

Heterogeneous Purchases, Heterogeneous Multipliers*

Umberto Muratori[†]

Pedro Juarros[‡]

Daniel Valderrama[§]

May 20, 2026

Abstract

Do the effects of local fiscal multipliers vary based on the type of government purchases? We find that government purchases of services generate larger increases in local employment, labor income, and output than purchases of goods. This variation stems from key factors: differences in sectoral characteristics —particularly centrality and markups— between good-producing and service-producing industries, and distinct impacts on business turnover and productivity. Service purchases generate larger aggregate productivity gains than goods purchases because firm turnover is concentrated in high-growth sectors, suggesting that extensive-margin reallocation and productivity improvements amplify fiscal responses more strongly for services. Our results highlight that fiscal policy effectiveness depends not only on scale but also on the composition of spending, with important implications for the design of government purchase programs.

JEL Codes: E10, E62, H57, R12

Keywords: Heterogeneous Local Fiscal Multipliers, Services and Goods Purchases

*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 University 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. The paper was previously circulated as “*Heterogeneous Spending, Heterogeneous Multipliers.*”

[†]Universidad del Rosario.

[‡]International Monetary Fund.

[§]World Bank.

1 Introduction

Until recently, the literature on fiscal multipliers treated government spending as a homogeneous good—an abstract aggregate collapsing tanks, hospital equipment, and cloud-computing contracts into the same unit of analysis.¹ This simplification, inherited from national accounting conventions, prevents us from assessing how the composition of government purchases shapes the transmission of fiscal shocks. While recent research has explored heterogeneity in private-sector responses, it has largely overlooked differences within government spending itself. This omission is increasingly important in an economy where both final consumers and producers have increased their demands for services at the expense of goods.²

Our paper fills this gap by examining whether the composition of government purchases influences the size and nature of fiscal multipliers—an issue long neglected due to the scarcity of granular spending data (Chodorow-Reich, 2019). Specifically, we answer two questions. First, does the size of the local fiscal multipliers depend on the type of government purchases? Second, what mechanisms explain this heterogeneity? To answer these questions, we compile a dataset covering 40 years of U.S. military procurement at the contract level. The dataset identifies the exact product or service purchased by the government, enabling us to classify purchases into goods and services. We exploit variation in procurement patterns across metropolitan statistical areas (MSAs) and over time to quantify the causal effects of the two types of public procurement purchases on employment, labor income and output (Nakamura and Steinsson, 2014).

Budget allocation remains a policy-driven decision, whereas production characteristics primarily affect the transmission of demand shocks. By isolating the multiplier effects associated with the allocation of purchases between the two categories of spending, we can more cleanly identify variation in the multiplier that is attributable to policymakers’ spending composition rather than to sector-specific (or firm-specific) characteristics. A better understanding of fiscal policy implications associated with the two categories of spending can assist policymakers in optimizing public resource allocation.

We document a shift in the composition of government defense spending from the early 1990s. Before this period, over 65% of defense procurement purchases were directed towards goods. However, in recent decades, the US federal government has allocated similar shares of resources to both goods and services.

We estimate employment, labor income and output local fiscal multipliers by category of purchases. An increase of one-percentage-point in services purchases (normalized by personal income) results in an on-impact employment growth of 0.25%. This impact strengthens over time, reaching 0.87% at the three-year horizon. Contrarily, employment responses for good purchases are positive but not statistically significant across any time horizon following the spending. These multipliers range from 0.03% on-impact to 0.12% at the four-year horizon, constituting between a quarter and a tenth of the estimates for services spending. Notably, unlike good purchases, service purchases stimulate significant increases in household labor income,

¹Cox et al. (2024) and Bouakez et al. (2025) are two recent works that made the case of considering spending as heterogeneous.

²In 2020, the service sector comprised almost 80% of the US production. (Herrendorf et al., 2014) document that services dominate household consumption, and (Francois and Hoekman, 2010; Duarte and Restuccia, 2010) show that services became the main intermediate inputs used by firms in production.

with increases between 0.31% and 1.53% at different horizons. Importantly, within a constrained sample due to limitations in data availability, an increase of one-percentage-point in service purchases yields a output response that varies between 0.3 on-impact and reaches the peak of 1.8 at the three-year horizons. In contrast, the output responses associated with the purchase of goods are not statistically significant, exhibiting values that range from approximately no effect on-impact to around 0.1 at the two-year horizon.

To gauge a better sense of the size of the employment estimates, we analyze the newly-created jobs for \$100,000 spent in procurement contracts. The number of new jobs created substantially differs between the two types of purchases. \$100,000 in procurement contracts that acquire services create 0.58 new jobs at impact, rising to 1.71 after two years, while good purchases only yield 0.08 new jobs initially and 0.15 after two years. These findings align with [Chodorow-Reich \(2019\)](#), who reports job creation estimates between 0.76 and 3.93 for the two-year horizon.

We also investigate whether employment effects concentrate solely within industries directly benefiting from procurement purchases or if the effects spill to non-receiving sectors. We observe that both purchase categories generate positive direct impacts on the leading recipients of such expenditure.³ Moreover, our findings indicate that indirect effects, manifesting as positive employment responses in sectors different from the leading recipient, are primarily associated with service purchases. The surge in manufacturing sector employment after service purchases is substantial, whereas other sectors witness moderate effects. Good purchases, however, do not trigger significant employment responses outside the manufacturing sector. This observation aligns with the growing use of services as intermediate inputs ([Duarte and Restuccia, 2010](#)).

We turn next to the causes of different responses for services and goods purchases. We explore three avenues: differences in production characteristics, extensive margin and productivity gains, and locality-specific characteristics. We find that production characteristics, and the extensive margin play an important role in determining the heterogeneous multiplier effects, while locality-specific characteristics do not seem to contribute to these differences.

We explore the role of production characteristics —contribution to final consumption, contribution to investment, centrality in the production network, labor intensity, markup, price rigidity, and tradability —identified by previous studies as important drivers of the fiscal multiplier ([Auerbach et al., 2020](#); [Bouakez et al., 2023](#); [Bouakez et al., 2025](#); [Cacciatore and Traum, 2022](#) [Cox et al., 2024](#)). First, we document that a larger share of service purchases compared to good purchases is allocated to industries with a larger contribution to final consumption and investment, higher centrality —upstream sectors that supply intermediate inputs to other industries —and higher labor intensity. Conversely, a greater share of good purchases compared to service purchases are channeled towards sectors with higher markups, higher frequency of price adjustments, and more tradable. Motivated by these differences, we assess the influence of these industry-attributes on the different fiscal responses after the two categories of purchases. Our empirical findings highlight centrality as a key explanatory factor. Spending in downstream sectors —more common for goods

³Manufacturing firms secure approximately 80% of good-purchase contracts, while contracts in service industries receive 55% of service purchases.

spending —generates stronger ripple effects through the production network. At the same time, higher markups in goods-producing industries appear to dampen their multipliers, though mainly by increasing estimation uncertainty. Controlling for these sectoral characteristics narrows the goods-services multiplier differential only modestly, leaving a substantial portion of the gap, especially for employment and labor income, unexplained.

Government purchases may also generate demand-driven growth by facilitating firm dynamism and productivity gains. We document that service purchases boost the establishment’s entry and exit into sectors with high labor productivity growth. By contrast, goods purchases increase the entry and exit rates in sectors with low labor productivity growth. These results suggest that service purchases generate an aggregate improvement in labor productivity larger than goods purchases because firm turnover mostly occurs in high-growth sectors. These results provide suggestive evidence that the extensive margin and productivity gains might amplify differently the fiscal responses to the two categories of purchases.

By contrast to the previous channels, we do not find support for locality-specific characteristics to matter in explaining the differences in the responses between categories of purchases. The allocation of spending in the two categories is uneven across locations, with some receiving a relatively larger share of spending in one of the two categories than the other. The differential geographical dispersion of goods-producing and services-producing firms may present a challenge for isolating the differential effects of goods-versus-services spending from differential effects arising from economic differences in the geographies in which spending occurs. We address this concern by calculating the correlations between the share of spending and location-specific characteristics that matter in quantifying the size of the fiscal response, such as the share of small firms ([Juarros, 2020](#)) or the MSA-size. We find these correlations are close to zero, implying no relationship between the share of a type of purchase and location-specific characteristics.

Our empirical analysis provides suggestive evidence that the heterogeneous multipliers by spending category are driven by multiple channels acting simultaneously (centrality, markups, business dynamics, and productivity gains). The goods-versus-services distinction serves as an effective shorthand for these factors. The key reason for emphasizing the goods-versus-services distinction is its direct relevance for policymakers. Government budgets are allocated by expenditure category (goods or services), whereas other characteristics are not directly observable or actionable in real time for fiscal authorities. In other words, policymakers typically cannot decide the spending based on these characteristics as easily as they can adjust the composition of spending towards services or goods. A major contribution of our paper is to show that this readily observable choice has significant implications for the multiplier, a point that would be less obvious if one focused solely on abstract characteristics. For these reasons, we believe that the goods-versus-services distinction provides clearer and more policy-relevant insights that help the design of fiscal interventions by emphasizing a dimension of spending that is both empirically tractable and directly tied to government decision-making.

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 private 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. 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.⁴ We classify spending into good and service purchases rather than consumption and investment. The importance of this classification is twofold. First, the US economy has shifted from a good economy to a service 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 response to fiscal stimuli. Our findings provide complementary insights to those already documented in the literature.⁵

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 (Ilzetzi et al., 2013; Corbi et al., 2019), the exchange rate regime (Born et al., 2013), population demographics (Basso and Rachedi, 2021), labor market rigidity (Cole and Ohanian, 2004; Hall, 2009; Gorodnichenko et al., 2012), automatic stabilizers (Dolls et al., 2012; Galeano et al., 2021), public indebtedness (Ilzetzi 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., 2022) 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 leads to different multiplier effects.

Lastly, our paper is linked to the growing literature that studies, in the presence of heterogeneous agents, the aggregate effects of fiscal policy and its channels. This literature has investigated both the heterogeneity on the consumer side (Galí et al., 2007; Brinca et al., 2016; Hagedorn et al., 2019), the heterogeneity on the production side (Bouakez et al., 2023; Bouakez et al., 2025), and the heterogeneity on both sides (Flynn et al., 2021). Our results show that the aggregate fiscal multiplier is higher when they originate in the service sectors, as they generate stronger general-equilibrium (indirect) effects. We test the importance of several industries' characteristics in quantifying the employment response after fiscal spending. In this regard, our findings are complementary to Bouakez et al. (2025) that use a structural approach to establish the importance of sectors' characteristics in determining the aggregate effect of sector-specific fiscal interventions.

⁴For developing countries, Ilzetzi et al. (2013) show that distinguishing between government consumption (goods and services) and investment is an important determinant of the impact of government spending shocks on output.

⁵Another complementary branch of the literature is the literature on infrastructure multiplier. The infrastructure multiplier literature is not suitable for addressing the short economic impact of fiscal policies because infrastructure projects typically have a long lead time. Furthermore, infrastructure spending tends to focus on a single industry (construction). Unlike the infrastructure multiplier literature, the purchase multiplier literature explores the effect of fiscal policies directed at already established products. These differences allow us to study the short-term effects of fiscal policy and quantify the role of industry characteristics in determining differences in the responses to fiscal policy.

⁶Studies 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.

2 Data

2.1 Government Procurement Contracts

We assemble a dataset 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. 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.⁷

The data contain detailed information, including the contract identification number, the action dates, 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 government procurement spending in the following way. We define the year in which the government purchase occurs as the year of the signature date, which occurs when the contract is either awarded or modified.⁸ The modifications of existing contracts could consist of downward revisions of the contract obligation. These modifications are reported as negative entries. As in [Auerbach et al. \(2020\)](#), we consider contracts with obligations and de-obligations with magnitudes within 0.5% of each other to be null and void. For comparability over time, we convert the nominal transaction value into real values using the US Bureau of Labor Statistics’ Consumer Price Index.⁹

The Product Service Codes and the Federal Supply Codes play a crucial role in classifying government purchases as either service or good purchases. 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 purchase as “Purchase of Services.” If the first digit is a number, we classify that spending as “Purchase of Goods.” To our knowledge, previous studies have not used this data feature in their analysis.¹⁰

Finally, the identification strategy exploits the geographic variation in government purchases. As standard in the literature, the geographic allocation is based on the location where the tasks of the contracts are

⁷Appendix A.1 tests the similarity of our data to that used in previous studies. Our universe of military spending is very similar to one from previous studies in both the aggregate level and the geographic distribution.

⁸The government fiscal year has been defined from October 1st to September 30th 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.

⁹No official statistics accurately measure local differences in prices. The BEA has recently started to produce MSA-level GDP deflators. These deflators are constructed by applying national price indices to current dollar values of MSA-level GDP at the industry level. Hence, these statistics do not capture local differences in prices but in industry composition. Some recent works, such as [Stroebel and Vavra \(2019\)](#) or [Hazell et al. \(2022\)](#) have constructed better measures for local price indexes. Unfortunately, we cannot use these measures either because they cover different levels of geographic aggregation and time periods or because they exploit proprietary microdata, which we don’t have access to. The lack of reliable official local price indexes and the difficulty of constructing better indexes have led other researchers, for example, [Auerbach et al. \(2020\)](#) or [Bisgaard-Larsen et al. \(2025\)](#), to use nominal prices. In our case, as we exploit a much longer time series than previous studies, the use of nominal prices may induce serious measurement errors. Thus, we opted for deflating the nominal values by national price indexes. Furthermore, goods and services may face different price adjustments. As a robustness analysis, we estimate our main specification by using two different aggregate deflators for goods and services spending.

¹⁰Appendix A.2 reports the major product codes included in the two categories.

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 (NBER) to aggregate the county-level contracts into Metropolitan Statistical Areas (MSAs).

2.2 Economic Outcomes

We use three outcome variables to proxy changes in economic activity. The first variable is employment, measured in the number of jobs.¹² The employment figures are collected from the Bureau of Economic Analysis (BEA).¹³ The second outcome variable consists of salary and wages, and proxies for the labor income. These data are collected annually from the BEA.¹⁴ The third variable is the Gross Domestic Product (GDP). Different from the previous outcome variables, GDP at the MSA level is only available starting from 2001. Finally, we collect personal income from the BEA. 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.

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,¹⁵ or the industry in which the contractor operates. 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.

The second set of filters is applied to the data after aggregating purchases 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 purchases to personal income smaller than 1.5.¹⁶

¹¹Following [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.

¹²The employment figures cover the total of full-time and part-time workers in any industry, including government and government enterprises.

¹³The BEA estimates of employment and wages differ from the BLS data because BEA adjusts for employment and wages not covered, or not fully covered, by the state unemployment insurance (UI) and the unemployment compensation for Federal employees (UCFE) programs. First, BEA adds estimates of employment and wages to the BLS data to bridge small gaps in UI coverage. Second, BEA uses additional source data to estimate employment and wages. Third, BEA adjusts employment and wages for misreporting under the UI and UCFE programs.

¹⁴For consistency, salary and wages, as all other nominal variables, are converted into real terms to absorb price variations.

¹⁵Our analysis starts from 1979 because the product codes before 1979 are not consistent with those after that date.

¹⁶We initially started the construction of the sample considering Core-based statistical areas (CBSAs) rather than MSAs.

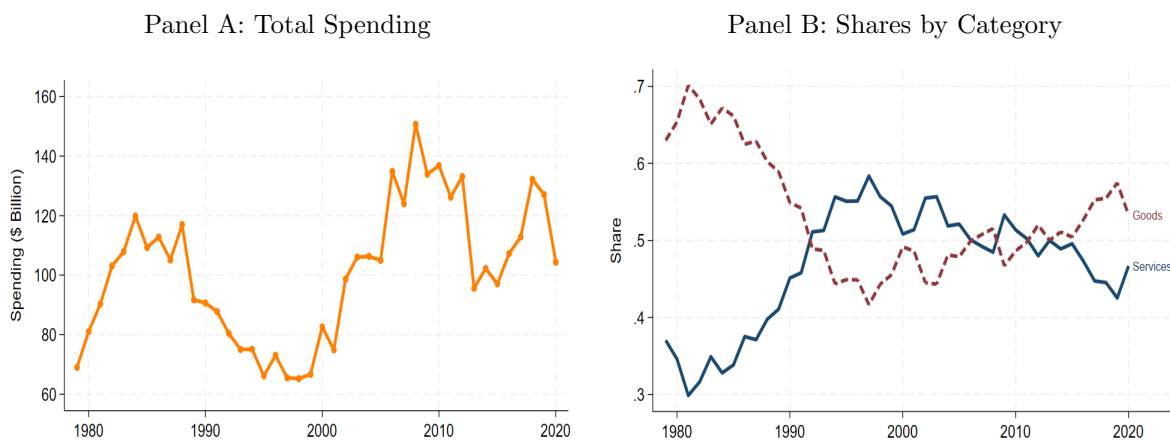
Finally, as in [Auerbach et al. \(2020\)](#) and [Demyanyk et al. \(2019\)](#), the analysis is carried out at the locality aggregated level rather than per capita.

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

3 Descriptive Statistics

Figure 1 reports spending over 1979 – 2019. Panel A shows the evolution of total 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 Afghan and Iraqi wars, and the last is in recent years, driven by the escalations in military buildups and competition with Russia and China.

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 as explained in Section 2.1. The classification of the purchases into goods and services is based on the Federal Product Classification. The figures are computed using the 344 MSAs in the final sample.

Panel B reports the shares of purchases by category. The division of the federal procurement purchases across the two categories highlights a reallocation of federal government purchases 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. Until the beginning of the 1980s, the spending on goods was about 60% of the procurement contracts. Since the 1990s, the spending has been split between goods and services with similar shares.¹⁷ Nevertheless, as showed in Table 1, the DoD awarded

The filters, particularly the minimum location size in terms of population, leave in the sample only a handful of micropolitan statistical areas. Therefore, we decided to restrict the analysis only to MSAs.

¹⁷A substantial portion of DoD spending is tied to traditional defense and military procurement. The DoD has also played a critical role in supporting technologies and services that, shortly after their creation, diffused broadly across the civilian economy. Well-known examples include the development of the internet, GPS, and medical innovations targeting public health threats such as endemic diarrheal diseases and viral infections. These cases highlight how DoD investments have historically anticipated and contributed to broader structural changes in the economy. To provide additional context, Table A.2 presents the list of product codes classified by spending category. These codes span a wide range of economic activities beyond core defense functions. To quantify changes over time, we calculate the growth in spending shares by category between two decades: 1981–1990 and 1991–2000. Among goods-related categories, 44 out of 78 experienced a decline in their share of total DoD spending. In contrast, 20 out of 24 service-related categories saw an increase in their share. The largest increases in services

most of its procurement contracts (91%) to purchase goods, and only a small share (9%) is directed to the acquisition of services, implying that service-purchase contracts are larger than good-purchase contracts.¹⁸

An interesting comparison is between the composition of government spending and private consumption expenditures. Figure B.1 shows that the share of private expenditures in services has also increased. Private consumers started to reallocate their spending from goods to service purchases at the beginning of the 1970s. Furthermore, the comparison highlights that government spending is more goods-driven than private consumption. This finding adds to the results of Cox et al. (2024) that emphasize the existence of a sectoral bias in government spending, by showing that the importance of goods and services purchases differs substantially between government spending and private consumption.

Table 1: Descriptive Statistics

	All	Goods	Services
Share of Spending (%)	-	54	46
Share of Contracts (%)	-	91	9
Average Value (Thousand USD)	173	102	860
90%-to-50%-percentile Ratio	39	34	21
Share of Spending for Top 1% (%)	88	92	66

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, as explained in Section 2.1. The classification of the spending into goods and services is based on the Federal Product Classification. The statistics are computed using the 344 MSAs in the final sample.

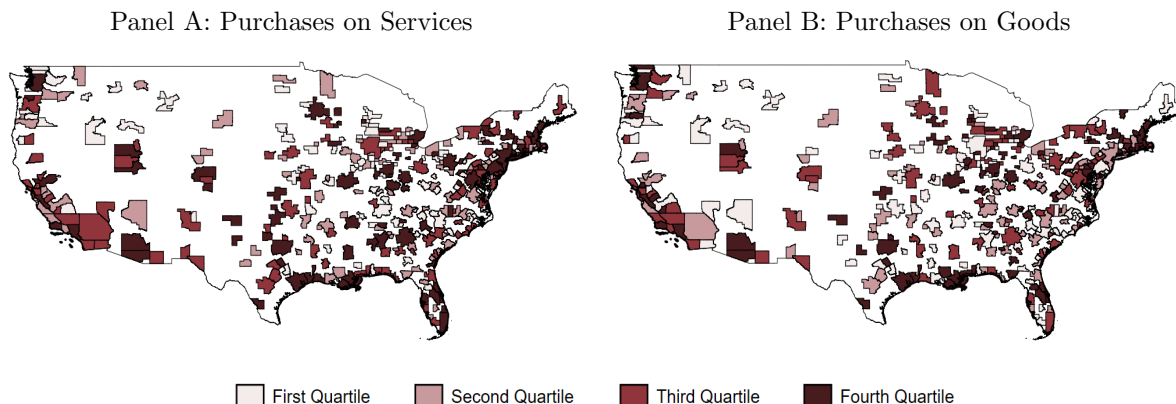
The last two rows of Table 1 explore the distributional characteristics of the purchases within categories. The distribution of contracts for the purchase of goods has a significantly fatter right-tail than the distribution for spending on services. In the case of spending for goods, contracts at the top decile of the distribution are about 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 is 21 times greater than the median contract. Among contracts to produce goods, the top 1% of contracts accounts for 92% of the goods spending. This result implies that although the majority of contracts buy 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 66% of the services spending, implying a more equal allocation of spending across contracts. Our findings are complementary to Cox et al. (2024), 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

were in categories Q - Medical Services, F - Natural Resources Management, and B - Special Studies/Analysis. These categories were negligible in the early 1980s but grew significantly over time.

¹⁸Services purchases includes R&D activities. R&D contracts are usually large. Removing R&D contracts from the service spending, the differences between the two categories are less marked. Nevertheless, the differences persist because the distribution of the two types of contracts differs, with service purchases more concentrated in the top tail of the distribution.

the actual performance of the tasks is granular. Our results highlight that, in addition to granularity in contract recipients, there exists a granularity in contract performance.¹⁹ Similarly to Cox et al. (2024), 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.

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 – 2019. The classification of the spending into goods and services is based on the Federal Product Classification. The figures are computed using the 344 MSAs in the final sample.

As the identification strategy exploits the cross-sectional variation in purchases, Figure 2 explores the geographic heterogeneity in spending allocation across MSAs. The figure shows the quartile to which an MSA belongs based on the average value of the military spending it received over 1979 – 2019. There are two main results. First, military spending is unequally distributed across MSAs. The top 40 MSAs in terms of awarded spending account for more than 35% 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.²⁰

These descriptive statistics provide evidence that although the DoD signs a large number of contracts, the largest share of government procurement spending is captured by a handful of contracts. Furthermore, these contracts are not equally distributed across MSAs. These findings imply that the distribution of spending at both contract and MSA levels is skewed to the right.

¹⁹We 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.

²⁰The Midwest has the highest concentration of production occupations with an average employment share in production jobs, almost 50% higher than the US average.

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 cross-sectional variation in military procurement purchase shares and time-series variation in national DOD spending to estimate impulse response functions of economic activity when local government purchases change.²¹ The benchmark specification is defined as follows:

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

where v is an outcome of interest in MSA l at horizon $t+k$ with $k = \{0, \dots, 4\}$. The endogenous variable $G_{l,t+k} - G_{l,t-1}$ measures the change in military spending normalized by the first lag of the personal income in location l , $Y_{l,t-1}$. The specification also includes MSA-fixed effects α_l^k to control for MSA-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.²² The coefficient β^k quantifies the change in the outcome of interest (employment, labor income, or output growth rate) in a window of k years when government purchase increases by a one percentage point of MSA's personal income.²³

We augment the benchmark specification to examine the heterogeneous effects of the different government spending components. In this respect, we estimate the specification (2):

$$\frac{v_{l,t+k} - v_{l,t-1}}{v_{l,t-1}} = \beta_g^k \frac{G_{l,t+k}^g - G_{l,t-1}^g}{Y_{l,t-1}} + \beta_s^k \frac{G_{l,t+k}^s - G_{l,t-1}^s}{Y_{l,t-1}} + \alpha_l^k + \delta_{t+k} + \varepsilon_{l,t+k} \quad (2)$$

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

4.2 Instrumental variables research design

Procurement spending is potentially endogenous due to political or economic factors. The observed changes in procurement spending may respond to unobserved local shocks, also affecting the outcome variables. For example, firms in locations with more 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 procurement

²¹In 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.

²²Standard errors are clustered by MSA.

²³Our empirical strategy provides estimates for the open economy relative local fiscal multiplier. Estimates of the open economy relative 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.

spending with a measure of comparative advantage that certain MSAs have in obtaining contracts. The instrument, $Z_{l,t+h}$, is defined as:

$$Z_{l,t+h} = \kappa_l \frac{G_{t+k} - G_{t-1}}{Y_{l,t-1}} \quad (3)$$

where κ_l is the average share of spending in MSA l . As the empirical specifications are in changes and we use MSA-fixed effects, the conditional variation 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:

$$Z_{l,t+h}^g = \kappa_l^g \frac{G_{t+k}^g - G_{t-1}^g}{Y_{l,t-1}}; \quad Z_{l,t+h}^s = \kappa_l^s \frac{G_{t+k}^s - G_{t-1}^s}{Y_{l,t-1}} \quad (4)$$

with Z^g being the instrument for changes in goods spending and Z^s the instrument for changes in services spending. κ_l^g and κ_l^s measure the shares of spending in goods and services captured by MSA l , respectively.

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 estimates. Our identification assumption is 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 errors induced by outsourcing 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.

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 MSAs without contracts. The importance of the bias in our estimates depends on the direction of the geographic spillovers. The spillovers are positive when the demand for final consumption or intermediate goods increases in the neighboring MSAs due to input-output linkages. 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, the spillovers could be negative, and our estimates would suffer from an upward bias.

The third concern relates to 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 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 Results

In this section, we first investigate the local outcome responses to local total government procurement spending. Then, we present the responses in the local outcome variables to government procurement spending divided by the two categories of interest —services and goods. In the last part of the section, we implement a set of checks to assess the robustness of our empirical findings.

5.1 Main Results

We start by estimating the response of employment, labor income, and output to local total government procurement spending.²⁴ The estimates in Panel A of Table [B.1](#) suggest a positive and significant effect on employment. Indeed, we estimate an effect of 0.091 on-impact and 0.268 at the two-year horizon. Similarly to Panel A, Panel B 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 0.13% on-impact and above 0.51% after a two-year horizon. Finally, Panel C reports the output response to shocks to the total spending. We document a positive but non-significant effect of government spending on GDP on-impact, but a positive and significant effect of 0.4% at the two-year horizon.²⁵

Now, we turn to the core results of the paper. We explore the effects of the two types of government

²⁴These estimates have been called in different ways in the literature. For example, [Nakamura and Steinsson \(2014\)](#) refers to it as the “open economy relative multiplier.” By contrast, [Auerbach et al. \(2020\)](#) refer to it as the “local fiscal multiplier.” In the remainder of the paper, we will use the notation introduced by [Auerbach et al. \(2020\)](#).

²⁵We compare our estimates to [Auerbach et al. \(2020\)](#), which is the closest paper to ours in terms of specification and sample construction. Our employment and income responses at the two-year horizon are 0.258 and 0.505, respectively. [Auerbach et al. \(2020\)](#) report 0.189 and 0.335 for employment and income, respectively. While the estimates for employment and income are relatively close, the discrepancy for the estimates in GDP is much larger. Indeed, we report 0.4 at $t + 2$, and [Auerbach et al. \(2020\)](#) 1.1 at the same horizon. The discrepancy is mostly due to the different geographic aggregation. [Auerbach et al. \(2020\)](#) focuses their analysis on CBSAs, while we restrict the sample to MSAs. Extending our analysis to CBSAs, we get an estimate for GDP at $t + 2$ of 0.9, which is close to the coefficient of [Auerbach et al. \(2020\)](#). Furthermore, if we consider states instead of MSAs, we get an estimate of 1.2 at $t + 2$, which is close to 1.4 documented by [Nakamura and Steinsson \(2014\)](#). Overall, we conclude that the differences in estimates can be explained by the different geographic aggregation.

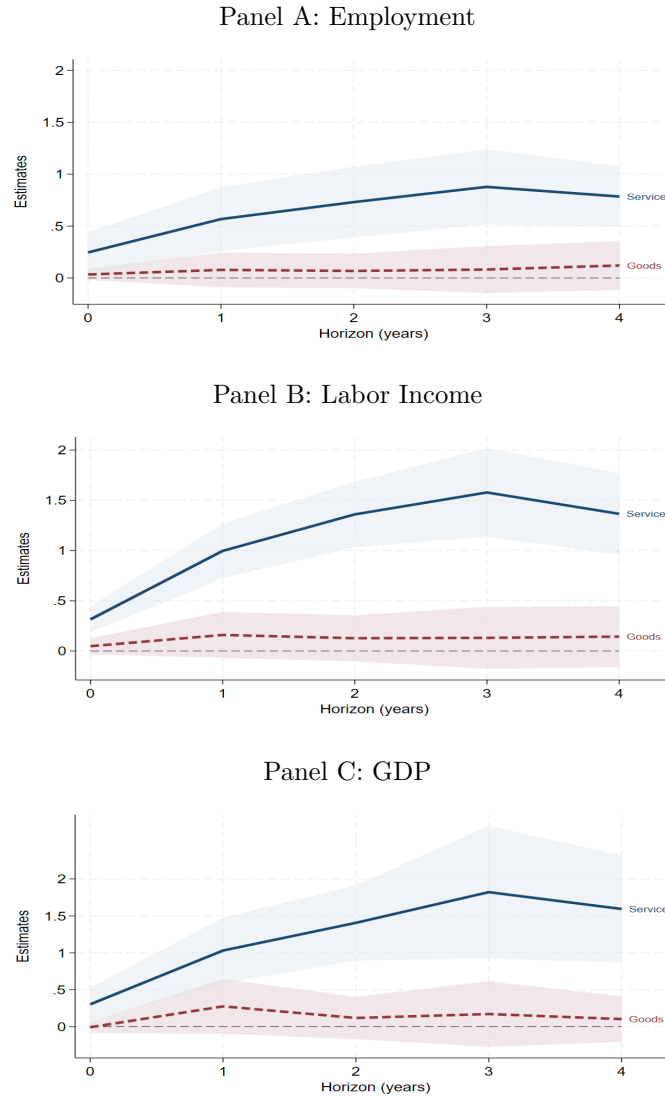
spending: spending on goods and services. Figure 3 plots the estimated effects reported in Table B.2. Panel A shows the effect on employment after goods and services spending shocks. We observe a remarkable difference between the two patterns. On the one hand, the estimates after a shock in services spending are positive and statistically significant. A one-percentage-point increase in normalized services spending by personal income generates an increase in employment of 0.24% on-impact and 0.72% 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 0.03% at impact and 0.07% at the two-year horizon, respectively. We test whether the gaps between two sets of coefficients are statistically different from each other, and we find that we reject the null hypothesis that the two estimates are equal.

One possible way to interpret these patterns is through a demand-driven mechanism. When demand for products rises, firms may respond by expanding their workforce or increasing hours worked. This could lead to higher labor income, which in turn may allow households to increase their consumption. In that case, the resulting increase in consumption could further reinforce demand for goods and services, potentially creating a feedback loop that supports additional labor demand. [Alonso \(2017\)](#) and [Bouakez et al. \(2025\)](#) show that the direct effect on the output of government spending shocks, consisting of output creation due to changes in government demand, is relatively small. Most of the aggregate impact comes from the indirect effect, consisting of changes in households' demand due to increases in their income as a consequence of government spending shocks. Our results suggest that this virtuous circle only activates after shocks to services spending but not goods spending. In Panel B of Figure 3 we show that there is a sizable increase in wages and salaries after a shock to services spending. Specifically, a one-percentage-point increase in normalized services spending generates a rise in labor income up to 1.3% within two periods after the shock. By contrast, labor income only slightly increases after goods spending shocks.²⁶ Finally, the last panel of Figure 3 shows the GDP impulse responses. Similarly to the previous two panels, we document a statistically positive effect of services spending on GDP, while the effect on GDP after goods spending shocks is close to zero and non-statistically significant at any standard levels.

If firm ownership goes largely beyond MSA boundaries, cross-MSA capital income may make the dividend income “less local.” As good producers rely on capital more than service producers, the spill of dividends to other locations may be more relevant for spending on goods and may partially explain the limited effects that we document after goods spending shocks. Although tracking ownership across MSA borders is hard, [Kolko and Neumark \(2010\)](#) show that 92.3% of establishments are locally-owned. Furthermore, [Mygrant \(2020\)](#) considers a subset of US firms and shows that the shares held within 20 miles of a firm establishment are 100% for family-owned businesses and 80% for employee-owned businesses. Only for publicly-traded firms that account for less than 1% of all US firms, the shares held within a 20 mile radius from a firm establishment are less than a quarter. These figures about the ownership structure of US businesses suggest

²⁶ 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 a combination of changes in labor income and dividends. We check the responses of dividends to the two types of shocks, and we find that firms do not significantly adjust dividends paid to households after spending shocks.

Figure 3: Impulse Responses - Purchase by Category



Notes: The panel includes 344 MSAs and consists of 12,024 observations for the period 1979 – 2019 for Panels A and B, and 4,676 observations for the period 2001 – 2019 for Panel C. Standard errors are clustered by MSA. The shaded areas represent the 90% confidence intervals.

that the spill of dividend income into neighboring locations is a relatively small issue. We further address this concern by quantifying the effect of both spending shocks on dividends in the neighboring locations. We find no effect on dividends in neighboring locations after either of the two types of spending, reinforcing the previous conclusion that the spill of dividend income into neighboring locations is relatively marginal.

As highlighted in [Auerbach et al. \(2020\)](#), DoD spending consists of two components: unanticipated spending associated with a new production —“demand shock” —and anticipated contract obligations that do not cause new production —“wealth transfers.” The negligible effect of DoD procurement spending on goods could be due to a weak first-stage instrument for goods spending. The weak instrument would not allow us to

isolate “demand shocks” from “wealth transfers” effectively. Alternatively, large goods producers may smooth production over fluctuations in national DoD spending more than services producers. Both cases would lead to underestimating the fiscal effects of goods spending. In Table B.2, we report the conditional first-stage F-statistics for the validity of the instruments. Since the core of the paper is testing whether the multipliers differ across types of spending, we need to show that we can identify both coefficients separately. This requires demonstrating that both instruments are valid separately for identifying their respective coefficients. We test whether a particular endogenous regressor alone is identified by implementing the test developed by Sanderson and Windmeijer (2016) that is based on the “partialled-out” procedure proposed by Angrist and Pischke (2009).²⁷ The 10% critical value for the conditional first-stage F-statistics for the validity of the instruments in a perfectly-identified model with two endogenous variables is 7.03. The reported F-statistics emphasize that we can reject the hypothesis that the instruments are weak, minimizing the concern of not effectively identifying “demand shocks.”

The specifications in equations (1) and (2) ignore the dynamics of spending and the outcome variables for the years between the periods of study. In other words, when the change in spending is measured as the difference in spending between $t + k$ and $t - 1$, the two specifications ignore the changes in the years $t \in \{t - 1, t + k\}$. 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 $t + k$. We refer to the cumulative response of local outcome variables to the cumulative local spending as the “cumulative local fiscal multiplier.” We construct the dependent variables as the cumulative change in an outcome variable between $t - 1$ and $t + k$:

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

and similarly, the cumulative local spending as the sum of the spending by category that a location receives between $t - 1$ and $t + k$:

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

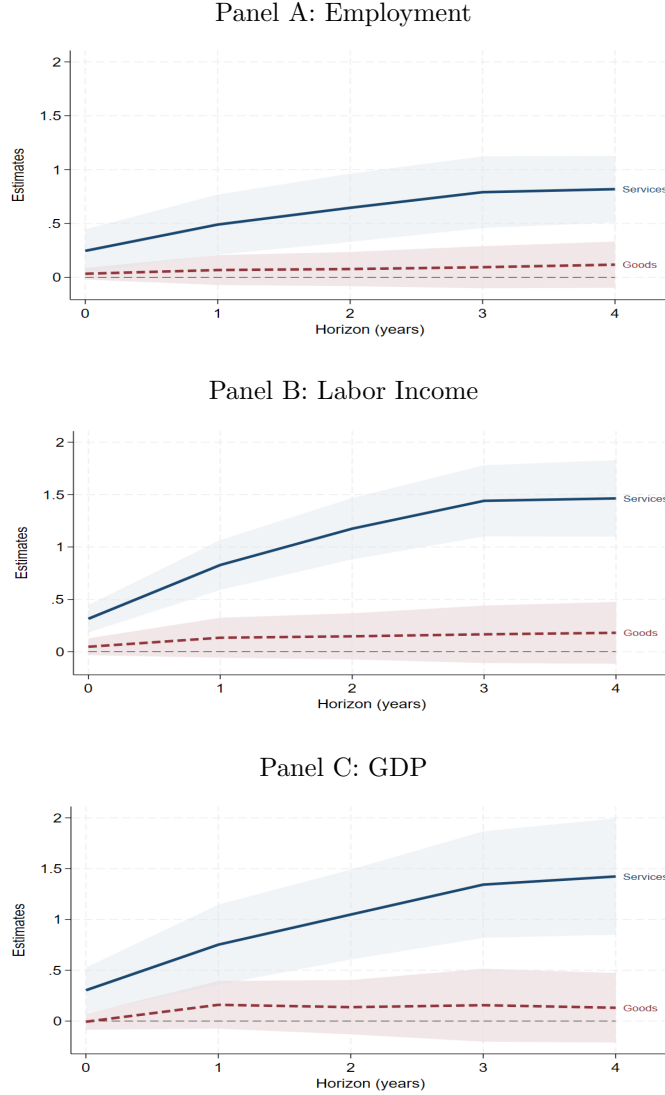
The estimates showed in Figure 4 come from Table B.3. The estimates highlight a stark difference between goods and services spending. Specifically, we document positive and significant responses of the cumulative local outcome variables to the cumulative local spending in services. By contrast, the cumulative effects in response to the cumulative local spending in goods are positive but non-significant. Furthermore, the increases in GDP and labor income are greater than one after the two-year horizon.

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,²⁸ and we aggregate them by MSAs and major NAICS

²⁷We also implement tests using only goods spending as an endogenous regressor. In all these cases, the F-statistics is above the Stock-Yogo critical value of 10, implying we can reject the presence of weak instruments.

²⁸Eckert et al. (2020) develop an algorithm to impute several suppressed employment by industry from the County Business Patterns. The suppression is due to confidentiality risks. Data are available until 2016.

Figure 4: Cumulative Fiscal Multiplier



Notes: The balanced panel includes 344 MSAs and consists of 12,024 observations for the period 1979 – 2019 for Panels A and B, and 4,676 observations for the period 2001 – 2019 for Panel C. Standard errors are clustered by MSA. The shaded areas represent the 90% confidence intervals.

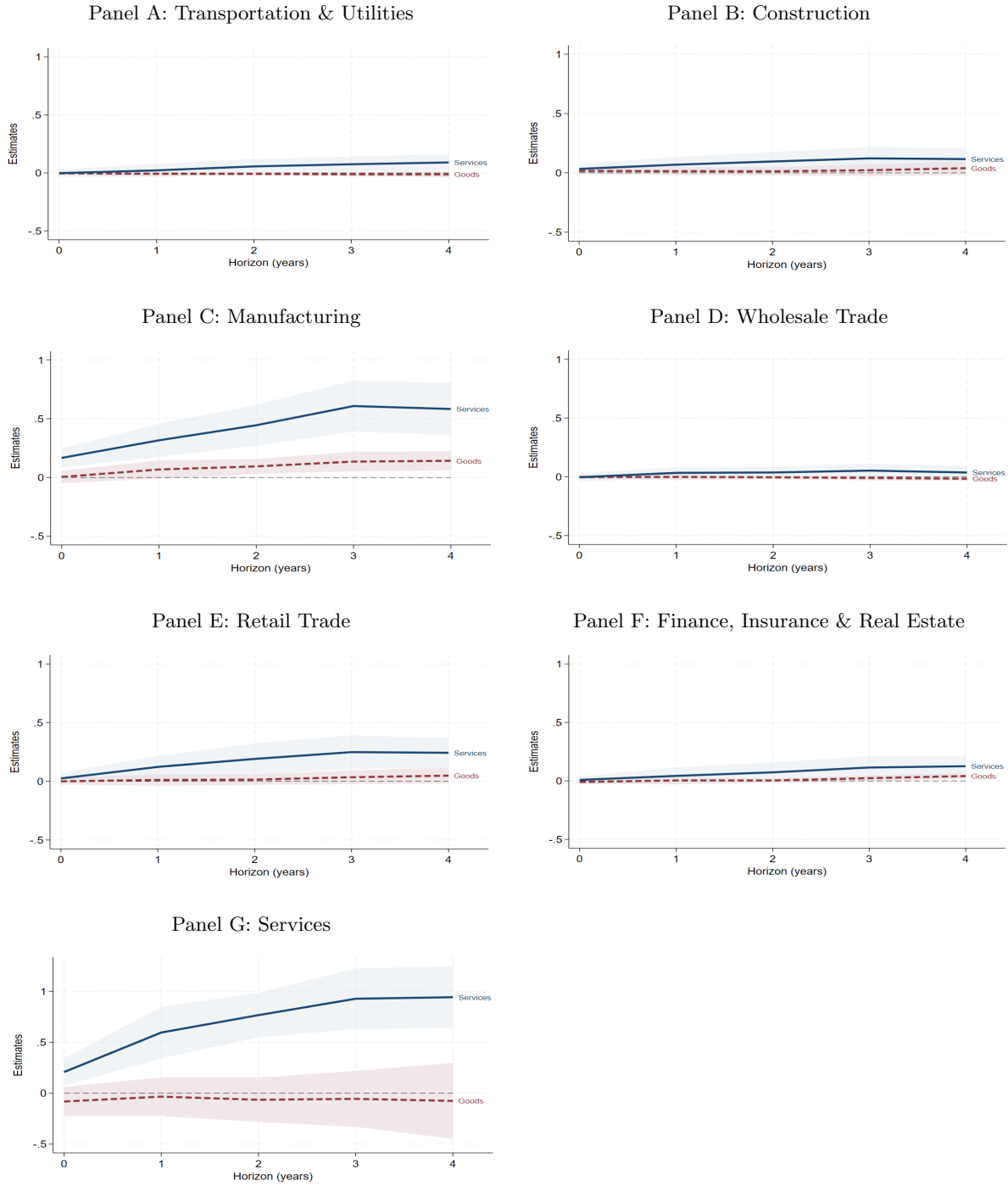
groupings.²⁹ As industries differ in their size, for better comparability across industries, we report the newly-created jobs rather than employment responses. The number of newly-created jobs per \$100,000 in procurement spending is calculated as

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

with $\hat{\beta}_i^k$ representing the estimate at different horizons k of the specification (2) estimated for each industry.

²⁹We 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.

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 344 MSAs for the period 1979 – 2019. Standard errors are clustered by MSA. The shaded areas represent the 90% confidence intervals.

$E_{l,t-1}$ is the level of employment in MSA l the year before the spending shock, $Y_{l,t-1}$ is the one year lag of the MSA's personal income, and T and N are the number of years and MSAs in our sample, respectively.

The estimates for the newly-created jobs are showed in Table B.4. Before describing the industry results, we briefly discuss the estimates for the entire economy.³⁰ The number of new jobs created substantially differs between the two types of spending. \$100,000 in procurement contracts that acquire services create 0.58 new jobs at impact, and 1.71 after two years. By contrast, \$100,000 used to purchase goods generate 0.07 new jobs at impact and 0.15 after two years. Figure 5 summarizes the calculations for the two types of spending by industry. First, our estimates suggest both types of spending generate positive and significant direct effects on the industry that receives most of that spending.³¹ Specifically, goods spending shocks generate 0.09 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 of positive employment responses that spill to other sectors occur mainly after services spending. Indeed, the employment response after services spending is large in the manufacturing sector and moderate in the other industries. By contrast, goods spending shocks do not generate significant employment responses in sectors other than manufacturing. The presence of indirect effects, which materialize as positive employment responses spilling to other sectors only after services spending shocks, is consistent with the fact that the services share of intermediate inputs has risen over time (Duarte and Restuccia, 2010).

5.2 Robustness

In the benchmark results in Table B.2, we deflate both types of spending by the same aggregate price index. Nevertheless, it is plausible that goods and services face different price adjustments across locations. To fully address this concern, we would need MSA-specific price indexes for goods and services that are unavailable. Thus, we provide some evidence that different price adjustments for goods and services do not matter in determining our main result by deflating the goods and services procurement spending series by the aggregate goods and services deflators.³² Table B.5 highlights the same pattern as in the benchmark specification. Specifically, we document a positive and significant effect after service spending in all three outcome variables and a non-significant and close to zero effect after goods spending.

Another concern relates to the size distribution of government procurement 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. Furthermore, as we showed in Section 3, services contracts are, on average, larger than goods contracts. Therefore, the size of the individual contracts may cause the differential fiscal multipliers rather the category of spending.

³⁰Briganti et al. (2025) show that the estimates of cost-per-job per year based on the employment multipliers derived from DoD procurement contracts are larger than those associated with other types of government spending.

³¹About 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.

³²We collect price indexes for services, durable and non-durable goods from the U.S. Bureau of Labor Statistics. We construct the price index for goods by weighted averaging durable and non-durable goods price indexes with weights equal to the relative weights of each consumption basket.

We address these concerns by taking three different approaches that help ensure that our estimates are not disproportionately driven by large federal procurement contracts.

The first approach consists of excluding from the regression analysis all contracts in the top 1% of the value distribution within each MSA-year. We adopt this conservative approach for two main reasons. First, the top 1% filter ensures that we remove genuinely large contracts, regardless of whether they are for goods or services, avoiding the potential bias of targeting only specific categories. Second, defining “large” relative to the MSA-year level accounts for substantial variation in contract size across both geography and time. A contract of average national size might be unusually large for a small MSA in a given year, and vice versa. Similarly, average contract size can vary significantly over time. Since our empirical strategy relies on both cross-sectional and temporal variation, we use a local and time-specific threshold in applying this restriction. The estimates excluding contracts in the top 1% of the value distribution within each MSA-year reported in Table B.6 reinforce the previous conclusion.

We acknowledge that the exclusion of contracts from our analysis may potentially introduce a bias in our estimates. This exclusion reduces the spending at the MSA-year level while leaving the outcome variables unchanged, potentially leading to an overestimation of the effect of each dollar spent through procurement contracts. To address this concern, we implement a complementary strategy that aims to mitigate this bias. Specifically, we exclude MSAs that receive “large” procurement contracts over the period 1979–2019. We define a “large” contract as one exceeding one million USD in real terms, and we exclude MSAs where such contracts represent more than 1% of total number of procurement contracts during the sample period.³³ The estimates excluding the MSAs with a large share of large contracts are showed in Table B.7. We find that removing MSAs with a large share of large contracts does not drive the differences between the service and good spending multipliers. If anything, we document that the gap between the estimates for the two categories of spending is greater when we restrict the sample of MSAs.

Finally, we adopt a third strategy to tackle the concern with the large contracts. We explicitly control for the lagged average contract sizes at the MSA level by type of spending.³⁴ This specification allows us to assess the extent to which differences in responses between goods and services purchases are driven by the larger size of service contracts. The coefficients controlling for the average contract size by type of spending are described in Table B.8. We document positive effects on the outcome variables due to service spending and no effects due to good spending. These results rule out the hypothesis that the differences in fiscal multiplier estimates between goods and services purchases are driven by differences in contract size.

6 Mechanisms

Spending on goods and services differs along several dimensions. This section investigates several mechanisms that could explain the heterogeneity in the previous estimates by category of spending.

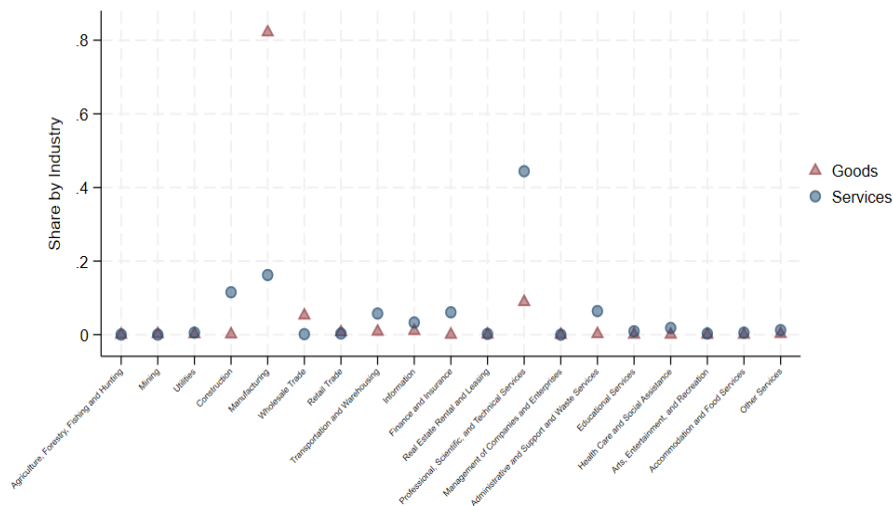
³³We have also conducted robustness checks using alternative thresholds for both the contract size used to define a “large” procurement and the percentage used to classify MSAs with negligible exposure.

³⁴We normalize the lagged average contract size by the lagged personal income at the MSA level

6.1 Industry-specific characteristics

Firms differ in their characteristics and produce distinct products using different production methods, which can influence the impact of spending in different ways, and lead to different amplifications. Our data lack information on firm characteristics and production inputs, but they do report the industry in which the work falls. As a workaround, we construct industry-level measures of these characteristics —contribution to final consumption, contribution to investment, centrality in the production network, labor intensity, markup, price rigidity, and trade costs. We then assign these measures to contracts based on the industry in which the work falls.³⁵

Figure 6: Share of Spending by Sectors



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.

Figure 6 shows the distribution of the two categories of spending by 2-digit NAICs codes. Goods spending is predominantly awarded to contractors that operate in the “*Manufacturing*” sector (80%), while services spending spans various sectors, with the “*Professional, Scientific, and Technical Services*” receiving about 40% of the DoD spending on services.³⁶ The marked differences in the sectoral allocation of the two types of spending make it important to examine whether the industries receiving these procurement contracts also differ in their characteristics.

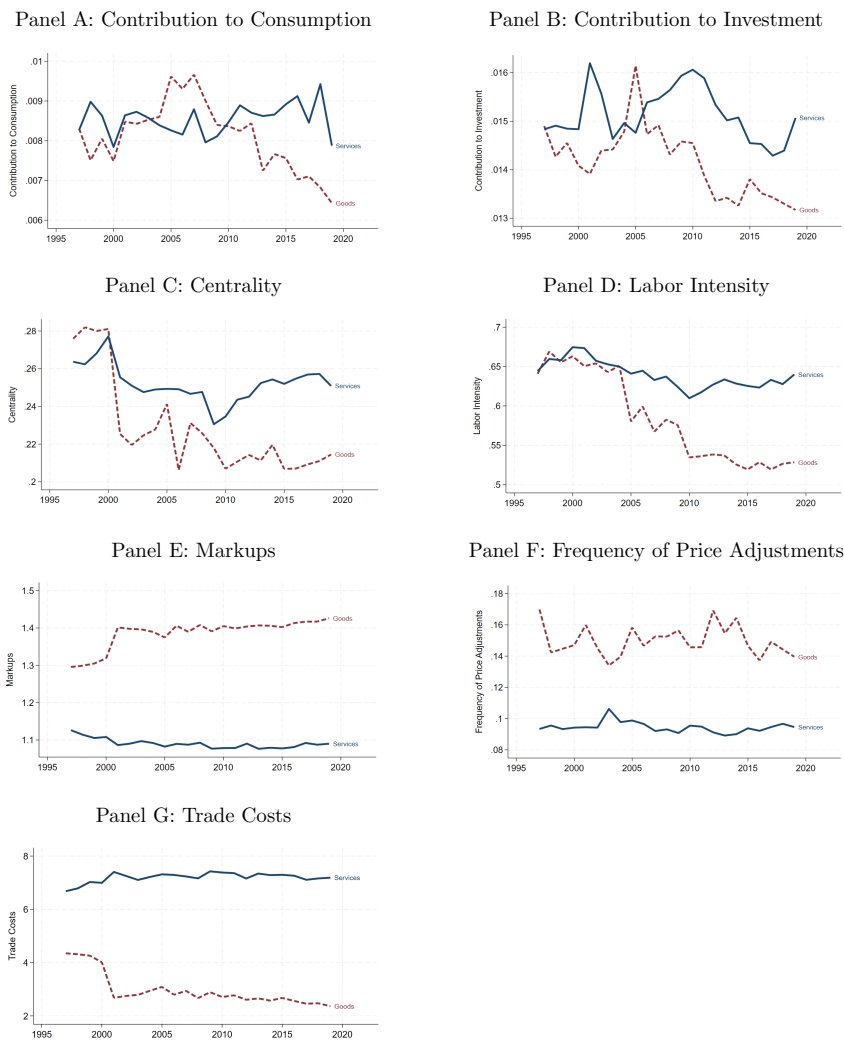
Each panel of Figure 7 presents the weighted average by the two types of spending of one of the seven sectoral characteristics —contribution to consumption, contribution to investment, centrality in the production network, labor intensity, markup, price rigidity, and trade costs. The weights are constructed as the

³⁵Assigning industry characteristics based on the contract rather than the contractor improves the measurement of the production process. Assigning the characteristics at the contract-level captures the production attributes of the contracted work rather than those of the firm. In the hypothetical case where characteristics were assigned based on the industry of the contractor, we would attribute the same production structure to both goods and services produced by the same firm. By contrast, our measures of the production process can vary within the same firm.

³⁶Auerbach et al. (2020) shows a similar ordering of industries receiving contracts without separating by products codes.

share of spending going to a specific industry relative to the total spending by category. These statistics highlight that firms receiving the two types of spending operate in industries that significantly differ along these characteristics.

Figure 7: Industry Confounders 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 spending allocated to the 334 MSAs in the final sample.

Panel A reports the contribution to final consumption demand, measured as the share of private consumption expenditure served by that industry. Sectors with higher consumption shares are typically better able to absorb increases in public demand without crowding out private consumption, supporting larger local fiscal multipliers. Empirically, prior to 2010, services spending is directed to industries contributing less to private consumption than industries receiving goods spending; since then, the consumption contribution of service-sector industries has remained broadly constant, while that of goods-producing industries has declined sharply.

Panel B considers the sectoral contribution to investment. A higher investment share can reduce or delay

the short-run multiplier, as investment goods involve longer production processes and capital constraints may limit firms’ ability to expand output quickly, especially under financing frictions (Bouakez et al., 2025). Before the mid-2000s, goods and services spending is directed to industries with similar investment contributions (around 1.5%). Since then, industries receiving services spending have exhibited higher investment shares than those receiving goods spending.

Panel C measures sectoral importance in the production network using the Katz–Bonacich centrality index.³⁷ Spending directed toward less central (downstream) sectors typically generates larger multipliers, as demand propagates upstream through input linkages (Baqae and Farhi, 2019; Flynn et al., 2021). Services contracts tend to be allocated to more upstream (more central) industries.

Panel D reports labor intensity, measured as employee compensation’s share in value added.³⁸ Higher labor intensity can raise local multipliers by strengthening employment and income channels (Nakamura and Steinsson, 2014; Alonso, 2017). An increase in public purchases boosts product demand, prompting firms to increase hiring. Consequently, labor and personal incomes grow, fueling higher consumer spending. Until the mid-2000s, labor intensity was similar across sectors receiving goods and services contracts (around 0.65); thereafter, it remained stable for the industries which receive services spending while declining sharply for the sectors which receive goods spending.³⁹

Panel E examines product-market competition using sectoral markups constructed following Hall (2018), which utilizes sector-level data on output and production inputs from the BEA. Markups affect fiscal multipliers through competing channels. High-markup sectors may amplify spending responses via pricing power (Rotemberg and Woodford, 1999; Baqae and Farhi, 2020), while lower distortions in competitive sectors can also yield larger output effects (Bouakez et al., 2025). Empirically, industries receiving goods spending exhibit higher markups (around 1.4) than those receiving services spending (around 1.1), with the gap increasing over time.

Panel F reports price rigidity using sectoral frequencies of price adjustment from Pasten et al. (2020).⁴⁰ Greater price stickiness can raise multipliers by limiting the inflation response and reducing monetary crowding out (Cox et al., 2024). Goods spending is awarded to sectors with more frequent price changes (about 15% monthly) than services spending (about 10%). Under Calvo pricing, this implies average price durations

³⁷We follow Acemoglu et al. (2012) in calculating the Katz–Bonacich measure of centrality as $\frac{\alpha_H}{S} (\mathbf{I} - \alpha_H \mathbf{W}')^{-1} \mathbf{1}$, with α_H being the average gross-output intensity of intermediate inputs, S the number of sectors, and $\mathbf{W} = \{\omega_{H,sx}\}_{s,x=1}^S$ the BEA Input–Output matrix.

³⁸Imperfect competition could bias this measure if monopsony power leads firms to underpay workers, mechanically lowering measured labor intensity. To assess this concern, we regress labor intensity on industry employment concentration ratios (with year and industry fixed effects) and find only a weak negative relationship, suggesting that monopsony power is not a key driver of measured labor intensity.

³⁹Hall (2009) and Bouakez et al. (2025) highlight different channels through which labor intensity affects fiscal multipliers. In a one-sector model with decreasing returns, Hall (2009) shows that spending directed to more labor-intensive sectors yields higher multipliers by mitigating labor substitution effects. In a multi-sector setting, Bouakez et al. (2025) argue that labor-intensive sectors generate larger multipliers because they rely more on labor, the non-marked-up input. By contrast, standard neoclassical models with short-run inelastic capital supply predict that shocks to more labor-intensive sectors can raise prices and crowd out activity, lowering multipliers.

⁴⁰The frequency price adjustments are measured at the input-output industry codes. We use the BEA concordance tables between NAICS codes and input-output industry codes to recover frequency price adjustments at the NAICS codes. A few NAICS codes are not matched.

of 6.2 months for goods sectors and 9.5 months for services sectors,⁴¹ consistent with stronger local effects for services spending.

Panel G reports trade costs from [Gervais and Jensen \(2019\)](#). Higher trade costs imply lower tradability and fewer geographic leakages, which traditionally support larger local multipliers ([Auerbach et al., 2020](#)). Alternative channels, however, imply that trade openness can also amplify multipliers ([Cacciatore and Traum, 2022](#); [Moretti, 2010](#)).⁴² Empirically, trade costs for services spending are more than twice those for goods spending and remain stable over time, implying that the products generated through services spending are less tradable.

Overall, [Figure 7](#) shows that services procurement is systematically concentrated in sectors characterized by higher contribution to final consumption and investment, more centrality in the production network, higher labor intensity, lower markups, and lower frequency of price adjustment than the sectors receiving goods spending.

As the goods and services spending are directed to industries with different characteristics, these differences may affect the transmission of a fiscal shock, generating different multiplier effects for the two types of spending. To investigate the relevance of these industry characteristics, we follow a strategy similar to [Bouakez et al. \(2025\)](#). We test the importance of the industry characteristics in explaining the differences in the fiscal multiplier between goods and services spending by running the following specification:

$$\begin{aligned} \frac{v_{l,t+k} - v_{l,t-1}}{v_{l,t-1}} = & \beta_g^k \frac{G_{l,t+k}^g - G_{l,t-1}^g}{Y_{l,t-1}} + \beta_s^k \frac{G_{l,t+k}^s - G_{l,t-1}^s}{Y_{l,t-1}} \\ & + \gamma^k \frac{G_{l,t+k} - G_{l,t-1}}{Y_{l,t-1}} \cdot \frac{(\mathcal{C}_{l,t-1} - \bar{\mathcal{C}})}{\hat{\mathcal{C}}} + \psi \mathcal{C}_{l,t-1} + \alpha_l^k + \delta_{t+k} + \varepsilon_{l,t+k} \end{aligned} \quad (5)$$

where v is an outcome of interest in location l at horizon $t + k$. The interaction term is constructed by multiplying the normalized changes in total spending by the demeaned weighted average of one of the sectoral characteristics. $\mathcal{C}_{l,t}$ is computed at the MSA level by summing, across all sectors, the product of each sector's characteristic and the share of spending directed to that sector in MSA l at time t , relative to the total spending received by that MSA in a given year. $\bar{\mathcal{C}}$ and $\hat{\mathcal{C}}$ are the average and the standard deviation of $\mathcal{C}_{l,t}$ over all MSAs and all years, respectively.⁴³

On the one hand, the coefficients β_g and β_s are interpreted as the local fiscal multiplier for an MSA whose weighted-average sectoral characteristic of total spending is equal to the average in the sample. In other

⁴¹The average duration is calculated as $-\frac{1}{\log(1-FPA)}$ with FPA being the mean monthly frequency of price adjustments.

⁴²First, [Cacciatore and Traum \(2022\)](#) show analytically and quantitatively that government spending multipliers can be larger in more open economies. In their view, lower trade costs imply stronger trade linkages that amplify spending through private-demand crowd-in or changes in the terms-of-trade. Second, if procurement is concentrated into more tradable industries, those industries may behave like exporters. In other words, they bring money into the local economy, which boosts the local wages and also creates jobs and spending in less tradable sectors ([Moretti, 2010](#)).

⁴³Our specification differs from that of [Bouakez et al. \(2025\)](#) in two aspects. First, we decompose the total expenditures into the first two terms to capture the heterogeneous multiplier effect between goods and services purchases. Second, [Bouakez et al. \(2025\)](#) interact the change in total spending with the sectoral characteristic at the horizon $t + k$. Implicitly, they assume that the evolution of sectoral characteristics is exogenous to the spending shock. In our specification, we interact the spending shock with the sectoral characteristic at $t - 1$, making it pre-determined at the time of the shock.

Table 2: Impulse Responses - Sectoral Characteristics at $t + 3$

	Baseline	Contribution to Consumption	Contribution to Investment	Centrality	Labor Intensity	Markups	Frequency of Price Adjustments	Trade Costs	All
Panel A: Employment									
Services Spending	0.884*** (0.232)	0.887*** (0.176)	0.846*** (0.157)	0.897*** (0.176)	0.900*** (0.182)	0.842*** (0.200)	0.879*** (0.174)	0.935*** (0.174)	0.821*** (0.166)
Goods Spending	-0.100 (0.067)	-0.091 (0.079)	-0.063 (0.067)	-0.161 (0.122)	-0.086 (0.060)	-0.008 (0.411)	-0.065 (0.070)	-0.132 (0.207)	0.089 (0.333)
Spending * Contribution to Consumption		0.012 (0.069)							-0.011 (0.077)
Spending * Contribution to Investment			0.095 (0.163)						0.086 (0.124)
Spending * Centrality				-0.149 (0.183)					-0.163 (0.162)
Spending * Labor Intensity					-0.035 (0.063)				-0.065 (0.060)
Spending * Markup						-0.042 (0.214)			-0.155 (0.158)
Spending * Frequency of Price Adjustment							-0.042 (0.080)		-0.078 (0.069)
Spending * Trade Costs								-0.053 (0.215)	-0.109 (0.179)
Observations	6,346	5,997	6,002	6,003	6,007	6,005	5,992	6,000	5,981
$\beta_s - \beta_g$	0.983	0.978	0.908	1.058	0.986	0.850	0.944	1.066	0.731
Test $\beta_s = \beta_g$	[.000]	[.000]	[.000]	[.000]	[.000]	[.136]	[.000]	[.002]	[.113]
Panel B: Labor Income									
Services Spending	2.045*** (0.545)	2.025*** (0.422)	1.984*** (0.405)	2.080*** (0.424)	2.058*** (0.434)	1.959*** (0.426)	2.020*** (0.433)	2.295*** (0.384)	1.914*** (0.332)
Goods Spending	-0.106 (0.123)	-0.076 (0.175)	-0.056 (0.129)	-0.375 (0.235)	-0.077 (0.116)	0.044 (0.869)	-0.045 (0.134)	-0.310 (0.452)	0.255 (0.622)
Spending * Contribution to Consumption		0.002 (0.167)							0.008 (0.162)
Spending * Contribution to Investment			0.090 (0.382)						0.163 (0.225)
Spending * Centrality				-0.615 (0.386)					-0.636** (0.322)
Spending * Labor Intensity					-0.069 (0.116)				-0.095 (0.111)
Spending * Markup						-0.066 (0.452)			-0.447 (0.278)
Spending * Frequency of Price Adjustment							-0.066 (0.135)		-0.118 (0.123)
Spending * Trade Costs								-0.277 (0.472)	-0.322 (0.375)
Observations	6,346	5,997	6,002	6,003	6,007	6,005	5,992	6,000	5,981
$\beta_s - \beta_g$	2.151	2.101	2.039	2.455	2.135	1.915	2.065	2.605	1.659
Test $\beta_s = \beta_g$	[.000]	[.000]	[.000]	[.000]	[.000]	[.101]	[.000]	[.000]	[.053]
Panel C: GDP									
Services Spending	1.820*** (0.544)	1.574*** (0.526)	1.974*** (0.560)	1.947*** (0.456)	1.796*** (0.533)	1.522 (0.945)	1.775*** (0.549)	2.287*** (0.434)	1.330 (0.990)
Goods Spending	0.171 (0.269)	0.373 (0.360)	0.074 (0.203)	-0.286 (0.287)	0.168 (0.266)	0.552 (1.454)	0.234 (0.287)	-0.179 (0.552)	0.870 (1.338)
Spending * Contribution to Consumption		-0.410 (0.301)							-0.205 (0.262)
Spending * Contribution to Investment			-0.417 (0.461)						-0.005 (0.447)
Spending * Centrality				-0.843** (0.422)					-0.885* (0.481)
Spending * Labor Intensity					0.033 (0.143)				-0.120 (0.156)
Spending * Markup						-0.200 (0.696)			-0.689 (0.812)
Spending * Frequency of Price Adjustment							-0.122 (0.179)		-0.165 (0.161)
Spending * Trade Costs								-0.374 (0.470)	-0.352 (0.488)
Observations	4,676	4,669	4,672	4,672	4,674	4,673	4,665	4,668	4,657
$\beta_s - \beta_g$	1.648	1.201	1.900	2.233	1.628	0.969	1.541	2.466	0.460
Test $\beta_s = \beta_g$	[.005]	[.078]	[.001]	[.000]	[.006]	[.676]	[.013]	[.004]	[.840]

Notes: The panel includes 344 MSAs. Given that certain industry characteristics are only available starting in 1997, we restrict the baseline specification results to this period for consistency. The discrepancy between the baseline estimates in Table B.2 and those reported here arises from the different time periods used in the analysis. 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.

words, β_g and β_s measure the response in one of the outcome variables in an MSA with sectoral characteristic $\mathcal{C}_{i,t-1} = \bar{C}$ following a dollar increase in goods and services spending in that MSA, respectively. On the other hand, the coefficient of the interaction term informs us on how the local fiscal multiplier varies with each sectoral characteristic. Our specification implies that there are no heterogeneous effects of the sectoral characteristics on the local fiscal multiplier by category. Equivalently, sectoral characteristics affect the coefficients of the two categories of spending only through the degree of exposure to these characteristics.

Table 2 explores the contribution of the sectoral characteristics evaluated at the horizon $t + 3$, which represents the peak of the multiplier effects, with Panel A for employment, Panel B for labor income, and Panel C for GDP. The first column reproduces the estimates from equation (2) restricting the sample period from 1997 due to the fact that certain industry characteristics are only available from that year. The next eight columns add sectoral characteristics one at a time. The final column includes all jointly.

The first two panels suggest that controlling for sector traits individually does not alter the goods-vs-services multiplier gap. In every specification, the service-spending coefficient remains larger and statistically significant. Even with all characteristics included, the local multiplier on services remains markedly higher. This suggests that the greater impact of service purchases on jobs and income is robust and not attributable to any single observed sectoral characteristic —such as markups, centrality, price rigidity, contribution to final demand, labor intensity, or tradability.

The last panel reports the estimates for output and shows a slightly different pattern. Including the centrality interaction (column 3) yields a negative and statistically significant coefficient, indicating that spending in more central (upstream) sectors is associated with smaller local output effects because they generate weaker propagation of the government spending through the network. This aligns with network-based multiplier theory (Baqae and Farhi, 2019; Flynn et al., 2021). In Panels A and B, the centrality coefficient is negative but non-significant when tested alone, yet becomes significant in the specification with all sectoral characteristics. Thus, across all outcomes, centrality emerges as a key structural factor. Government spending in less central industries consistently yields higher local multipliers, helping explain part of the services-goods disparity given that our data show service contracts are disproportionately awarded to firms in upstream, high-centrality sectors, while goods contracts more often go to downstream sectors. The negative centrality effects suggest that this upstream bias in service spending should, if anything, reduce its multiplier advantage. Yet the service advantage persists, implying other mechanisms are at play.

Returning to Panel C of Table 2, the goods-vs-services gap is noticeably affected when we control for markups, while the other individual controls leave the gap qualitatively intact. When we include the markup interaction (column 6), or all characteristics (column 8), the point estimate for services remains higher but loses statistical significance. In these cases, we cannot reject the null hypothesis that goods and services have equal output responses. However, this loss of significance is largely due to inflated standard errors —caused by reduced degrees of freedom and multicollinearity —rather than a substantial change in point estimates. Thus, while high-markup sectors (more common in goods-producing industries) may dampen

output multipliers, the markup control primarily increases uncertainty rather than altering the underlying estimates. This is consistent with the idea that firms with greater market power adjust quantities less in response to demand shocks, attenuating the fiscal multiplier. Still, the markup channel does not fully explain the stronger service multiplier.

We compare the estimates from Panel C with those of [Bouakez et al. \(2025\)](#), who study aggregate multipliers by sector characteristics. Qualitatively, our findings align with those of [Bouakez et al. \(2025\)](#). Indeed, output responses are larger when spending targets sectors that are more downstream (low centrality), have lower markups, stickier prices, and possibly higher labor intensity and greater final demand shares.⁴⁴ Differences in coefficient magnitudes and significance likely reflect the use of more granular geographic units and aggregate rather than per capita outcomes.

In summary, we have shown that sectors receiving service contracts differ systematically from those receiving goods contracts. Service procurement is more concentrated in upstream, high-centrality sectors with lower markups, greater labor intensity, and larger contributions to final consumption and investment. Goods procurement is more common in downstream sectors with higher markups, more flexible prices, and lower trade costs. Our regressions highlight centrality as a key explanatory factor. Spending in downstream sectors—more common for goods spending—generates stronger ripple effects through the production network. At the same time, higher markups in goods-producing industries appear to dampen their multipliers, though mainly by increasing estimation uncertainty. Controlling for these sectoral characteristics narrows the goods-services multiplier differential only modestly, leaving a substantial portion of the gap, especially for employment and labor income, unexplained. This suggests that additional mechanisms are at work. In the next subsections, we explore further channels that may help explain the remaining difference in local multiplier effects between goods and services procurement.

6.2 Extensive margin and productivity gains

Government spending may affect firms’ entry and exit decisions, and therefore, the initial demand shock can be amplified through the extensive margin. Increases in spending may reduce the uncertainty about future profits and ease credit constraints, leading to higher turnover rates of firms and productivity gains.⁴⁵

Different types of spending may generate different firm dynamics patterns, leading to heterogeneous multiplier effects. Several reasons could explain these different patterns. First, government spending may lower entry costs and ease credit constraints. 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, 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. Second, after each type of spending, new firms may enter

⁴⁴[Bouakez et al. \(2025\)](#) use the Calvo probability of not adjusting prices, while we use the frequency of price changes; thus, a negative sign on our measure corresponds to a positive effect of price stickiness in their framework.

⁴⁵Recent empirical studies have provided insights into the relationship between fiscal policy, business dynamism, and growth. [Ferraz et al. \(2015\)](#) show that firms that win government procurement contracts grow more than their competitors. [Lewis and Winkler \(2017\)](#) show that net firm entry rises after an expansion in US government spending. [Bisgaard-Larsen et al. \(2025\)](#) find positive effects on TFP and labor productivity after government spending shocks.

sectors that have a different contribution to aggregate productivity.

Table 3: Entry and Exit Rates

	(1)	(2)	(3)	(4)	(5)
	t	t+1	t+2	t+3	t+4
<i>Entry Rate - Sectors High Growth</i>					
Services Spending	2.297** (1.082)	2.586 (1.576)	2.612** (1.197)	2.254*** (0.841)	2.212*** (0.540)
Goods Spending	-0.129 (0.392)	-0.334 (0.297)	-0.183 (0.226)	-0.054 (0.168)	-0.035 (0.175)
<i>Entry Rate - Sectors Low Growth</i>					
Services Spending	-0.611 (0.436)	-0.484 (0.363)	-0.109 (0.401)	0.314 (0.446)	0.493 (0.428)
Goods Spending	0.023 (0.110)	0.339** (0.163)	0.263** (0.115)	0.273* (0.145)	0.406** (0.168)
<i>Exit Rate - Sector High Growth</i>					
Services Spending	-0.681 (1.263)	1.753* (1.029)	2.217** (0.975)	2.457** (1.083)	2.241*** (0.704)
Goods Spending	0.565 (0.400)	0.331 (0.241)	0.296** (0.136)	0.089 (0.147)	0.192 (0.164)
<i>Exit Rate - Sector Low Growth</i>					
Services Spending	-0.261 (0.399)	0.079 (0.403)	0.294 (0.330)	0.681* (0.388)	0.804** (0.344)
Goods Spending	0.256 (0.265)	0.227 (0.158)	0.308** (0.124)	0.392** (0.180)	0.335** (0.169)

Notes: The panel consists of 12,024 observations and includes 344 MSAs. 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.

To dig into the importance of the extensive margin and productivity gains in explaining the different multiplier effects, we proceed as follows. We collect data series on establishment entry and exit from the County Business Patterns at the MSA-sector level. As we cannot construct labor productivity series at the MSA level from 1979 because data are available only from the early 2000s, we follow a different strategy to quantify the effect of establishment entry and exit on productivity. We compute labor productivity at the sector level, dividing sector GDP by its employment using data from the BEA.⁴⁶ We then classify sectors as sectors with high (low) labor productivity growth if their labor productivity growth rates are above (below) the median growth rate across the sectors. Finally, we aggregate the establishment entry and exit to get MSA series of entry and exit in sectors with high or low labor productivity growth. This strategy allows us to explore whether entry and exit occur in sectors that have a greater contribution to productivity.⁴⁷

⁴⁶Data with a consistent definition of sectors are available from 1997 onward.

⁴⁷We recognize that recent studies have shown a large dispersion in firms' productivity even in narrow industries. Our data do not allow us to identify if the entry or exit is high or low productivity within their sector defined at the 2-digit NAICS. The

Table 3 reports the estimates for establishment entry and exit by sector’s labor productivity growth. Both types of spending incentivize business turnover. We document that service purchases boost the establishment’s entry and exit into sectors with high labor productivity growth. By contrast, goods purchases increase the entry and exit rates in sectors with low labor productivity growth, consistent with the fact that goods purchases are directed to contractors in higher markup sectors. These results suggest that service purchases generate an aggregate improvement in productivity larger than goods purchases because firm turnover mostly occurs in high-growth sectors. These results provide suggestive evidence that the extensive margin and productivity gains might amplify differently the fiscal responses to the two categories of purchases.

Our findings are consistent with Bisgaard-Larsen et al. (2025), which also document increases in net entry and productivity improvements as a consequence of government spending. With respect to them, we highlight the different responses by the type of spending, and we provide new insights about the mechanisms of how spending shocks impact business dynamism and productivity.

6.3 Location-specific characteristics

Figure 2 emphasizes that the allocation of spending in the two categories is uneven across locations, with some receiving a relatively larger share of spending in one of the two categories than the other. The differential geographical dispersion of goods production and services production may present a challenge for isolating the differential effects of goods-versus-services spending from differential effects arising from economic differences in the geographies in which spending occurs. Indeed, if locality-specific characteristics are correlated with the probability of receiving a type of spending and the determination of the size of the fiscal multiplier, then the different effects between services and goods spending that we have documented may be caused by some location-specific characteristics.

In the main specification in equation (2), we include location fixed-effects that capture the time-invariant characteristics of a specific location. Thus, our estimates isolate the effects of the different types of spending after absorbing location-specific confounders that may affect the outcome variable of interest.⁴⁸

To further dig into this potential explanation, we test the correlation between the allocation of spending types and the location-specific characteristics. Juarros (2020) documents that DoD spending multipliers are higher with larger shares of small firms.⁴⁹ If locations that receive more services spending also have a larger share of small firms, the larger multiplier effect coming from services spending could be attributed to the firm-size distribution rather than the type of spending. Similarly, if locations receiving more spending on services are more populated, the larger local market may generate larger multiplier effects. To rule out these hypotheses, we compute the correlation between the year-location share of spending on services and location characteristic. The resulting correlation is virtually zero. Overall, location-specific characteristics seem not

implicit assumption is that the entry and exit are close to the average productivity of the sector to which they belong. We are aware of this important caveat in the interpretation of our results.

⁴⁸Furthermore, in constructing the sample, we apply a series of commonly-used filters, described in Section 2, that remove outliers locations from our sample and make the locations more comparable along their characteristics.

⁴⁹We collect data on the share of small firms from the Business Dynamics Statistics files. Small firms are defined as firms with less than 20 employees.

to be relevant confounders of the multiplier differences between services and goods spending.

7 Conclusions

The Great Recession renewed interest in the effectiveness of fiscal policy. Most of the studies that estimate local fiscal multipliers find positive effects of fiscal spending on outcomes like employment, income, 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 US economy has seen two significant structural transformations related to the role of services: the increasing share of services in the consumer basket and the increasing use of services as intermediate inputs.

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 differences in centrality, markups, and business turnover are important drivers of these multiplier differences.

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.

References

- Acemoglu, D., Carvalho, V. M., Ozdaglar, A., and Tahbaz-Salehi, A. (2012). The Network Origins of Aggregate Fluctuations. *Econometrica*, 80(5):1977–2016.
- 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.
- Angrist, J. D. and Pischke, J.-S. (2009). *Mostly Harmless Econometrics: An Empiricist’s Companion*. Number 8769 in Economics Books. Princeton University Press.
- Auerbach, A., Gorodnichenko, Y., and Murphy, D. (2020). Local Fiscal Multipliers and Fiscal Spillovers in the USA. *IMF Economic Review*, 68(1):195–229.
- Baqae, D. R. and Farhi, E. (2019). The macroeconomic impact of microeconomic shocks: Beyond hulten’s theorem. *Econometrica*, 87(4):1155–1203.
- Baqae, D. R. and Farhi, E. (2020). Productivity and Misallocation in General Equilibrium. *The Quarterly Journal of Economics*, 135(1):105–163.
- Barnichon, R., Debortoli, D., and Matthes, C. (2022). Understanding the size of the government spending multiplier: It’s in the sign. *The Review of Economic Studies*, 89(1):87–117.
- 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.
- Bisgaard-Larsen, R., Ravn, S. H., and Santoro, E. (2025). House prices, endogenous productivity, and the effects of government spending shocks. *European Economic Review*, 172.
- 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. (2023). The Government Spending Multiplier in a Multisector Economy. *American Economic Journal: Macroeconomics*, 15(1):209–239.

- Bouakez, H., Rachedi, O., and Santoro, E. (2025). The sectoral origins of heterogeneous spending multipliers. *Journal of Public Economics*, 248:105404.
- Briganti, E., Dwyer, H., Duque Gabriel, R., and Sellemi, V. (2025). Breaking down the us employment multiplier using micro-level data. Staff working papers, Bank of Canada.
- Brinca, P., Holter, H. A., Krusell, P., and Malafry, L. (2016). Fiscal multipliers in the 21st century. *Journal of Monetary Economics*, 77(C):53–69.
- Buchheim, L., Watzinger, M., and Wilhelm, M. (2020). Job creation in tight and slack labor markets. *Journal of Monetary Economics*, 114(C):126–143.
- Cacciatore, M. and Traum, N. (2022). Trade flows and fiscal multipliers. *The Review of Economics and Statistics*, 104(6):1206–1223.
- Chodorow-Reich, G. (2019). Geographic cross-sectional fiscal spending multipliers: What have we learned? *American Economic Journal: Economic Policy*, 11(2):1–34.
- 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.
- 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. (2024). Big g. *Journal of Political Economy*, 132(10):3260–3297.
- 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.
- Duarte, M. and Restuccia, D. (2010). The Role of the Structural Transformation in Aggregate Productivity. *The Quarterly Journal of Economics*, 125(1):129–173.
- 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.
- 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.
- Flynn, J. P., Patterson, C., and Sturm, J. (2021). Fiscal policy in a networked economy. NBER Working Papers 29619, National Bureau of Economic Research, Inc.
- Francois, J. and Hoekman, B. (2010). Services Trade and Policy. *Journal of Economic Literature*, 48(3):642–692.
- 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.
- Galí, J., López-Salido, J. D., and Vallés, J. (2007). Understanding the Effects of Government Spending on Consumption. *Journal of the European Economic Association*, 5(1):227–270.
- Gervais, A. and Jensen, J. B. (2019). The tradability of services: Geographic concentration and trade costs. *Journal of International Economics*, 118(C):331–350.
- 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.
- Hall, R. E. (2009). By How Much Does GDP Rise If the Government Buys More Output? *Brookings Papers on Economic Activity*, 40(2):183–249.
- Hall, R. E. (2018). New Evidence on the Markup of Prices over Marginal Costs and the Role of Mega-Firms in the US Economy. NBER Working Papers 24574, National Bureau of Economic Research, Inc.
- Hazell, J., Herreño, J., Nakamura, E., and Steinsson, J. (2022). The Slope of the Phillips Curve: Evidence from U.S. States. *The Quarterly Journal of Economics*, 137(3):1299–1344.
- Herrendorf, B., Rogerson, R., and Valentinyi, A. (2014). Growth and Structural Transformation. In Aghion, P. and Durlauf, S., editors, *Handbook of Economic Growth*, volume 2 of *Handbook of Economic Growth*, chapter 6, pages 855–941. Elsevier.
- Ilzetzki, E., Mendoza, E. G., and Végh, C. A. (2013). How big (small?) are fiscal multipliers? *Journal of Monetary Economics*, 60(2):239–254.
- Juarros, P. (2020). Fiscal stimulus, credit frictions and the amplification effects of small firms.

- Kolko, J. and Neumark, D. (2010). Does local business ownership insulate cities from economic shocks? *Journal of Urban Economics*, 67(1):103–115.
- Lewis, V. and Winkler, R. (2017). Government Spending, Entry, And The Consumption Crowding-In Puzzle. *International Economic Review*, 58(3):943–972.
- Moretti, E. (2010). Local multipliers. *American Economic Review Papers & Proceedings*, 100(2):373–377.
- Mygrant, M. S. (2020). Keeping profits at home: A study of firm ownership and the geographical concentration of capital gains in the united states. *Local Economy*, 35(5):460–481.
- Nakamura, E. and Steinsson, J. (2014). Fiscal Stimulus in a Monetary Union: Evidence from US Regions. *American Economic Review*, 104(3):753–792.
- Pasten, E., Schoenle, R., and Weber, M. (2020). The propagation of monetary policy shocks in a heterogeneous production economy. *Journal of Monetary Economics*, 116(C):1–22.
- 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.
- Rotemberg, J. J. and Woodford, M. (1999). The cyclical behavior of prices and costs. In Taylor, J. B. and Woodford, M., editors, *Handbook of Macroeconomics*, volume 1 of *Handbook of Macroeconomics*, chapter 16, pages 1051–1135. Elsevier.
- Sanderson, E. and Windmeijer, F. (2016). A weak instrument F-test in linear IV models with multiple endogenous variables. *Journal of Econometrics*, 190(2):212–221.
- Stroebe, J. and Vavra, J. (2019). House Prices, Local Demand, and Retail Prices. *Journal of Political Economy*, 127(3):1391–1436.
- Suárez-Serrato, J. C. and Wingender, P. (2016). Estimating Local Fiscal Multipliers. NBER Working Papers 22425, National Bureau of Economic Research, Inc.
- Woodford, M. (2011). Simple Analytics of the Government Expenditure Multiplier. *American Economic Journal: Macroeconomics*, 3(1):1–35.

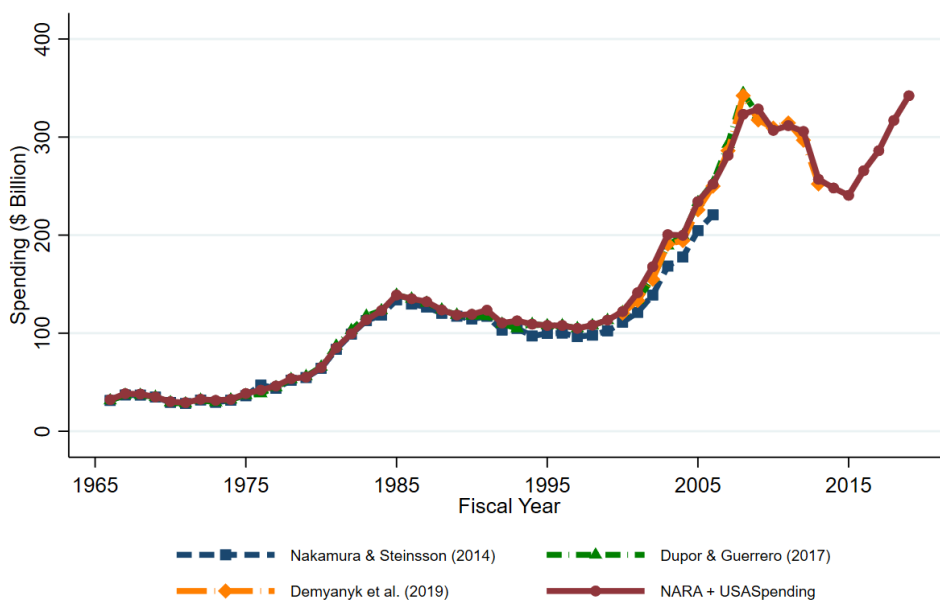
Appendix

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 total 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)	(2)	(3)
	Nakamura and Steinsson (2014)	Dupor and Guerrero (2017)	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
<i>Within</i> – 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 total 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

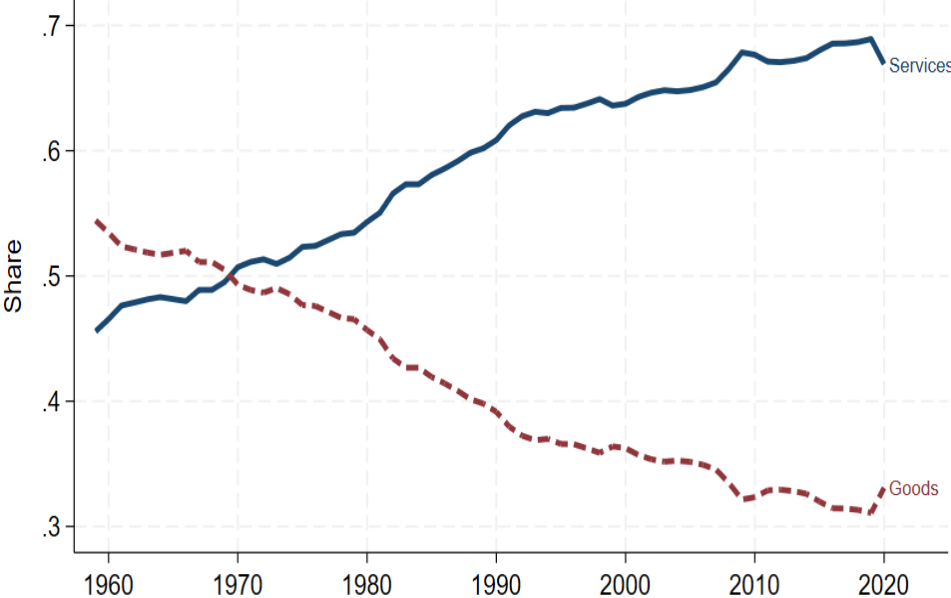
Our dataset includes 3239 distinct 4-digit product codes, of which 2547 are Product Services Codes (PSCs), identifying spending in services, and 692 are the Federal Supply Codes (FSCs), identifying spending in goods. PSCs are grouped into 24 sub-groups, while FSCs are grouped into 85 groups. Table A.2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

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: Impulse Responses - Total Spending

	(1)	(2)	(3)	(4)	(5)
	t	t+1	t+2	t+3	t+4
<i>Panel A: Employment</i>					
Spending	0.092* (0.050)	0.245** (0.114)	0.268** (0.126)	0.361* (0.192)	0.384** (0.181)
F-stat	9.684	22.946	15.961	11.037	12.369
<i>Panel B: Labor Income</i>					
Spending	0.134** (0.064)	0.467*** (0.163)	0.505*** (0.188)	0.657** (0.289)	0.652** (0.270)
F-stat	9.684	22.946	15.961	11.037	12.369
<i>Panel C: GDP</i>					
Spending	0.055 (0.038)	0.452*** (0.148)	0.404** (0.164)	0.559* (0.329)	0.462** (0.220)
F-stat	11.921	37.565	17.718	5.578	10.854

Notes: The panel includes 344 MSAs and consists of 12,024 observations for the period 1979 – 2019 for Panels A and B, and 4,676 observations for the period 2001 – 2019 for Panel C. 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: Impulse Responses - Spending by Components

	(1)	(2)	(3)	(4)	(5)
	t	t+1	t+2	t+3	t+4
<i>Panel A: Employment</i>					
Services Spending	0.246** (0.124)	0.560*** (0.191)	0.722*** (0.212)	0.867*** (0.224)	0.771*** (0.179)
Goods Spending	0.033 (0.033)	0.075 (0.101)	0.065 (0.101)	0.078 (0.136)	0.118 (0.142)
S-W F-stat Services	14.435	86.299	65.444	42.787	40.807
S-W F-stat Goods	21.199	41.414	32.187	30.537	53.276
$\beta_s^k - \beta_g^k$	0.214	0.486	0.657	0.789	0.654
Test $\beta_g^k = \beta_s^k$	[.111]	[.033]	[.008]	[.002]	[.004]
<i>Panel B: Labor Income</i>					
Services Spending	0.307*** (0.081)	0.960*** (0.166)	1.315*** (0.200)	1.528*** (0.272)	1.315*** (0.248)
Goods Spending	0.047 (0.048)	0.157 (0.139)	0.126 (0.139)	0.129 (0.186)	0.143 (0.183)
S-W F-stat Services	14.435	86.299	65.444	42.787	40.807
S-W F-stat Goods	21.199	41.414	32.187	30.537	53.276
$\beta_s^k - \beta_g^k$	0.260	0.804	1.189	1.399	1.173
Test $\beta_g^k = \beta_s^k$	[.011]	[.001]	[.000]	[.000]	[.000]
<i>Panel C: GDP</i>					
Services Spending	0.303** (0.135)	1.031*** (0.270)	1.406*** (0.309)	1.820*** (0.544)	1.592*** (0.438)
Goods Spending	-0.007 (0.048)	0.275 (0.225)	0.118 (0.173)	0.171 (0.269)	0.103 (0.187)
S-W F-stat Services	6.582	47.196	31.130	4.128	7.758
S-W F-stat Goods	18.892	88.286	93.962	23.699	42.772
$\beta_s^k - \beta_g^k$	0.310	0.756	1.288	1.648	1.489
Test $\beta_g^k = \beta_s^k$	[.045]	[.040]	[.000]	[.005]	[.002]

Notes: The panel includes 344 MSAs and consists of 12,024 observations for the period 1979 – 2019 for Panels A and B, and 4,676 observations for the period 2001 – 2019 for Panel C. 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: Cumulative Local Fiscal Multiplier - Spending by Components

	(1)	(2)	(3)	(4)	(5)
	t	t+1	t+2	t+3	t+4
<i>Panel A: Employment</i>					
Services Spending	0.246** (0.124)	0.485*** (0.173)	0.637*** (0.196)	0.780*** (0.205)	0.806*** (0.190)
Goods Spending	0.033 (0.033)	0.065 (0.084)	0.075 (0.096)	0.091 (0.118)	0.115 (0.129)
$\beta_s^k - \beta_g^k$	0.214	0.420	0.563	0.688	0.691
Test $\beta_g^k = \beta_s^k$	[.111]	[.039]	[.014]	[.004]	[.003]
<i>Panel B: Labor Income</i>					
Services Spending	0.307*** (0.081)	0.799*** (0.145)	1.133*** (0.178)	1.392*** (0.209)	1.412*** (0.224)
Goods Spending	0.047 (0.048)	0.132 (0.115)	0.146 (0.133)	0.164 (0.165)	0.179 (0.177)
$\beta_s^k - \beta_g^k$	0.260	0.667	0.987	1.227	1.233
Test $\beta_g^k = \beta_s^k$	[.011]	[.001]	[.000]	[.000]	[.000]
<i>Panel C: GDP</i>					
Services Spending	0.303** (0.135)	0.752*** (0.238)	1.048*** (0.268)	1.342*** (0.317)	1.422*** (0.346)
Goods Spending	-0.007 (0.048)	0.159 (0.143)	0.136 (0.162)	0.155 (0.217)	0.130 (0.207)
$\beta_s^k - \beta_g^k$	0.310	0.593	0.912	1.186	1.291
Test $\beta_g^k = \beta_s^k$	[.045]	[.044]	[.005]	[.002]	[.002]

Notes: The panel includes 344 MSAs and consists of 12,024 observations for the period 1979 – 2019 for Panels A and B, and 4,676 observations for the period 2001 – 2019 for Panel C. 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: Newly-created Jobs

	(1)	(2)	(3)	(4)	(5)
	t	t+1	t+2	t+3	t+4
<i>Panel A: All Industries</i>					
Services Spending	0.582 (0.294)	1.325 (0.451)	1.708 (0.501)	2.050 (0.529)	1.823 (0.424)
Goods Spending	0.077 (0.077)	0.178 (0.239)	0.154 (0.239)	0.184 (0.322)	0.278 (0.334)
<i>Panel B: Transportation & Utilities</i>					
Services Spending	-0.002 (0.019)	0.025 (0.037)	0.060 (0.041)	0.080 (0.042)	0.095 (0.044)
Goods Spending	-0.003 (0.009)	-0.008 (0.012)	-0.009 (0.008)	-0.012 (0.011)	-0.013 (0.015)
<i>Panel C: Construction</i>					
Services Spending	0.032 (0.028)	0.068 (0.039)	0.095 (0.049)	0.120 (0.058)	0.114 (0.056)
Goods Spending	0.016 (0.013)	0.014 (0.021)	0.012 (0.021)	0.021 (0.032)	0.039 (0.036)
<i>Panel D: Manufacturing</i>					
Services Spending	0.164 (0.050)	0.311 (0.085)	0.439 (0.106)	0.604 (0.131)	0.580 (0.134)
Goods Spending	0.003 (0.032)	0.066 (0.047)	0.092 (0.038)	0.132 (0.050)	0.139 (0.048)
<i>Panel E: Wholesale Trade</i>					
Services Spending	-0.005 (0.022)	0.030 (0.028)	0.033 (0.027)	0.050 (0.036)	0.035 (0.030)
Goods Spending	-0.002 (0.008)	0.000 (0.016)	-0.004 (0.010)	-0.010 (0.012)	-0.016 (0.011)
<i>Panel F: Retail Trade</i>					
Services Spending	0.024 (0.029)	0.121 (0.056)	0.189 (0.081)	0.247 (0.086)	0.241 (0.075)
Goods Spending	0.000 (0.017)	0.012 (0.030)	0.014 (0.027)	0.034 (0.035)	0.048 (0.041)
<i>Panel G: Finance</i>					
Services Spending	0.008 (0.018)	0.043 (0.045)	0.073 (0.052)	0.114 (0.058)	0.125 (0.050)
Goods Spending	-0.009 (0.011)	0.003 (0.018)	0.003 (0.013)	0.023 (0.014)	0.041 (0.016)
<i>Panel H: Services</i>					
Services Spending	0.205 (0.080)	0.594 (0.156)	0.768 (0.134)	0.928 (0.180)	0.945 (0.184)
Goods Spending	-0.082 (0.086)	-0.034 (0.115)	-0.066 (0.132)	-0.058 (0.166)	-0.078 (0.226)

Notes: The estimates for each industry are based on 344 MSAs for the period 1979 – 2019. The standard errors are calculated using the delta method.

Table B.5: Impulse Responses - Different Price Adjustments

	(1)	(2)	(3)	(4)	(5)
	t	t+1	t+2	t+3	t+4
<i>Panel A: Employment</i>					
Services Spending	0.258* (0.142)	0.573** (0.227)	0.752*** (0.250)	0.918*** (0.257)	0.820*** (0.204)
Goods Spending	0.030 (0.027)	0.075 (0.091)	0.061 (0.090)	0.065 (0.120)	0.095 (0.122)
$\beta_s^k - \beta_g^k$	0.227	0.498	0.691	0.852	0.725
Test $\beta_g^k = \beta_s^k$	[.127]	[.052]	[.013]	[.003]	[.002]
<i>Panel B: Labor Income</i>					
Services Spending	0.300*** (0.089)	0.986*** (0.192)	1.399*** (0.233)	1.647*** (0.301)	1.414*** (0.276)
Goods Spending	0.042 (0.040)	0.147 (0.126)	0.109 (0.122)	0.100 (0.161)	0.108 (0.157)
$\beta_s^k - \beta_g^k$	0.258	0.839	1.290	1.547	1.307
Test $\beta_g^k = \beta_s^k$	[.013]	[.001]	[.000]	[.000]	[.000]
<i>Panel C: GDP</i>					
Services Spending	0.349** (0.159)	1.254*** (0.323)	1.722*** (0.370)	2.285*** (0.673)	1.987*** (0.533)
Goods Spending	-0.011 (0.043)	0.203 (0.196)	0.080 (0.146)	0.109 (0.204)	0.059 (0.150)
$\beta_s^k - \beta_g^k$	0.361	1.051	1.642	2.176	1.928
Test $\beta_g^k = \beta_s^k$	[.040]	[.008]	[.000]	[.002]	[.001]

Notes: The panel includes 344 MSAs and consists of 12,024 observations for the period 1979 – 2019 for Panels A and B, and 4,676 observations for the period 2001 – 2019 for Panel C. 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: Impulse Responses - Excluding top 1% contracts by MSA-year

	(1)	(2)	(3)	(4)	(5)
	t	t+1	t+2	t+3	t+4
<i>Panel A: Employment</i>					
Services Spending	0.761*** (0.166)	1.120*** (0.275)	1.393*** (0.413)	1.489*** (0.450)	1.330*** (0.405)
Goods Spending	0.147 (0.178)	0.325 (0.282)	0.562 (0.415)	0.831 (0.546)	1.254* (0.694)
$\beta_s^k - \beta_g^k$	0.614	0.795	0.831	0.658	0.075
Test $\beta_g^k = \beta_s^k$	[.021]	[.069]	[.183]	[.376]	[.930]
<i>Panel B: Labor Income</i>					
Services Spending	0.871*** (0.188)	1.685*** (0.318)	2.285*** (0.397)	2.395*** (0.434)	2.027*** (0.373)
Goods Spending	0.280 (0.268)	0.518 (0.434)	0.895 (0.681)	1.250 (0.881)	1.762 (1.083)
$\beta_s^k - \beta_g^k$	0.591	1.167	1.390	1.145	0.265
Test $\beta_g^k = \beta_s^k$	[.106]	[.053]	[.100]	[.265]	[.825]
<i>Panel C: GDP</i>					
Services Spending	1.611*** (0.428)	2.904*** (0.823)	3.084*** (1.110)	3.332*** (1.250)	3.189*** (1.227)
Goods Spending	-0.414 (0.687)	-0.426 (0.705)	-0.183 (0.540)	-0.313 (0.635)	-0.472 (0.660)
$\beta_s^k - \beta_g^k$	2.025	3.331	3.267	3.644	3.661
Test $\beta_g^k = \beta_s^k$	[.023]	[.005]	[.015]	[.015]	[.014]

Notes: The panel includes 344. 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.7: Impulse Responses - Excluding MSAs with large contracts

	(1)	(2)	(3)	(4)	(5)
	t	t+1	t+2	t+3	t+4
<i>Panel A: Employment</i>					
Services Spending	0.247*	0.592***	0.767***	0.931***	0.838***
	(0.132)	(0.201)	(0.221)	(0.229)	(0.175)
Goods Spending	0.018	0.039	0.031	0.034	0.061
	(0.034)	(0.102)	(0.103)	(0.131)	(0.136)
$\beta_s^k - \beta_g^k$	0.230	0.553	0.736	0.897	0.777
Test $\beta_g^k = \beta_s^k$	[.105]	[.019]	[.004]	[.001]	[.000]
<i>Panel B: Labor Income</i>					
Services Spending	0.299***	0.959***	1.327***	1.567***	1.365***
	(0.082)	(0.166)	(0.201)	(0.295)	(0.277)
Goods Spending	0.014	0.084	0.059	0.042	0.040
	(0.052)	(0.140)	(0.139)	(0.171)	(0.168)
$\beta_s^k - \beta_g^k$	0.286	0.874	1.267	1.524	1.325
Test $\beta_g^k = \beta_s^k$	[.007]	[.000]	[.000]	[.000]	[.000]
<i>Panel C: GDP</i>					
Services Spending	0.348**	1.075***	1.399***	1.862***	1.618***
	(0.152)	(0.293)	(0.323)	(0.618)	(0.482)
Goods Spending	0.000	0.282	0.131	0.197	0.125
	(0.045)	(0.230)	(0.176)	(0.285)	(0.195)
$\beta_s^k - \beta_g^k$	0.348	0.793	1.268	1.665	1.493
Test $\beta_g^k = \beta_s^k$	[.038]	[.036]	[.001]	[.010]	[.004]

Notes: The panel includes 199 MSAs. Standard errors are clustered by MSA. The symbols *, **, and *** represent the 10%, 5%, and 1% significance levels, respectively. All regressions include MSA and year-fixed effects.

Table B.8: Impulse Responses - Controlling for the Average Contract Size

	(1)	(2)	(3)	(4)	(5)
	t	t+1	t+2	t+3	t+4
<i>Panel A: Employment</i>					
Services Spending	0.241** (0.121)	0.546*** (0.185)	0.704*** (0.206)	0.848*** (0.220)	0.759*** (0.178)
Goods Spending	0.034 (0.031)	0.083 (0.095)	0.074 (0.098)	0.096 (0.137)	0.129 (0.143)
$\beta_s^k - \beta_g^k$	0.207	0.463	0.630	0.752	0.630
Test $\beta_s^k = \beta_g^k$	[.109]	[.031]	[.008]	[.003]	[.005]
<i>Panel B: Labor Income</i>					
Services Spending	0.302*** (0.078)	0.942*** (0.157)	1.284*** (0.190)	1.497*** (0.261)	1.296*** (0.241)
Goods Spending	0.048 (0.046)	0.168 (0.127)	0.142 (0.131)	0.159 (0.182)	0.159 (0.184)
$\beta_s^k - \beta_g^k$.254	.774	1.143	1.338	1.137
Test $\beta_s^k = \beta_g^k$	[.009]	[.000]	[.000]	[.000]	[.000]
<i>Panel C: GDP</i>					
Services Spending	0.306** (0.136)	1.048*** (0.272)	1.411*** (0.309)	1.813*** (0.540)	1.604*** (0.443)
Goods Spending	-0.008 (0.051)	0.229 (0.233)	0.094 (0.181)	0.134 (0.255)	0.074 (0.194)
$\beta_s^k - \beta_g^k$	0.314	0.820	1.317	1.679	1.531
Test $\beta_s^k = \beta_g^k$	[.046]	[.029]	[.000]	[.004]	[.002]

Notes: The panel includes 344 MSAs. 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.