table B.1 in Appendix B reports the point estimates.

Figure 3: Local Fiscal Multipliers - Aggregate Spending



Notes: All panels report the estimates from equation (1) for different outcome variables. 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 shaded areas represent the 90% confidence intervals. For graphical comparability, the y-axis is on the same scale.

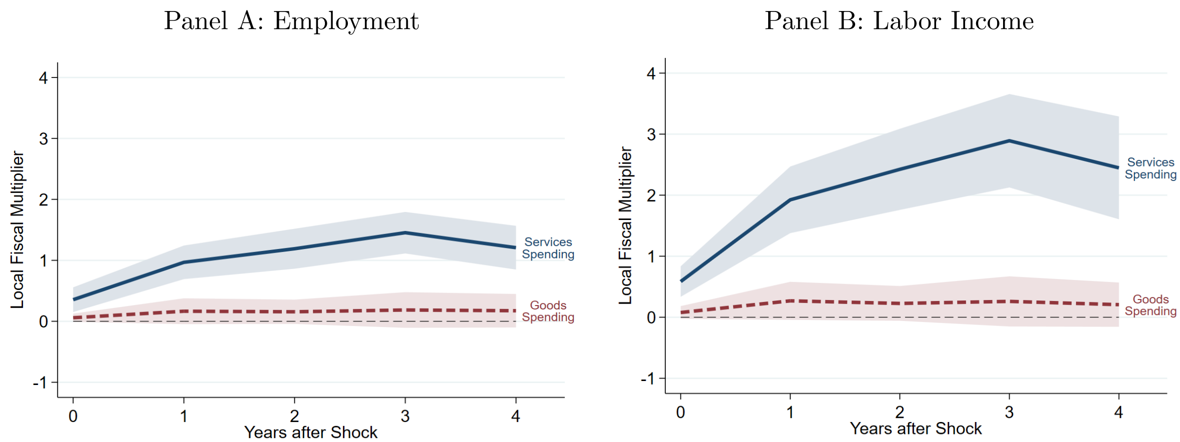
The estimates in Panel A suggest a positive and significant employment multiplier.[25] Our estimates for aggregate spending are reasonably close to the effects estimated by other authors. Specifically, Auerbach et al. (2020) report the on-impact employment multiplier for local spending shocks at the CBSA level to be 0.186. This estimate is very close to ours (0.139). However, the discrepancy between the two sets of estimates increases at further horizons, with ours being greater than theirs. This discrepancy is due to two sample differences. First, our period is from $1979-2019$, while their sample starts in the early 2000s.

[24] As output by MSA is not available before the 2000s, consistent with previous studies, we report the employment multiplier rather than the output multiplier. Focusing on employment rather than the 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 employment to output multiplier is to divide output per worker by the cost per job.

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

There is a large literature arguing that the national fiscal multiplier has decreased over time (Blanchard and Perotti, 2002).[26] 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 some spending spills into other geographic areas attenuating the estimates of local fiscal shocks. Similarly to Panel A, Panel B of Figure 3 also shows a positive and significant effect of procurement spending on labor income at any horizon. A one-percentage-point increase in government spending increases labor income by 0.83% after two years and above 1% at the three-year horizon.

Figure 4: 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 is on the same scale.

Now, we turn to the core results of the paper. We explore the effects of shocks on two types of government spending: spending in goods and services. Figure 4 plots the estimated effects. Table B.2 in Appendix B reports the point estimates. Panel A shows the employment local multipliers after shocks to goods and services spending. We observe a remarkable 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% level. A one-percentage-point increase in normalized services spending by personal income generates an increase in employment of 0.36% on-impact and 1.19% two periods after the shock. On the other hand, goods spending shocks have small effects on the employment multiplier. The estimates are statistically non-significant and range between 0.059% at impact and 0.186% at the two-year horizon, respectively.[27]

The increase in the demand for products leads firms to hire more workers. As firms hire more workers (or similarly, workers work more hours), the labor income increases and households can increase their consumption. The increased demand for products further increases labor demand, leading to a multiplier effect.

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

[27] We also test whether the two sets of coefficients are statistically different from each other. We reject the null hypothesis that the two estimates are equal at the 5% significance level.

Our results suggest that this virtuous circle only activates after shocks to services spending but not to goods spending. Panel B of Figure 4 explores the effects of the two types of shocks on labor income. There is a noticeable sharp and sizable increase in wages and salary after a shock to services spending. Specifically, a one-percentage-point increase in normalized services spending generates a rise in labor income up to 2.5% within four periods after the shock. By contrast, the response of wages to goods spending shocks is very different. Labor income only slightly increases after goods spending shocks.[28]

To grasp a better sense of the size of our estimates, we calculate the jobs creation for $100,000 spent in procurement contracts. The number of newly-created jobs per $100,000 in procurement spending is calculated as

$$\Delta E_{t+k} = \beta_i^k \frac{1}{TN} \sum_{t=1980}^{T} \sum_{l=1}^{N} \frac{100,000}{Y_{l,t-1}} E_{l,t-1} \quad \forall i \in \{s, g\}. \tag{5}$$

Procurement contracts of $100,000 value generate 0.32 (s.e. 0.12) new jobs at impact and 1.03 (s.e. 0.35) at the two-year horizon.[29] The number of new jobs created substantially differs between the two types of spending. $100,000 in procurement contracts that acquire services create 0.81 (s.e. 0.28) new jobs at impact, and 2.71 (s.e. 0.45) after two years. By contrast, $100,000 used to purchase goods generate 0.13 (s.e. 0.09) new jobs at impact and 0.36 (s.e. 0.28) after two years. These estimates are in line with the literature. Chodorow-Reich (2019) reports a range from 0.76 to 3.80 of new jobs created at the two-year horizon estimated on variation due to the 2009 American Recovery and Reinvestment Act. The number of new jobs created by services spending is close to the upper bound of that range, while the new jobs generated by goods spending is slightly below the lower bound. This finding suggests that the composition of the government purchase bundle could help explain the large range of newly-created jobs reported in the literature.

We also investigate whether the employment responses occur in the industries that receive the procurement spending or the effects spill over to industries that do not receive the spending. To explore this aspect, we collect employment data by county and industry,[30] and we aggregate them by MSAs and major NAICS groupings.[31] As industries differ in their size, for better comparability across industries, we report the newly-created jobs rather than employment growth rates. We calculate the newly-created jobs by first running specification (2) for each industry separately, and we use equation (5) to calculate the creation of jobs per 100K spending in either services or goods. Figure 5 summarizes those calculations.

First, our estimates suggest both types of spending generate positive and significant direct effects on the

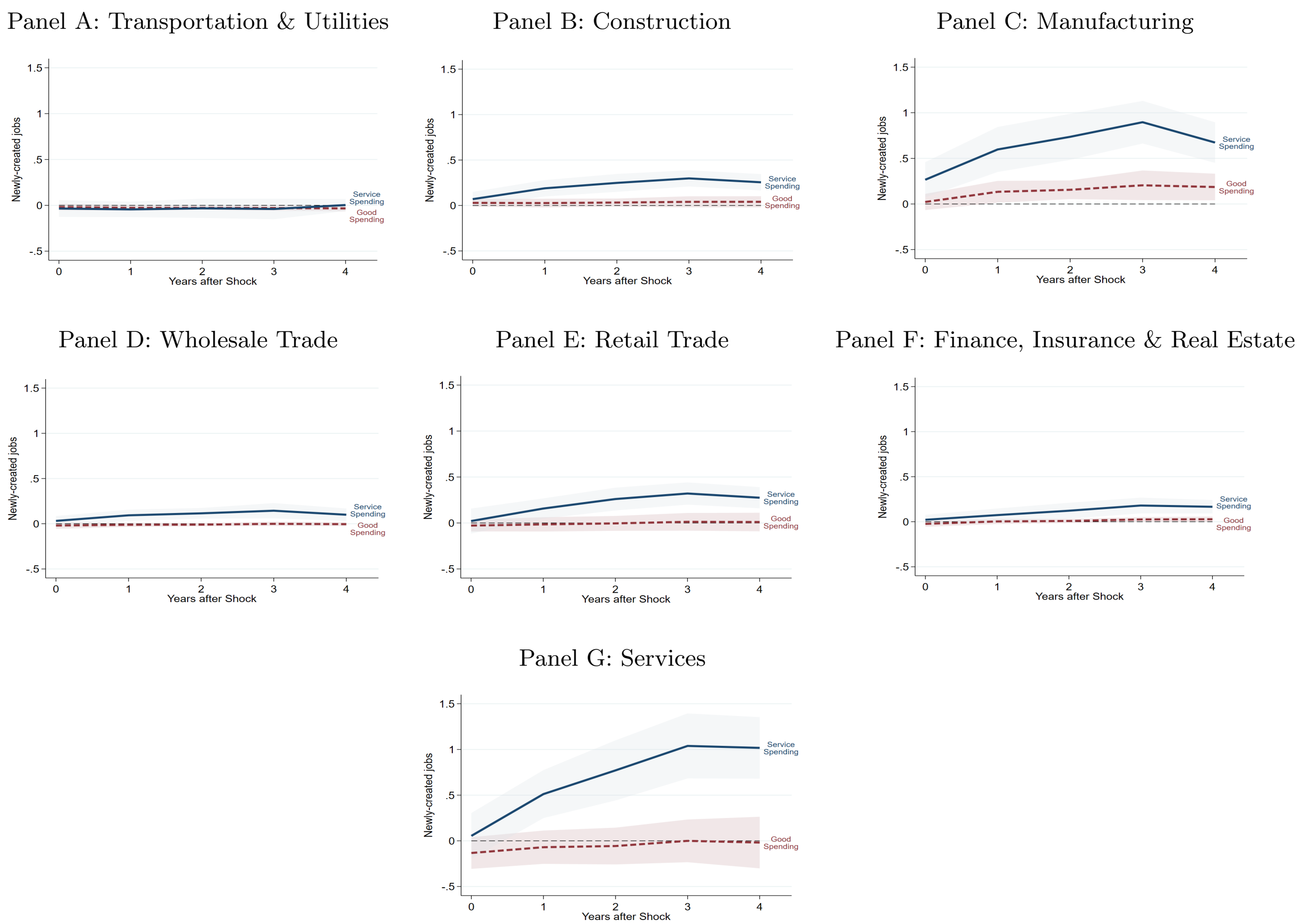
[28] Although not reported, the responses to personal income closely track the responses to labor income, with the only difference being that the effect size is smaller for personal income. This stems from the fact that a combination of changes in labor income and dividends gives changes in personal income. We check the responses of dividends to the two types of shocks, and we find that firms do not adjust dividends paid to households after spending shocks.

[29] The standard errors are calculated using the delta method.

[30] Eckert et al. (2020) develop an algorithm to impute several suppressed employment by industry from the County Business Patterns. The missing employment figures are suppressed by the US Census Bureau to protect confidentiality.

[31] We use the grouping reported by Nakamura and Steinsson (2014). We exclude Agriculture and Mining because several MSAs have missing employment figures for these sectors due to confidentiality risks. The inclusion of all major NAICS groupings would result in having only 294 MSAs with non-missing information. We have replicated the analysis by including all industries but restricting the set of MSAs to 294. The results remain unchanged, and they are available upon request.

Figure 5: Newly-created Jobs by Industry



Notes: All panels report the estimates from equation (2) for different industries separately. 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 is on the same scale.

industry that receives most of that spending.[32] Specifically, goods spending shocks generate 0.15 new jobs at the two-year horizon in the manufacturing sector, and services spending shocks create 0.77 new jobs after two years in the services sector. Furthermore, for both types of spending, the employment responses are the largest in the industry that receives the largest share of the procurement contracts. Second, our findings suggest that indirect effects consisting of positive employment responses that spill to other sectors occur only after services spending shocks. Precisely, services spending generates positive employment responses in all other industries, with the exception of the "Transportation and Utilities" sector. The employment response is large in the manufacturing sector—0.73 newly-created jobs after two years—and moderate in the other industries—between 0.1 and 0.25 newly-created jobs at the two-year horizon. By contrast, goods spending shocks do not generate significant employment responses in sectors other than manufacturing. Alonso (2017)

[32] About 80% of the goods spending is allocated to contractors belonging to the manufacturing sector, and 55% of the services spending is directed to contractors in the services industry.

and Bouakez et al. (2020) show that the direct effect of government spending shocks is relatively small and that most of the aggregate impact comes from the indirect effect of households' responses to increases in their income. Consistently with this view, Figure 4 documents labor income increases only after services spending shocks. The marginal increases in labor income after goods spending shocks causes the alleviation of the multiplier effects and the absence of indirect effects. These results highlight the relevance of the transmission of fiscal spending to the household's budget constraint through changes in labor income. We will dig further into the importance of this transmission mechanism in explaining the difference between services and goods spending employment local fiscal multipliers in Section 6.1.

5.2 Robustness

One concern is that the specification ignores the dynamics of spending and employment for the years between the periods 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.3 in Appendix B. All coefficients are quantitatively similar to the ones presented in our main specification, i.e., services spending generates large employment multipliers. In contrast, the estimated effects of a shock to goods spending are small and, in many cases, non-significant.

A second concern is linked 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 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.4 in Appendix B.[33] The employment multiplier is positive and significant only after shocks in services spending.

A third concern relates to the distribution of government contracts. Government procurement spending is granular, and a relatively small share of contracts captures a large share of the total spending. As contracts are concentrated in a small number of firms, our results could be driven by the nature of the specific contractors rather than the differences between goods and services. We rule out this possibility by excluding

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

the top percentile of contracts by value in each location and period and re-estimating the specification (2). The estimates are reported in Table B.5 in Appendix B. The new set of estimates highlights the same patterns in the responses to services and goods spending shocks, implying the characteristics of the contracts play only a marginal role in explaining the differences between the two categories of spending.

6 Mechanisms

This section investigates three mechanisms—labor intensity, tradeability, and productivity gains—that could explain the heterogeneity in the previous estimates by category of spending.

6.1 Labor intensity

Differences in the local fiscal multipliers generated by spending on goods and services could be due to the transmission of the shock to the household's budget constraint. We show that the labor intensity of the spending is an important determinant in quantifying the effect of fiscal interventions. Intuitively, in labor-intensive industries, a larger share of government spending passes through the households. Public spending pushes the demand for products, and firms hire more workers to match the demand. Labor and personal income rise, enabling households to spend more on consumption, further pushing up product demand. In non-labor-intensive industries, this pass-through mechanism is alleviated because the reaction of labor demand to a government spending shock is smaller.

Our dataset does not contain information on the 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 employee compensation by industry.[34] We then compute labor intensity as the contribution of the employee 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 in Appendix B classifies the available industries between labor- and non-labor-intensive. Our classification aligns with economic intuition. Indeed, industries such as healthcare, education, hospitality, and food service are classified as labor-intensive, while manufacturing and retail are categorized as non-labor-intensive.

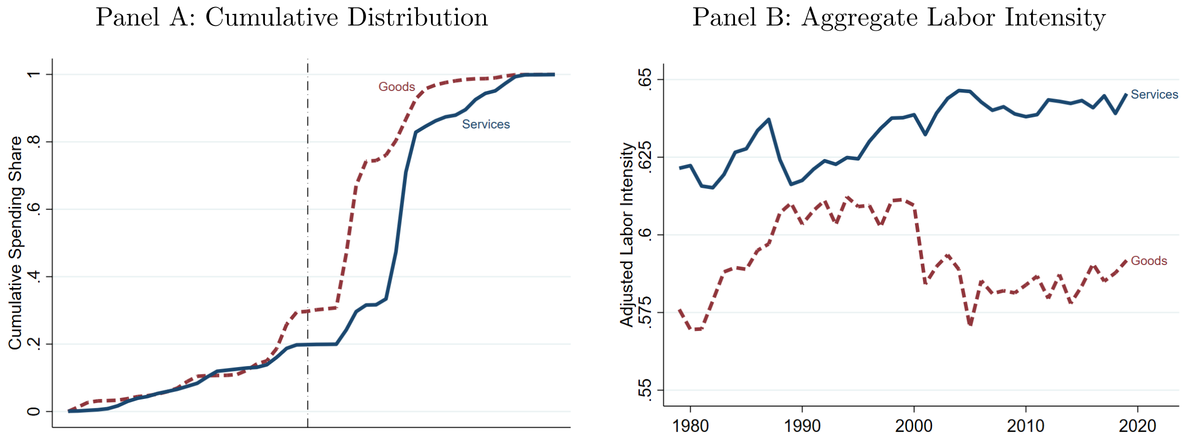
Imperfect competition may also be a confounder in our classification. Firms with more monopsony power may pay less for their employees. As our industry classification is based on employee compensation, industries with firms that use labor intensively but pay their employees little due to their monopsony 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 on the importance of imperfect competition as a confounder, we use employment concentration shares as proxies for the degree of imperfect competition, and we explore the relationship between labor intensity and measures of concentration.[35] Specifically, we regress labor intensity against the measures of concentration after having absorbed the contribution of the year and industry fixed-effects. The results in Table B.6 in Appendix B document a weak negative relationship

[34]The data contain 66 NAICS codes starting from 1997. Although the value-added and its components are not available from the BEA, we use an imputation strategy described in Appendix A.4 to assign industry to contracts awarded before 1997.

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

between employment concentration and labor intensity at the industry level. We interpret these findings as evidence that monopsony power is not an important confounder of our industry classification.

Figure 6: Labor Intensity by Spending 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 labor- and non-labor-intensive industries based on the classification in Table A.3.

Figure 6 shows the distribution of spending by components and labor intensity. Panel A shows the cumulative distribution of contracts, sorting them in ascending order 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 on goods and services. The share of goods spending allocated to non-labor-intensive industries is 50% higher than the share allocated from spending in services (30% vs. 20%). Not only is the cumulative distribution at the cutoff higher for goods spending, but the cumulative distribution of services spending is also shifted to the right, implying that a larger share of contracts with services spending is directed to labor-intensive industries.

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.[36] Formally, the aggregate labor intensity measure is computed as follows:

$$AggLI_{i,t} = \sum_{j \in \mathcal{J}} \omega_{i,j,t} \overline{LI}_j \quad \forall i \in \{s, g\} \tag{6}$$

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

[36]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.

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 6 reports the measures of aggregate labor intensity. The aggregate labor intensity is higher for services spending than for 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.[37]

These results point out that the composition of the spending in the two categories differs in terms of the labor intensity of the industries that receive the contracts. This heterogeneity in 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 in determining the differences in the local fiscal multipliers across categories of spending by separately considering the components of spending for each category that go to labor- and non-labor-intensive industries. We formally test this hypothesis by implementing the following specification:

$$\frac{v_{l,t+k} - v_{l,t-1}}{v_{l,t-1}} = \sum_{i \in \{s,g\}} \sum_{p \in \{L,N\}} \gamma^k_{i,p} \frac{G^{i,p}_{l,t+k} - G^{i,p}_{l,t-1}}{Y_{l,t-1}} + \alpha^k_l + \delta_{t+k} + \varepsilon_{l,t+k} \tag{7}$$

with the independent variables instrumented by

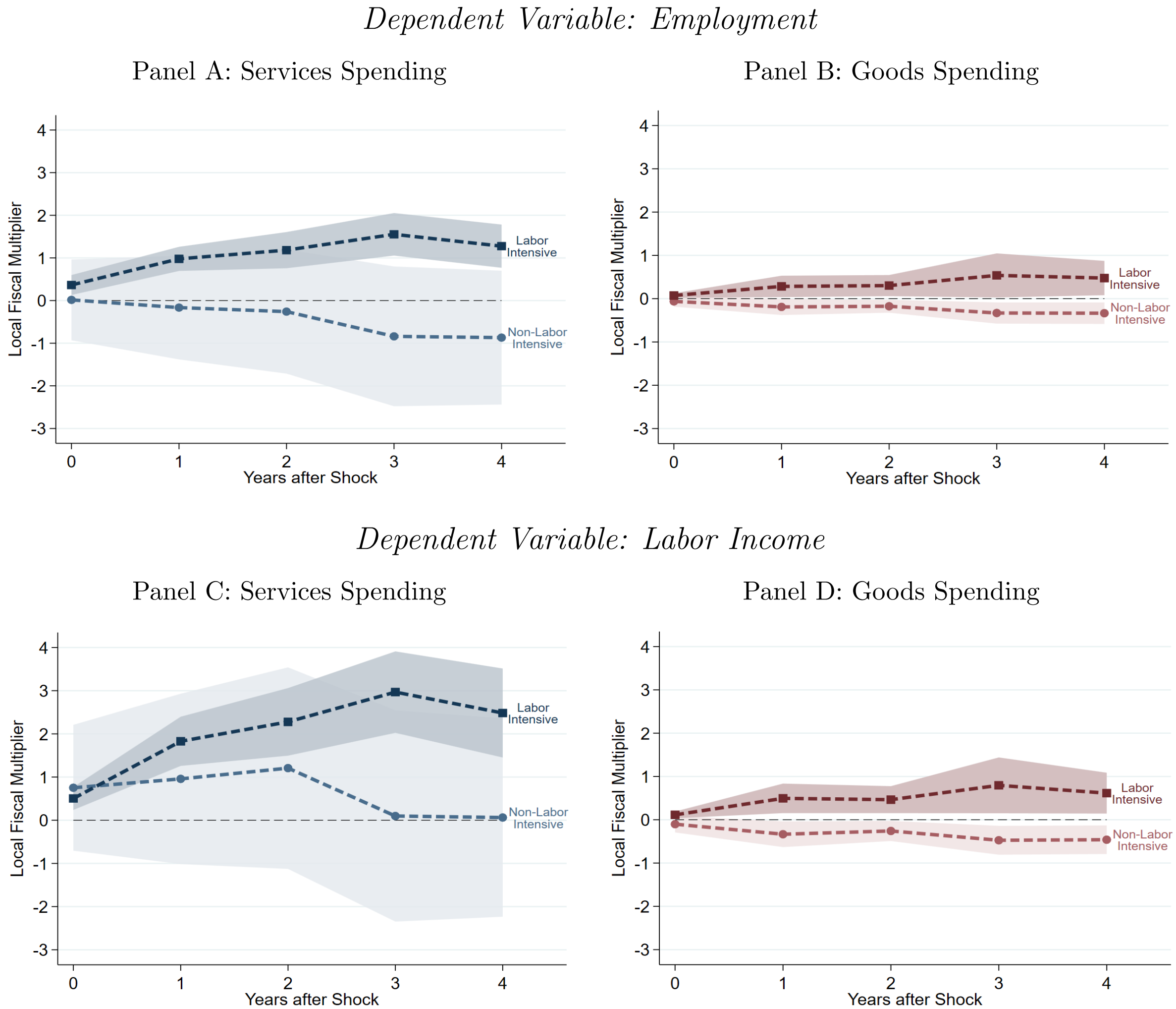
$$Z^{i,p}_{l,t+h} = s^{i,p}_l \frac{G^{i,p}_{t+k} - G^{i,p}_{t-1}}{Y_{l,t-1}} \quad \forall i \in \{s, g\} \text{ and } \forall p \in \{L, N\}, \tag{8}$$

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 labor income in Figure 7 provide support for the hypothesis that the heterogeneous 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.[38]

Spending on both services and goods directed to labor-intensive industries generates positive and highly significant effects on employment. Panel A of Figure 7 shows that the estimates for the employment multiplier for labor-intensive services spending range between 0.36% and 1.5%. Although the employment multiplier coming from goods spending, presented in Panel B, is also positive and significant, the size of this effect is significantly smaller, with estimates between 0.07% and 0.54%.[39] The effect of spending in non-labor-intensive industries is negative for both categories of spending (goods and services), 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.[40] It is worth noticing that the estimates for

[37] We cannot assess how sizable 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.
[38] Table B.7 in Appendix B reports the point estimates.
[39] We test the null hypothesis that the two coefficients are the same, and we reject this hypothesis at any horizon.
[40] There is no consensus on whether government spending crowds-in or crowds-out private consumption. There are advocates

Figure 7: Local Fiscal Multipliers - Labor Intensity



Notes: All panels report the estimates from equation (7). The instruments are computed as in equation (8). 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 is on the same scale.

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 each other out. Panels C and D of Figure 7 present the estimated effects on labor income. Spending directed to labor-intensive industries generates increases in labor income, while the effects of spending in non-labor-intensive industries are either negative or non-significant.[41] These results are consistent with the findings previously discussed. In other words, when a large share of the spending passes through the household in the form of increases in labor income, households increase their consumption and firm production, generating sizable employment responses.

for both sides, namely, a crowding-out effect (Barro, 1981) and a crowding-in effect (Mountford and Uhlig, 2009).

[41] The estimates for effects on employment and labor income in non-labor-intensive industries after services spending shocks are less precisely estimated because the share of spending directed to these industries is relatively small.

To summarize, the previous analysis provides support that services and goods spending differ in the labor intensity used in production and that labor intensity matters for quantifying the fiscal multiplier. Overall, these findings imply that labor intensity is an important driver of the differences in local fiscal multipliers between services and goods spending shocks.

6.2 Tradeability

Goods are more tradeable than services.[42] If the production of goods occurs in the neighboring locations, one could observe smaller 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 in production are done by sub-contractors outside the performance location, we would geographically misallocate part of the spending. In addition, the definition of the place of performance varies slightly across categories of products.[43] The measurement errors in allocating contracts to MSAs could affect our previous findings. In this sub-section, we show that even after accounting for these factors, our main findings do not change. These results imply that the heterogeneous reactions of employment to different types of spending do not depend on the products' tradeability or geographic spillovers.

We undertake two strategies to quantify the importance of the geographic allocation of contracts. The literature has argued extensively that the geographic misallocation of contracts becomes less important as one moves to larger geographic aggregation. Specifically, it becomes a minor issue at the state level. Isard (1962) argues that geographic disaggregation at the state level does not contain significant measurement errors. Nakamura and Steinsson (2014) use US Census Bureau data from 1963 to 1983 on shipments from defense industries to the government to verify that, on average, there is a one-for-one relationship between prime contracts allocated to a state and the actual shipments from that state. 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 those using state-level data. If tradeability and geographic allocation of spending to MSAs are not important drivers of our estimates, one should expect the state-level analysis to lead to the same conclusions as the MSA-level analysis. The state-level estimates are in Table 2. We do not find any effect of goods spending on employment, while the effects of services spending remain sizable and significant. As output data are

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

[43] The location of the majority of manufacturing contracts reflects the location of the plant where the product is finally assembled or processed. The location of construction contracts corresponds to the location where 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.

Table 2: 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)
F-stat	9.945	17.988	19.649	11.219	9.246
	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)
F-stat	9.945	17.988	19.649	11.219	9.246

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.

available at state-level since 1979, we also report the output multiplier.[44] We find strong positive effects on output following shocks to services spending and no effects due to changes in the goods spending. Similar to Demyanyk et al. (2019) and Auerbach et al. (2020), the effects are larger when we use 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 spills into/from other areas. The potential measurement error in subcontracting outside an 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 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:

$$\frac{\tilde{Y}_{l,t+k} - \tilde{Y}_{l,t-1}}{\tilde{Y}_{l,t-1}} = \tilde{\beta}_g^k \frac{G^g_{l,t+k} - G^g_{l,t-1}}{\tilde{Y}_{l,t-1}} + \tilde{\beta}_s^k \frac{G^s_{l,t+k} - G^s_{l,t-1}}{\tilde{Y}_{l,t-1}} + \alpha_l^k + \delta_{t+k} + \varepsilon_{l,t+k}, \tag{9}$$

[44] 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.

Table 3: 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)
F-stat	0.248	18.771	1.103	1.869	0.865
	Dependent Variable: Labor Income				
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)
F-stat	0.248	18.771	1.103	1.869	0.865

Notes: All panels report the estimates from equation (9). The instruments are computed as in equation (10). 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.

where $\tilde{Y}_{l,t+k}$ is the outcome variable for the neighboring locations of l. We define the neighboring locations to l as MSAs whose center is located within a 50 miles radius from the center of l.[45] Similarly, as before, we instrument the main regressors with a shift-share instrument defined as

$$Z^g_{l,t+h} = s^g_l \frac{G^g_{t+k} - G^g_{t-1}}{\tilde{Y}_{l,t-1}}; \quad Z^s_{l,t+h} = s^s_l \frac{G^s_{t+k} - G^s_{t-1}}{\tilde{Y}_{l,t-1}}. \tag{10}$$

Results for employment and labor income are reported in Table 3.[46] The findings show that the higher tradeability of goods cannot explain the differences in the multipliers.[47] On the one hand, the estimates point out that 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

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

[46]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.

[47]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.

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 that the estimates we reported in Table B.2 are downward biased, implying that they should be considered as lower bounds of the effects of government spending.

6.3 Productivity Gains

Government spending may improve productivity. Increases in spending may reduce the uncertainty about future profits, ease credit constraints leading to higher turnover rates of firms, or generate a faster growth of incumbents. These channels, typical in growth models with firm heterogeneity, may generate productivity improvements.[48] If different types of spending generate different firm dynamics, the heterogeneity in the fiscal multipliers could be partially explained by differences in productivity gains.

Table 4: Entry and Exit Rates

	(1) t	(2) t+1	(3) t+2	(4) t+3	(5) t+4
	Dependent Variable: Entry Rate				
Services Spending	-0.180	0.121	0.394	0.637**	0.602**
	(0.399)	(0.281)	(0.256)	(0.293)	(0.238)
Goods Spending	-0.070	0.224*	0.155*	0.168	0.183
	(0.117)	(0.135)	(0.084)	(0.108)	(0.125)
F-stat	6.325	12.472	14.034	17.648	29.439
	Dependent Variable: Exit Rate				
Services Spending	-0.574	-1.183**	-1.102***	-1.104***	-0.775***
	(0.560)	(0.542)	(0.317)	(0.295)	(0.256)
Goods Spending	0.159	-0.232	-0.058	-0.194	-0.224
	(0.240)	(0.189)	(0.187)	(0.194)	(0.195)
F-stat	6.325	12.472	14.034	17.648	29.439

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. The first stage consists of the same specification and sample as in the main regression, therefore, the F-statistics for the first stage are the same as those reported in Table B.2.

Government spending may lower entry costs, ease credit constraints, incentivize firm entry, and improve the allocation of production factors. As goods producers usually face larger entry costs than services producers, spending directed to goods producers may not be sufficient to lower the entry costs. In this scenario,

[48] Recent empirical studies have provided insights into the relationship between fiscal policy, business dynamism, and growth. Ferraz et al. (2015) show that firms who win government procurement contracts grow more than their competitors. Lewis and Winkler (2017) argue that net firm entry rises after an expansion in US government spending. Slavtchev and Wiederhold (2016) show that the 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.

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 4 reports the impact of the different types of spending on the establishment entry rate.[49] The establishment entry rate increases a few periods after the occurrence of a shock to the services spending. Consistent with the higher entry costs hypothesis, the effects are smaller after shocks to goods spending.[50]

Government demand may also change firms' incentives over exiting the market and keeps low-productivity firms active that otherwise would have left the market. If low-productivity 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 bottom panel of Table 4 explores the effect of the different types of spending on the exit rates. The estimates suggest a decline in the exit rates after shocks to services spending. Shocks to the goods spending have marginal effects on the exit rates.

Table 5: 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)
F-stat	7.628	16.917	17.043	26.988	28.410
	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)
F-stat	7.628	16.917	17.043	26.988	28.410

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.

The last source of productivity growth comes from the improvement of technology via innovation activities. We proxy for innovation with patenting activities.[51] The estimates for both types of spending reported

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

[50] We dig further into this result, and we explore the impacts of 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-sized firms. Indeed, the increase in entry rates is driven by the entry of small-size establishments. Although the results are not reported in the paper, they are available upon request.

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

in Table 5 show a decline in the number of patents firms applied for and the quality of the granted patents in terms of forward citations. Although the effects after both types of spending are negative, the declines are larger after shocks to services spending compared to goods spending.

The empirical estimates suggest that shocks to services spending increase the establishment entry rates, decrease the exit rates, and hurt the innovation activities of incumbent firms. Shocks to goods spending positively affect 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 low-productivity firms alive that otherwise would have exited the market. The worsening of innovative activities amongst incumbent firms may lower 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 and their effects cancel out, this contribution is expected to be relatively marginal.

7 Conclusions

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

We explore the differences in multipliers generated by two types of government purchases: goods and services. The breakdown into goods versus services is a natural distinction because the government's core decision is to allocate its budget between these two types of products.

We show that purchases of services generate positive and significant multiplier effects, while shocks to goods spending have a smaller and non-significant impact. We find that the heterogeneity 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 investigate the importance of other mechanisms, such as tradeability and productivity improvements, in explaining these differences and find little support, if any, for these alternative channels.

Our findings suggest that there is room for governments to redesign their fiscal spending and obtain higher multipliers by reallocating dollars from goods toward 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 what to purchase. Furthermore, as the composition of the spending matters, the declining effectiveness of fiscal policy in recent decades may be due to changes in the composition of the spending or an economy-wide decline in the importance of labor in production. Future works should further quantify the extent to which these structural changes have contributed to the declining effectiveness of fiscal policy at the national level.

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.

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.

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.

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.

Eckert, F., Fort, T. C., Schott, P. K., and Yang, N. J. (2020). Imputing Missing Values in the US Census Bureau's County Business Patterns. NBER Working Papers 26632, National Bureau of Economic Research, Inc.

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.

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. A. and Shapiro, M. D. (1998). Costly capital reallocation and the effects of government spending.

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.

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.

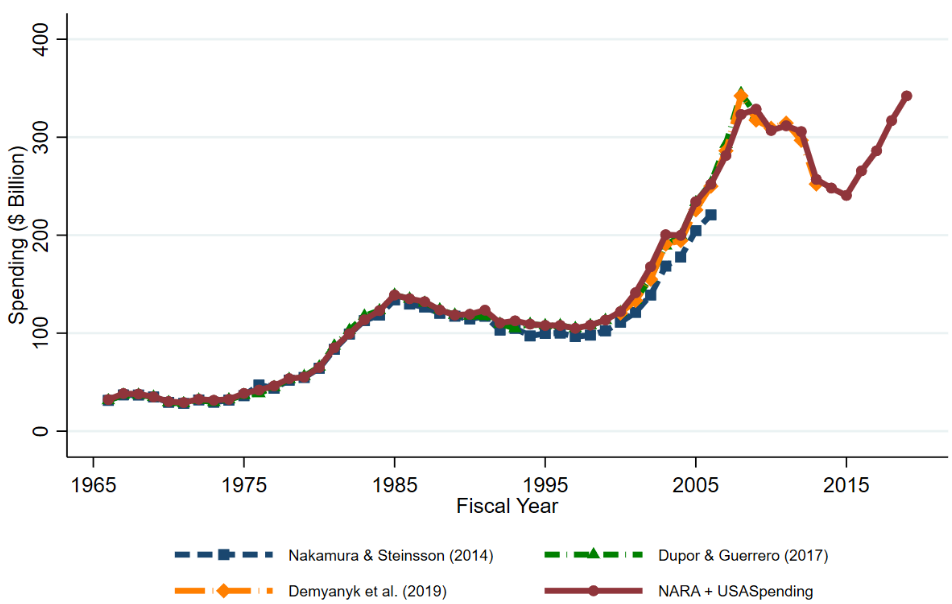
Appendix

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 aggregate military spending in nominal terms at the national level by fiscal years. We compare aggregate military procurement spending derived from our microdata with those calculated by Nakamura and Steinsson (2014), Dupor and Guerrero (2017), and Demyanyk et al. (2019). Overall, we conclude our data closely match 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 in 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 spending at the 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 data 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 either the upper or lower bound of our estimates of β. The largest discrepancy is between our data and that 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 aggregate spending, while ours are similar to the other sources of comparison. The within-R^2 is 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). Because our data are constructed using NARA until 2006, and USASpending 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 that 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), identifying spending in services, and 692 are the Federal Supply Codes (FSCs), identifying spending in goods. PSCs 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 classify an industry as a high-labor intensive industry if its average is above the median; and a low-labor intensive industry 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 is 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 estimate precisely 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 to which a contractor belongs is only available from 1988. We use the product codes to assign the industry to contracts for which this information is not reported. Intuitively, we assign a contract with no information about the industry of the contractor to the dominant industry in the production of the product code associated with the contract. We follow several sequential steps to infer 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 information. As the BEA industry classification includes several narrower industries, if there are two or more BEA industries with 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 contracts 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, accounting for 16% of the total spending, are imputed using the first step of our imputation strategy. The second and third 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 as either low- or high-labor intensive as data on value-added; its components are reported only from 1997. To overcome this, we apply the classification of a given industry based on data from 1997 to prior years. For example, if an industry is classified as low-labor intensive using the data for $1997 - 2020$, we also classify that industry as low-labor intensive 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 produces the product. To address this concern, for each product code, we compute the share of the spending covered by the industry that we assign as the 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 the assigned industry receiving a large share of this spending.

The second assumption is that the classification of industries as high or low labor-intensive is stable over time. We test the stability by re-classifying industries as high or low labor-intensive each year and comparing this new classification with the benchmark used in the paper. We find that the classification based on the annual labor intensity matches the classification based on the full sample for more than 91% of the cases, implying that 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 industries' technologies. 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 citation 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 our focus to patents with an application year starting from 1976. We do not observe the exact date on which innovation occurs. As is 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) changes in the propensity to cite, 2) the lifespan of a patent, and 3) 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 has experienced a dramatic acceleration. This increase in the use of patents is viewed as a response to the increase in the protections offered to patent holders 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 issue that impacts the over-time comparison of patent statistics is linked to the lifespan of a patent. Older patents have a longer period of 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 affects mechanically 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 less cited mechanically, independently of their innovative nature. The review process essentially takes 8 years to be fully completed.[52] Thus, we restrict the analysis to patents that applied by the end of 2006. We undertake this very cautious approach because we want to ensure that citing patents in the 5-year window have completed their review process.

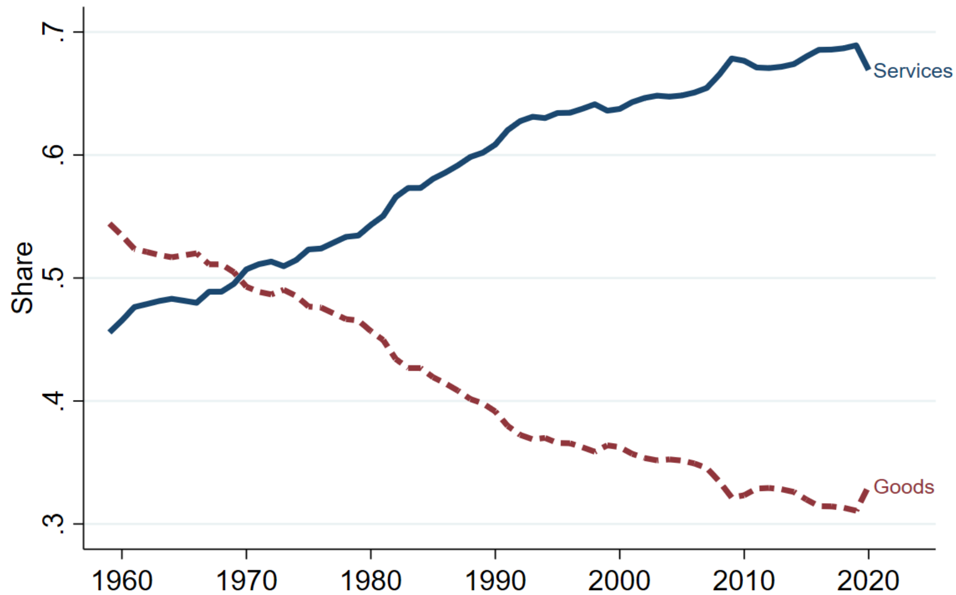
We associate a patent with an MSA by using the latitude and longitude of the address of the patent inventors. We geolocate the latitude and longitude into MSAs by using the 2010 TIGER/Line Shapefile constructed by the US Census Bureau. When patents have multiple inventors whose addresses are in different MSAs, we split those patents equally across MSAS according to the number of inventors. In the geographical aggregation, we restrict our sample to patents with at least one inventor who is located in the US mainland. The final data are 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 window after the grant year. The first statistic measures the number of patents filed at the USPTO, while the second measures the quality of the stock of patents in a specific year.

[52]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. Additional Empirical Results

Figure B.1: Private Consumption Expenditures by Category



Note: Data on the private consumption expenditures by goods and services have been downloaded from the Bureau of Economic Analysis.

Table B.1: Local Fiscal Multiplier - Aggregate Spending

	(1) t	(2) t+1	(3) t+2	(4) t+3	(5) t+4
	Dependent Variable: Employment				
Spending	0.139***	0.414***	0.453***	0.550**	0.499***
	(0.051)	(0.153)	(0.153)	(0.236)	(0.192)
F-stat	8.299	9.330	10.101	7.680	10.675
	Dependent Variable: Labor Income				
Spending	0.253***	0.834***	0.873***	1.046**	0.941***
	(0.095)	(0.286)	(0.285)	(0.424)	(0.339)
F-stat	8.299	9.330	10.101	7.680	10.675

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.

Table B.2: 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)
F-stat	6.325	12.472	14.034	17.648	29.439
Test $\beta_g^k = \beta_s^k$	[.0333]	[.0003]	[.0000]	[.0000]	[.0001]
	Dependent Variable: Labor Income				
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)
F-stat	6.325	12.472	14.034	17.648	29.439
Test $\beta_g^k = \beta_s^k$	[.0048]	[.0001]	[.0000]	[.0000]	[.0001]

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.3: 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)
F-stat	6.325	17.508	16.134	15.929	22.052
Test $\beta_g^k = \beta_s^k$	[.0333]	[.001]	[.0001]	[.0000]	[.0001]
	Dependent Variable: Labor Income				
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)
F-stat	6.325	17.508	16.134	15.929	22.052
Test $\beta_g^k = \beta_s^k$	[.0048]	[.0002]	[.0000]	[.0000]	[.0001]

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.4: 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)
F-stat	18.538	21.348	12.688	8.445	6.815
Test $\beta_g^k = \beta_s^k$	[.0189]	[.0176]	[.0549]	[.0931]	[.0662]
	Dependent Variable: Labor 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)
F-stat	18.538	21.348	12.688	8.445	6.815
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.5: Local Fiscal Multipliers - Excluding top 1% contracts by value

	(1) t	(2) t+1	(3) t+2	(4) t+3	(5) t+4
	Dependent Variable: Employment				
Services Spending	1.348***	2.435***	2.759***	2.843***	2.497***
	(0.300)	(0.522)	(0.651)	(0.752)	(0.643)
Goods Spending	0.391	0.611	0.774	0.904	1.186
	(0.417)	(0.599)	(0.839)	(0.900)	(0.959)
F-stat	9.718	7.453	3.29	5.117	13.972
Test $\beta_g^k = \beta_s^k$	[.0916]	[.0397]	[.0841]	[.1222]	[.272]
	Dependent Variable: Labor Income				
Services Spending	2.059***	4.336***	5.044***	5.145***	4.523***
	(0.750)	(1.241)	(1.393)	(1.610)	(1.425)
Goods Spending	0.506	0.684	0.680	0.832	1.176
	(0.650)	(0.884)	(1.161)	(1.238)	(1.317)
F-stat	9.718	7.453	3.29	5.117	13.972
Test $\beta_g^k = \beta_s^k$	[.1885]	[.0438]	[.0393]	[.0666]	[.1264]

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.6: 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.7: Local Fiscal Multipliers - 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)
F-stat	5.679	8.652	9.84	11.611	15.717
	Dependent Variable: Labor Income				
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)
F-stat	5.679	8.652	9.84	11.611	15.717

Notes: Both panels report the estimates from equation (7). The instruments are computed as in equation (8). 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.