MACROECONOMIC FRAMEWORKS: RECONCILING EVIDENCE AND MODEL PREDICTIONS FROM DEMAND SHOCKS

Alan J. Auerbach
UC Berkeley and NBER

Yuriy Gorodnichenko*
UC Berkeley and NBER

Daniel Murphy
University of Virginia
Darden School of Business

January 23, 2021

Abstract: How do demand shocks affect the economy? We exploit detailed data on U.S. defense spending to examine a large set of outcome variables in response to well-identified local demand shocks, jointly examining new outcomes (e.g., firm entry and housing rents) and other key macroeconomic outcomes and elasticities that previously have been estimated separately or in settings with weaker identification. We find that government spending crowds in employment, firm entry, private consumption, and labor productivity while also increasing local housing rents. To reconcile the evidence with theory, we study a model of economic slack.

JEL: E32, E62, H5

Keywords: fiscal multiplier, business cycles

Acknowledgement: We thank Chris Boehm, Pascal Michaillat, and Benjamin Schoefer for helpful comments and Mathieu Pedemonte for excellent research assistance.

Corresponding author: Department of Economics, University of California, Berkeley, CA 94720-3880, USA; ygorodni@econ.berkeley.edu

1. Introduction

What are the key adjustment margins of an economy to a demand shock, and how should these adjustments inform macroeconomic models? Partial answers to this question are found across studies that vary in the source of identification and the outcome measure that is studied. Such idiosyncrasies in approaches and empirical settings can render the forest difficult to ascertain through the trees. Indeed, despite decades of research, economists have yet to reach a consensus on how key macro metrics that discipline macro models adjust in response to a demand shock.

Much of this state of ambiguity can be attributed to a dearth of data. In aggregate empirical settings there exists data on a range of relevant outcome variables but identification of exogenous variation ("structural shocks") is notoriously difficult and, more generally, macroeconomic time series have relatively modest variation and quickly run into the curse of dimensionality.[1] Researchers have therefore directed their attention toward subnational variation, which permits improved identification but has the limitation that there is less data on outcomes of interest. For example, a benchmark study with strong identification of local demand shocks, Nakamura and Steinsson (2014), requires a long panel of state-level data (spanning back to the 1960s) and is therefore limited in the outcomes of interest that can be examined.[2]

In this paper we exploit more recent city-level data to offer a comprehensive perspective on the effects of a well-identified demand shock with large variation both at short-run (1-year) frequencies and over longer horizons. Our ability to exploit a large set of outcome variables is made possible by the availability of new city-level data on Department of Defense (DOD) spending, which we use as a source of exogenous variation in local demand. The large cross-sectional variation in the city-level data provides sufficient statistical power to obtain precise

[1] Macroeconomic research has proceeded under two broad approaches. The first approach is model-based inference, whereby researchers take a particular model as given and use data to discipline the model parameters. While providing an internally consistent interpretation of the data, this approach is often criticized for opaque identification and high sensitivity of estimates to alternative assumptions (e.g., Summers 1991). Inference under the second approach is often conducted on times series data (e.g., Galí 1999, Christiano, Eichenbaum, and Evans 2005) for which identification of macroeconomic shocks is difficult. In empirical settings that permit stronger identification, inference is often conducted on a small subset of relevant outcome variables, which limits the ability of these studies to offer a broad perspective on the types of models that are consistent with the data (and hence limits their ability to inform the first approach).

[2] Due to data limitations, few studies have examined the elasticity of outcomes with respect to output in response to well-identified aggregate demand shocks. Nekarda and Ramey (2011) is one of the few, which exploits industry-level variation in government spending to examine components of the labor share. Nakamura and Steinsson (2014) focuses on government spending multipliers. Boehm and White (2019) focus on labor market outcomes (earnings, employment, etc.) but do not examine the response of labor share, prices, output, firm entry, etc. Thus, while close in spirit to this paper, these studies focus on smaller sets of outcome variables.

estimates without relying on extreme events such as World Wars or other major military conflicts. Furthermore, local DOD spending likely does not directly enhance local productivity (as it neither enters as an input in firms' production nor contributes to local public infrastructure), nor does it enter in the utility function of local households. This feature implies that DOD spending affects outcomes of interest only through its effect on output—and hence renders its variation an ideal natural experiment for studying the effects of aggregate demand shocks.

Our comprehensive approach generates estimates for a large set of outcomes of interest that share a common empirical setting and source of identification. The set of outcome variables includes components of key "macro metrics" that are often used to discriminate among competing macroeconomic frameworks—the labor share (which is typically linked to the inverse of markups) and the household labor wedge (between the marginal value of labor to households and firms). These components include hours, employment, wages, output, earnings, and consumption. We also produce the first (to our knowledge) estimates of the effects of a pure (DOD) local demand shock on firm entry and local housing rental prices. The response of firm entry helps to distinguish among competing macroeconomic models (Campbell and Lapham 2002), and the housing rental price response is important for both the local consumer price index (part of the household labor wedge) and more generally for macroeconomic models that emphasize housing as a key transmission mechanism (e.g., Garriga, Kydland, and Sustek 2019).

This paper is a step toward providing a joint perspective on how these and other "macro metrics" respond to exogenous variation in aggregate demand. Our city-level analysis is particularly conducive to studying the transmission of demand shocks because, apart from regional data providing more variation than aggregate data, an open-economy (city) setting imposes fewer unrelated restrictions on joint comovement than does a closed-economy (national setting). For example, in a closed-economy setting, consumption (a component of the household labor wedge) and output (a component of the labor share) are tied together by a goods market clearing condition that need not hold in an open economy. Similarly, at the national level the consumer price index (a component of the labor wedge) is closely linked to the producer price index (a component of the labor share), whereas these indices can differ substantially at the city level. With these comovements less restricted by forces beyond households' and firms' optimality conditions, city-level responses thus in principle are more closely related to the underlying microfoundations of interest.

Our focus on contemporaneous (within-year) conditional outcomes also distinguishes our analysis from traditional analyses based on impulse responses using time series data. Most of our data are available at the annual frequency (and hence not conducive to estimating monthly or quarterly impulse responses), so rather than estimate high-frequency (e.g., quarterly) impulse responses, we instead report effects over one-year and multi-year horizons. Our estimates are relatively stable across horizons (consistent with a DOD shock that is highly persistent), with the notable exception being the local population response, which is significant and positive only at longer (2-year+) horizons.

We find that, in response to a DOD-induced increase in Gross Domestic Product (GDP) at an annual horizon, the labor share falls slightly and the household labor wedge plummets (consumption and hours increase while real worker wages fall). The relatively mild response of the labor share masks a large increase in labor productivity, which tends to decrease marginal costs and increase markups (decrease labor shares). Counteracting this effect on the labor share is an increase in nominal wages. Furthermore, the extensive margin of labor (employment) is by far the strongest adjustment in the labor market (double the response of hours per employee). The local population response is indistinguishable from zero at the one-year horizon (suggesting impediments to short-run labor mobility) but positive over longer horizons. Both firm entry and local housing prices exhibit strong responses.

How do these estimates compare to the predictions of workhorse macroeconomic models? Our empirical setting maps most directly into open-economy macroeconomic models, so we simulate prominent open-economy models for comparison. We also simulate workhorse closed-economy models to compare the elasticity of hours with respect to output, a metric that is independent of openness. We document that two key elements of our empirical evidence challenge each of the workhorse macroeconomic frameworks. First, there is a large increase in labor productivity: a one percent increase in real GDP in response to a demand shock is associated with just over a 0.50 percent increase in hours. Second, there is a large decline in the household labor wedge. The large increase in consumption accompanied by an increase in consumer prices and hours imply a large decrease in the marginal utility of work that, according to prominent macroeconomic frameworks, should contract labor supply and result in small fiscal multipliers (contrary to the large multipliers that we document).[3] We also document an increase in firm entry, a margin that is absent from workhorse macroeconomic models.

Taken together, our results suggest that demand shocks increase output without requiring large increases in worker effort or real worker wages. Positive (DOD) spending shocks improve measured productivity, reduce labor market slack, and induce firm entry. In short, government spending crowds in (rather than crowds out) economic activity. These results add to the evidence from our prior work (based on the same empirical setting) that DOD spending removes slack in the economy. In particular, in Auerbach et al (2020a, henceforth AGM), we focus exclusively on income multipliers and show that DOD spending in an industry does not crowd out activity in other industries (in the same and nearby cities) but rather has small but positive crowding-in general equilibrium effects on other industries. The evidence in this paper expands on the AGM evidence with estimates for a much broader set of outcomes.

As a first step in reconciling theory with the evidence, we therefore turn to a recent class of models for which increases in output are associated with declines in economic slack. Michaillat and Saez (2015) and Murphy (2017) propose models in which aggregate demand shocks increase measured labor productivity and—by reducing idleness—can accommodate increased output without significantly affecting households' marginal rate of substitution between leisure and consumption (hence permitting a large fall in the measured household labor wedge). To evaluate this class of models quantitatively with respect to our empirical evidence, we extend and calibrate the negligible marginal costs (NMC) model in Murphy (2017). The NMC framework is suitable for such an analysis because it can accommodate two prominent features of our empirical evidence, firm entry and local output multipliers greater than unity.[4] In addition to being an important feature

[3] We simulate both open-economy models and closed-economy models and discuss why comparisons to closed-economy models are informative based on recent work that maps consumption responses to fiscal stimulus in an open-economy setting into lower bounds for closed-economy consumption responses (Chodorow-Reich 2019). A key assumption for this mapping is that localized government spending does not cause a reallocation of production factors across locations. This assumption is consistent with the evidence in Auerbach, Gorodnichenko, and Murphy (2020a) that localized spending does not crowd out economic activity in nearby locations, as well as with the evidence presented below that the local population response to local DOD spending is statistically indistinguishable from zero over the short run. In the Appendix we document a significant population response over longer horizons.

[4] Introducing DOD spending in the Michaillat and Saez (2015) model would imply local output multipliers below unity. DOD spending (which does not enter firms' production function nor households' utility function) would tighten the goods market, reducing sales to the private sector. In the NMC framework, output multipliers can exceed unity even in small locations (such that households import all consumer goods) due to a positive firm entry margin.

of our evidence, firm entry is a key margin for evaluating the effectiveness of alternative fiscal policies during a downturn (Auerbach et al. 2020b).

Specifically, we extend the NMC model in Murphy (2017) to include different locations and to incorporate a local housing/land market. We calibrate the extended NMC model to match the DOD share of GDP, the labor share, and the relative response of earnings to the response of GDP, and then use the model to predict the response of macro metrics to a local DOD shock. While stylized in many dimensions, the model can account for key adjustment margins, and notably can explain a large multiplier and large increases in local land prices, consumption, the employment rate, firm entry, and measured labor productivity.

The theoretical exercise suggests that modeling slack (excess capacity, idleness) is important for reconciling theory with evidence. In addition to explaining our evidence, theories of slack can also rationalize a number of other important recently documented empirical regularities. For example, in addition to the industry-level evidence from AGM, fiscal multipliers have been found to be higher during periods of high unemployment (e.g., Auerbach and Gorodnichenko 2012) and during episodes of low aggregate demand (Demyanyk, Loutskina, and Murphy 2019, henceforth DLM).

Our paper is related to various other strands of the literature in addition to the broad set of work on the macro metrics (which we discuss in further detail below). Recent work has used regional data combined with existing theoretical frameworks to infer aggregate relationships (see, e.g., Chodorow-Reich 2019 for a review). We similarly exploit regional variation, but our objectives are different. First, our analysis of regional data establishes a set of conditional contemporaneous (within-year) co-movements for a wide set of indicators (not just fiscal multipliers). Second, we use this information to identify the set of microfoundations and modeling frameworks that are consistent with the data. This exercise provides a crucial input for mapping regional responses to aggregate responses. Indeed, these mappings require a theoretical framework, but which framework should a researcher use? We aim to provide a set of empirical moments to restrict the set of frameworks potentially useful for aggregation exercises.

2. Macro Metrics

Here we outline the various metrics that have been used to distinguish between macroeconomic theories and that we will examine in our empirical analysis of local demand shocks. We also discuss the interpretation of these metrics at the city level relative to an analysis at the national level.

A. Labor Share and the Household Labor Wedge

An expansive body of work has examined the cyclicality of the household labor wedge and, separately, the labor share to evaluate models of the macroeconomy. For example, Nekarda and Ramey (2011) find that industry-level defense spending is associated with no detectable effect on the labor share (which under some conditions is equivalent to the inverse of markups), leading them to reject the textbook New Keynesian model. On the other hand, the evidence in Bils, Klenow, and Malin (2012) is generally indicative of countercyclical markups (procyclical labor shares), consistent with the New Keynesian paradigm. Relatedly, Stroebel and Vavra (2019) and Murphy (2019) find that one component of marginal costs – the gap between nominal prices and nominal input costs – increases in response to demand shocks.

With respect to the household labor wedge, Shimer (2009) documents that it is countercyclical. This is consistent with the procyclical opportunity cost of labor documented in Chodorow-Reich and Karabarbounis (2016) and poses significant challenges to classes of search and matching models based on the Mortensen-Pissarides (1994) framework. Chodorow-Reich and Karabarbounis demonstrate that models typically used for understanding the cyclicality of the employment rate such as Hagedorn and Manovskii (2008) struggle to explain unemployment volatility in the presence of procyclical opportunity costs of work (countercyclical labor wedges).

To fix ideas, consider any environment that admits a representative household. Then the household labor wedge T^H represents the deviation of the household's marginal rate of substitution (between consumption and leisure) from the real wage:

$$\underbrace{H^{\frac{1}{\xi}} I(C)}_{MRS} \times \mathrm{T}^H = \frac{W}{P_C}, \quad (1)$$

where H is hours worked, C is consumption, W is workers' hourly nominal wage, and P_C is the nominal price of the consumption bundle. The function $I(C)$ depends on specifics of the representative household's utility function. For example, $I(C) = C$ in the case of separable preferences with log utility over consumption (such that changes in hours do not affect the marginal utility of consumption at any given level of C) and $I(C) = 1$ is the case of Greenwood, Hercowitz, and Huffman (1988; henceforth GHH) preferences. ξ is the Frisch elasticity of labor supply. Equation (1) is quite general and nests the household behavior implied by most standard neoclassical and New Keynesian models. In neoclassical models, the household labor wedge can arise from distortionary taxes, for example, while in New Keynesian models the household labor wedge can also capture the markup charged by workers when they have market power.

Next consider any environment that admits a representative firm with a production function $Q = ZH^{\beta}$, where $\beta \leq 1$ and Z includes other production factors. Deviations of the marginal product of labor (MPN), $\beta \frac{Q}{H}$, from the real wage are captured by the markup, or "firm labor wedge" T^{F}:

$$\mathrm{T}^{F} = \beta \frac{PQ}{WH}, \tag{2}$$

where P is the price of firms' output. Equation (2) is also quite general and nests the standard neoclassical and New Keynesian frameworks. In the New Keynesian model, for example, $\mathrm{T}^{F} > 1$ represents the markup charged by firms with market power. These equations also nest search models. As discussed in Shimer (2009), models with search-and-matching employment can in principle accommodate any real product wage that is less than the MPN ($\mathrm{T}^{F} \geq 1$) and any real worker wage that is greater than the MRS ($\mathrm{T}^{H} \geq 1$). When there is a common worker and product real wage, it must fall between the MRS and the MPN. The labor wedge and the markup consist of additional macro metrics that have guided theory, including labor productivity Q/H, real household wages W/P_C, and consumption C. Our empirical analysis below will examine how the labor wedge and the markup as well as their components respond to an aggregate demand shock.

Our unit of analysis will be the city, which, as discussed in the introduction, can be beneficial relative to a national analysis for a number of reasons. First, in a national (closed-economy) setting, households' consumption and firms' production are linked by a market clearing condition that in principle might restrict the relative movements in the variables beyond any restrictions implied by households' and firms' first-order conditions. Second, at the national level the consumer price P_C is often the same as the producer price P, whereas at the city (open-economy) level, these can differ. Conditional comovements are therefore less restricted at the city level by forces beyond those that are directly related to the microfoundations of household and firm behavior. Finally, identification is stronger at the city level due to the large and exogenous variation in DOD spending, as discussed below.

We will report many of the macro metrics as elasticities with respect to real GDP to emphasize that the response of these metrics applies generally to a demand-induced increase in output. Much of the prior literature on government spending shocks has reported fiscal multipliers: changes in GDP, employment, or earnings in response to a change in DOD spending. We also examine fiscal multipliers, but they are only a subset of the broader set of metrics of interest. In our setting, fiscal multipliers are relevant in that they verify that DOD spending indeed causes an increase in GDP. Having established that, our primary interest will be in how other metrics change in response to a demand-induced percent increase in GDP. It is important to note that the city-level multipliers we estimate can only be translated into national multipliers (often the object of policy interest) with the use of a theoretical framework (see, e.g., the discussion in Chodorow-Reich 2019). We view our analysis as helping to sort through the relevant frameworks that can be used, for example, to translate regional multipliers into national multipliers.

B. Other Macro Metrics

An additional set of metrics we examine are related to margins of adjustment in the labor market, as intensive and extensive labor market adjustments to demand shocks are relevant for distinguishing among theories of the macroeconomy. For example, motivated by evidence of large fiscal multipliers, researchers have proposed theories in which an increase in aggregate demand can lead to large increases in output due to a large response of hours (e.g., Nakamura and Steinsson 2012), due to a reduction in unemployment (e.g., Christiano et al. 2016), or due to a reduction in firm-level idleness or slack (e.g., Michaillat and Saez 2015; Murphy 2017). To inform the debate on the labor market transmission of demand-driven fluctuations, we decompose changes in hours into changes in hours per worker, changes in the employment rate (1 minus the unemployment rate), and changes in the labor force.[5]

A final metric we examine is firm entry. Procyclical firm entry at the national level has inspired theories in which productivity shocks drive down marginal costs and induce firm entry (e.g., Bilbiie, Ghironi, and Melitz 2012). But it remains unclear whether demand shocks induce firm entry, and if so, how one should interpret such evidence. As Campbell and Lapham (2004) highlight, firm entry in response to demand shocks challenges models driven by sticky prices, as it's unlikely that prices can be changed less frequently than establishments can be opened.

[5] Our city-level analysis admits an additional margin of adjustment relative to a closed economy: the labor force can change due to migration. Inward migration (if any) can generate larger fiscal multipliers than would arise in a closed-economy setting (due to the higher elasticity of labor supply), but it will not affect the interpretation of metrics related to the labor wedge or the labor share. As demonstrated below, the short-run labor force response to DOD spending is statistically insignificant and relatively small compared to other margins of labor force adjustment, suggesting that migration is likely not a relevant force behind our measured output response to DOD spending.

To summarize, our set of macro metrics includes components of households' optimality conditions (the household labor wedge) and firms' optimality conditions (inverse of the labor share), as well as indicators of labor market adjustment (e.g., the unemployment rate) and firm entry. To the best of our knowledge, our study is the first to offer joint evidence on the comovement of these metrics conditional on a well-identified demand shock.

3. Data and Methodology

Our analysis relies on variation in DOD spending, which constitutes over half of discretionary federal government spending. In addition to being a significant force for fiscal stimulus, DOD spending has the advantage that it neither enters directly in households' utility function nor contributes significantly to local productive public infrastructure, thus helping to isolate the potential channels through which it can affect the economy. In particular, it does not appear directly in the household labor wedge (equation (1)) or the markup (equation (2)).

We use a new dataset of city-level DOD spending that allows us to overcome some of the challenges faced in previous work (e.g., limited variation in government spending). Table 1 reports the variation in DOD spending and GDP at different levels of aggregation: city, state, and national. DOD spending is highly variable at the city level, especially compared to the variation in GDP. But DOD spending's standard deviation falls by more than 50 percent when aggregated to the state level and by an order of magnitude when aggregated to the national level. Furthermore, DOD spending is often concentrated in handful of locations in a given state, and we have precise information about the local incidence of government spending shocks. For example, Huntsville (a major center for rocketry) accounts for 70 percent of DOD spending in Alabama.

We complement government spending data with data on a wide range of economic outcomes. Table 2 summarizes these data, their sources, and available time periods. The unit of analysis is city-year, where city is defined as a core-based statistical area (CBSA).[6]

A. Government Spending Data

Our measure of government spending shocks uses data on DOD contracts, available at USAspending.gov. This data source contains detailed information on contracts signed since 2000,

[6] CBSA is geographic area defined by the Office of Management and Budget that consists of one or more counties (or equivalents) anchored by an urban center of at least 10,000 people plus adjacent counties that are socioeconomically tied to the urban center by commuting.

including the name and location (zip code) of the primary contractor, the total contracted amount (obligated funds), and the duration of the contract. In most cases, we also observe the primary zip code in which contracted work was performed.

These data offer several advantages relative to the data used to estimate state-level local fiscal multipliers. First, the detailed location data permit us to estimate multipliers at lower levels of economic geography. This increases the cross-sectional dimension of our study and allows us to examine localized outcome variables for which data are available for only a limited, more recent, period of time. Second, the information on the duration of each contract allows us to construct a proxy for outlays associated with each contract over time. This proxy captures the component of DOD contracts that directly affects output contemporaneously. Also, some of the spending is based on pre-determined contracts, which helps mitigate concerns about endogeneity.[7] AGM and DLM provide further discussion of this data source and the construction of the DOD spending series.

B. Data on Output, Prices, and Labor Market Outcomes

Our measure of output is from the Bureau of Economic Analysis (BEA), which provides data on nominal and real GDP for large cities in the U.S. since 2001. Real GDP is derived by applying chain-type price indices from BEA's industry accounts.[8]

To construct series on wage rates, hours worked, and employment, we rely on the American Community Survey (ACS), which contains information on respondents' city of residence beginning in 2005. Our city-level measures of hours, employment, and the labor force are based on weighted sums of the hours, employment status, and labor force status of respondents in each city. The city-level wage measure is the average of household wages, which are equal to labor income divided by

[7] To construct this spending/outlay measure by location, we derive a flow spending measure for each contract by allocating the contracted amount equally over the contract's duration. For example, for a $3 million contract that lasts three years we assign $1 million in spending for each year of the contract. We then aggregate spending across contracts in a location at each point in time to construct local measures of DOD spending. In addition to new contract obligations, the dataset also contains modifications to existing contracts, including downward revisions to contract amounts (de-obligations) that appear as negative entries. Many of these de-obligations are very large and occur subsequent to large obligations of similar magnitude. Furthermore, in many cases, de-obligations happen within days after obligations appear in the reporting system. When we observe obligations and de-obligations with magnitudes within 0.5 percent of each other, we consider both elements of the pair to be null and void as it is unlikely that any outlays were associated with these temporary obligations. This restriction removes 4.7 percent of contracts from the sample.

[8] The BEA's procedure for computing city-level GDP is available at https://www.bea.gov/sites/default/files/methodologies/GDPMetro2015.pdf. The BEA's city-level measure of real GDP is computed by adjusting nominal GDP by national industry-level price indices, which may not account for the full extent of local price changes in response to a local demand shock, particularly for non-traded goods.

hours worked. We also examine wage residuals derived from a Mincerian regression of wages on observable respondent characteristics, including age, education, occupation, and industry.

Our measure of employee income (pre-tax earnings) comes from the Bureau of Labor Statistics' Quarterly Census of Employment and Wages (QCEW). Consistent QCEW data at the county level are available since 1984. The QCEW data are based on employers' filings with state-level unemployment insurance agencies. In most cases earnings covered by the QCEW include wage earnings, bonuses, stock options, and severance pay. Employer contributions to insurance, pensions, and welfare funds are not included and, as a result, our measured labor share will be lower than in the aggregate NIPA data. Note that CBSAs (our definition of cities) are collections of counties and thus we can exactly match QCEW data to other city-level outcomes.

To infer the response of the city-level consumer price index, we follow Moretti (2011) and exploit data on local housing rents, which account for approximately 40 percent of household expenditure. Direct measures of city-level consumer price indices are generally not available, as the price data collected for the national Consumer Price Index are based on fewer than 40 urban areas and imputed for other cities (e.g., prices for all metropolitan areas in Pennsylvania are imputed with prices collected in Pittsburgh and Philadelphia).[9] In particular, we scale the local rental price response by 40 percent when inferring the consumer price response. This scaled measure accurately captures the local CPI response under the assumption that prices of other expenditure items (primarily tradables) are orthogonal to local DOD spending shocks. If other prices increase in response to a shock, then our scaled measure of local prices provides a lower bound for the actual response of consumer prices (and therefore, any observed decrease in the household labor wedge is a lower bound for the actual decrease in the labor wedge).

To construct a measure of local housing rental prices, we obtain residualized rental prices from the ACS as in Albouy (2012) and Murphy (2018). Residualizing the rental measure controls for observable variation in the quality of housing. In particular, we obtain the housing-cost differential for respondent j in location ℓ using a regression of gross rents, $r_{j\ell}$, on controls ($Z_{j\ell}$) for size, rooms, commercial use, kitchen and plumbing facilities, age of building, home ownership, and the number of residents per room: $\log(r_{j\ell}) = Z_{j\ell}\beta + \epsilon_{j\ell}$. Rents for homeowners are imputed

[9] Other measures, such as scanner data (Nielsen, IRI, etc.), cover 10-15 percent of consumer expenditures (food and small nondurables such as detergents, hygienic products, etc.), most of which have large tradable components that are less responsive to local conditions and often have prices set at the national level (DellaVigna and Gentzkow 2019).

using a discount rate of 7.85 percent (Peiser and Smith 1985). The residuals $\epsilon_{j\ell}$ are the rent differentials that represent the amount individual j pays for her apartment/home in location ℓ relative to the average cost of a similar apartment/home in the U.S. Our city-level measure of rental prices is constructed by averaging these residuals within a city.

Our measure of rental price growth provides a close approximation of rental prices for the subset of cities examined by the Bureau of Labor Statistics in computing the CPI. Among the 38 cities for which we have information on both measures, the median correlation between our rental price growth measure and the BLS measure of the implied cost of housing is 0.61. Furthermore, regressing growth in the BLS city-level CPI on the city-level housing cost measure yields a coefficient of 0.46, which is close to the scaling factor that we use in translating the DOD-induced response of our rental price measure into the response of the local consumer price index.

C. Data on Consumption and Firm Entry

We use a measure of consumption (auto registrations) that is commonly used to study consumption responses to local demand shocks (e.g., Mian, Rao, and Sufi 2013; DLM). This data set is provided by R.L. Polk and contains the number of new automobile registrations in a zip code in a month. The zip code is based on the address of the person who purchases the automobile rather than the address of the dealership. We aggregate the zip code-month-level data to the city-year level. One potential drawback to the auto registrations data is that autos are a durable good for which consumption need not equal expenditure. Mian et al. (2013) report that the growth rate response of credit card purchases on nondurables (another popular proxy for consumer spending) is about a quarter (0.34 relative to 1.31) of the growth rate response of auto purchases to a wealth shock. Therefore, we approximate the consumption response by scaling the auto response by a fourth.[10]

Our measure of firm entry is based on growth in the number of establishments in a city. The U.S. Census Bureau's County Business Patterns provides information on the number of establishments in each zip code. We aggregate their data to derive a series of city-level establishment growth rates. The establishment growth series contains extreme outliers (the

[10] This magnitude of adjustment is consistent with other data. For example, the growth rate of new car registrations is highly correlated (ρ=0.68) with the growth rate of consumer spending at the aggregate level but the standard deviation of the growth rate for new car registrations is approximately 6 times larger than the standard deviation for the growth rate of consumer spending. Regressing the growth rate of new car registrations on the growth rate of consumer spending at the aggregate level yields a slope of 4.

maximum is over 90 times the size of the 99th percentile). To remove the influence of these extreme observations, we winsorize the establishment growth rate series at the bottom and top 0.5 percent.

In addition to these metrics, it would also be useful to have data on capital expenditures, which would inform the extent to which investment responds to local demand shocks. To the best of our knowledge, such data are not available at the city level. In the absence of such data, we refer the reader to the industry-level estimates from Nekarda and Ramey (2011).

D. Econometric Specification

We estimate several econometric specifications to achieve two goals. First, we verify that government spending shocks influence output. Second, we examine how demand-driven changes in output translate into changes in our macroeconomic metrics.

Building on AGM, we use the following specification to achieve the first goal:

$$\frac{\Delta Y_{\ell t}}{Y_{\ell,t-1}} = \beta \frac{\Delta G_{\ell t}}{Y_{\ell,t-1}} + \psi_\ell + \alpha_t + error_{\ell t}, \quad (3)$$

where ℓ and t index locations (CBSA) and time (year), Y is a measure of output, G is a measure of defense spending, and ψ_ℓ and α_t are location and time fixed effects. Coefficient β measures the local DOD spending impact multiplier, that is, the dollar amount of output produced by a dollar of local DOD spending.

We instrument for variation in government spending $\frac{\Delta G_{\ell t}}{Y_{\ell,t-1}}$ using a Bartik instrumental variable (IV) shock, $\frac{\widehat{\Delta G}_{\ell t}}{Y_{\ell,t-1}} \equiv \frac{s_\ell \times (G_t - G_{t-1})}{Y_{\ell,t-1}}$, where s_ℓ is the location's average share of DOD contract spending over the relevant period and G_t is aggregate contract spending in period *t*. As discussed in AGM, the Bartik IV addresses potential endogeneity concerns (such as the influence of political factors, as discussed in Nakamura and Steinsson 2014) *and* isolates the component of DOD contracts that is actually associated with new production. Many DOD contracts represent payment for new production as well as payment for production that would have occurred anyway, either because the specific contract was anticipated or because firms smooth over lumpy contracts. AGM argue that the Bartik IV approach isolates the relevant component of $\frac{G_{\ell,t+h} - G_{\ell,t-1}}{Y_{\ell,t-1}}$ associated with *new* production by using information on changes in national production.[11] In addition, we report estimated specification (1) with $\frac{s_\ell \times (G_t - G_{t-1})}{Y_{\ell,t-1}}$ as a regressand (i.e., in the reduced form):

$$\frac{\Delta Y_{\ell t}}{Y_{\ell,t-1}} = \beta' \frac{\widehat{\Delta G}_{\ell t}}{Y_{\ell,t-1}} + \psi_\ell + \alpha_t + error_{\ell t}. \qquad (3')$$

Once we establish that β' is not zero, we can use the following specification to hit the second goal:

$$\frac{\Delta X_{\ell t}}{X_{\ell,t-1}} = \gamma \frac{\Delta Y_{\ell t}}{Y_{\ell,t-1}} + \psi_\ell + \alpha_t + error_{\ell t}, \qquad (4)$$

where X is an outcome variable of interest and $\frac{\Delta Y_{\ell t}}{Y_{\ell,t-1}}$ is instrumented with our Bartik shock $\frac{s_\ell \times (G_t - G_{t-1})}{Y_{\ell,t-1}}$ (specifically, (3') is the first-stage of (4)). To make our results consistent with our simulations of workhorse macroeconomic models, we will use real GDP as a measure of Y on the right-hand side of specification (4). Coefficient γ informs us about how variable X reacts to changes in output that are driven by demand-side factors.

Our identifying assumption (exclusion restriction) is that DOD spending affects changes in variable X only though its effect on GDP. We find this assumption to be reasonable given that local DOD spending is unlikely to directly enter households' utility functions or as inputs to firms' production functions at business cycle frequencies. Thus, we view violation of this assumption as an unlikely threat to our identification. Our approach can be also undermined if government spending does not affect output but affects other variables. For example, if labor supply is perfectly inelastic, a demand shock does not influence output but raises wages. This kind of problem should manifest itself in a weak first stage (equation (3)), which is not the case in our results.

We report effects over a one-year horizon in the body of the paper, as our key objective is to uncover the immediate comovements of different variables in response to a demand-induced change in output. The Appendix presents analogous results for effects over longer horizons.

4. Empirical Results

This section presents the effects of a local DOD shock on a range of outcomes, including the macro metrics and their components (e.g., for the household labor wedge, we examine effects on real

[11] One potential limitation of our IV approach is that it does not rule out all anticipation effects, as it is possible that local households forecast local demand based on anticipated changes in national DOD spending. Any unobserved anticipation effects would likely bias our estimates downward, as some changes in household spending would occur in advance of the observed local demand shock.

worker wages, hours, and consumption). We begin by presenting estimates of fiscal multipliers. Our estimates are above 1, consistent with strong output effects of DOD spending. We then examine the elasticities of our macro metrics with respect to the DOD-induced increase in GDP.

A. Baseline GDP and Income Multipliers

To begin our discussion of the effects of DOD spending shocks on different components of income, we show, in Table 3, the differential impact of government spending shocks on GDP. In columns 1 and 3, we use the Bartik shock as an instrument for spending. The first column of the table shows the impact response of nominal GDP to a spending shock. The third column shows the impact on real GDP. The magnitudes of the estimated fiscal multiplier (1.05-1.10) are similar to estimates of city-level multipliers (e.g., AGM and DLM). Columns 2 and 4 report results for reduced form regressions (3'). We find that the Bartik shock is a strong predictor of output changes.[12]

B. Labor Share (Firm Labor Wedge) and the Output-to-Labor Ratio

How does the earnings share respond to a demand shock relative to GDP? To assess the degree of comovement between output and earnings in response to demand shocks, we first estimate specification (4) with the change in earnings normalized by nominal GDP as the regressand and nominal GDP growth as the regressor. We find (column 2 of Table 4) that the change in earnings with respect to the change in GDP is 0.36.[13] This estimate is below the average labor share (0.41, column 1), implying that labor shares are mildly countercyclical, although we cannot reject the null hypothesis of a change equal to the average labor share. We arrive at a similar conclusion when we instead use earnings growth as the regressand (column 3 of Table 4). The estimated elasticity of earnings with respect to GDP is 0.93, and we cannot reject the null of a unit elasticity. Finally, in row 1 of Table 5 we examine the response of a direct measure of the labor share to demand-driven changes in GDP. Our estimate is negative but not statistically significant from zero. Each of these estimates points to a labor share that is approximately acyclical or mildly countercyclical and hence (under the assumption that the average wage equals the marginal wage)

[12] Our focus is on impact (one-year-horizon) effects of demand shocks, but it is informative to note that the effects of a government spending shock are very stable when considering longer-horizon changes on output and DOD spending. Table A1 reports 2-year and 3-year multipliers (in which the denominator of the multiplier is based on changes in DOD spending over the full horizon). The estimates are very similar to those reported in Table 3 and the long-horizon multipliers reported in AGM.

[13] We also examine earnings from the ACS and find that its response to a DOD shock (0.32) is of similar magnitude to the response of QCEW earnings.

a markup that is approximately acyclical or mildly procyclical. This finding is consistent with the industry-level evidence in Nekarda and Ramey (2011) and the time-series evidence in Hall (2009) and Karabarbounis (2014).

The inverse of the labor share $\frac{PQ}{WH}$ consists of a productivity component (Q/H) and a relative-price component (P/W). In row 2 of Table 5 we use specification (4) to examine the productivity component of the labor share. A percent increase in real GDP due to a demand shock is associated with a 0.57 percent increase in hours (or, equivalently, an elasticity of output with respect to hours of 1.75). This result suggests that labor productivity strongly increases in response to a demand shock, which all else equal drives down marginal costs (and the labor share $\frac{WH}{PQ}$) and pushes up the markup. Our evidence reinforces the evidence in Nekarda and Ramey (2011) that aggregate labor productivity increases in response to positive demand shocks, although our estimates of the productivity response are much larger, pointing to a strong increase in capital utilization rates and/or labor effort and declines in labor hoarding. Our finding of a relatively acyclical labor share along with strongly procyclical labor productivity suggests that the (producer) price-to-wage ratio is countercyclical.[14]

C. Employment

To understand sources of changes in output, we now study the reaction of various labor margins. Row 2 of Table 5 reports that the estimated elasticity of hours with respect to output is 0.57. When we decompose this elasticity into the extensive margin (the number of employees) and intensive margin (the number of hours per employee), we find that the bulk of the elasticity is accounted by the extensive margin (the elasticity of the number of employees is 0.41, row 3) rather than the intensive margin (the elasticity of hours per employee is not statistically different from zero, row 4). This estimate is similar to that in e.g. Shimer (2009), which is based on unconditional correlations. The fact that our estimates from identified demand shocks coincide with correlations over the business cycle is consistent with the notion that demand shocks account for the majority of business cycle fluctuations (e.g., Michaillat and Saez 2015).

[14] We do not have a direct measure of the local GDP deflator and therefore do not report its response in Table 5. The estimated response of the implied GDP deflator (by adjusting nominal GDP by real GDP) is an imprecisely estimated zero, which is consistent with an increase in real product wages.

To further explore the role for the extensive margin of employment, we decompose the change in employment into two components: changes in the labor force and changes in the employment rate. Rows 5 and 6 of Table 5 show the estimated elasticities for the employment rate and for the labor force. The employment rate response is economically and statistically significant. The labor force response is positive but not statistically different from zero.

We can further decompose the response of the labor force to changes in the labor force participation rate (row 7) and changes in population (row 8). We find that none of these margins has a statistically significant response to demand-driven changes in GDP at the one-year horizon. The mild population response is consistent with the evidence in Boehm and White (2019) and Suárez Serrato and Wingender (2016) that migration has low predictive power for fiscal multipliers and suggest that the relatively large city-level multiplier is driven by firm and labor market adjustments other than population inflows. It is also consistent with the relatively low degree of geographic mobility documented over recent decades (e.g., Molloy et al., 2016) and the assumption underlying the translation from open-economy multipliers to closed-economy multipliers in Chodorow-Reich (2019).[15] However, when we examine longer-horizon responses (reported in the Appendix), we detect a strong and significant population response. One implication is that translating local multipliers into national multipliers using longer-horizon data may require adjusting for the local population response.

D. Household Labor Wedge

The household labor wedge consists of nominal wages, the local CPI, consumption and hours. Using data from the American Community Survey, we use the nominal wage rate as a dependent variable in specification (4). The wage response to demand-induced changes in output is positive but not statistically different from zero (row 9 of Table 5). This "raw" wage response may be affected by changes in the composition of workers (e.g., Solon, Barsky, and Parker 1994). To address this concern, we use residual nominal wages, based on the Mincerian regression described in Section 2. The response of this composition-adjusted measure of nominal wages is larger than the "raw" wage response but also noisy (row 10).

[15] Appendix Table 2 shows that there is a strong population response over longer horizons, consistent with empirical evidence in Zou (2018).

Local consumer prices, on the other hand, exhibit an economically large and significant response. The estimates of the wage response (0.11) and the CPI response (0.26) imply a *negative* real household wage response.

Exacerbating the contribution of real household wages to a negative household labor wedge is a strong increase in consumption (row 12 of Table 5).[16] The rise in consumption, along with the rise in hours (row 4) should, given standard preference assumptions, increase the marginal valuation of leisure. Hence, the substantial decline in the real worker wage along with the increases in consumption and hours indicate a large decline in the household labor wedge.

The negative labor wedge response we document is consistent with the cyclicality of the labor wedge observed in aggregate data. Shimer (2009) discusses possible explanations of the countercyclical labor wedge, including cyclical tax policy, time-varying work disutility, and time-varying labor market power. An advantage of our setting is that (instrumented) local DOD spending is plausibly orthogonal to local preferences, market power, and taxes. Given that these factors are not responsible for the labor wedge response we document, they may also be unlikely candidates for understanding the aggregate time series cyclicality.

E. Firm Entry

How can labor productivity increase so much in response to a short-run increase in output induced by the DOD shock? Potential explanations include adjustments in labor effort (and/or a reduction in labor hoarding), increased capital utilization, and endogenous firm entry. Notably, Devereux, Head, and Lapham (1996) predict that, assuming increasing returns in production, even wasteful government spending can increase measured labor productivity (and lead to large multipliers) by inducing firm entry. We find (row 13 of Table 5) that the elasticity of the number of local establishments with respect to output is 0.15, meaning that some of increased employment occurs at new establishments. While much of the prior theoretical literature has focused primarily on technology shocks as drivers of entry (e.g., Bilbiie, Ghironi, and Melitz 2012), here we document

[16] The local consumption increase is qualitatively consistent with regional evidence from government purchases in Dupor et al. (2019). They find that a dollar increase in spending from the American Recovery and Reinvestment Act (ARRA) is associated with \$0.29 of additional local consumption. To compare our estimate to theirs, we can convert our estimated consumption elasticity into a multiplier: $\frac{\%\Delta C}{\%\Delta Y} = \frac{dC}{dY}\frac{Y}{C}$. Setting $\frac{C}{Y} \approx 0.6$ (based on the consumption share of national income) and noting that our multiplier estimate implies that $dY = 1.05 dG$, our estimate of the consumption elasticity implies that $\frac{dC}{dG} = 1.05 \times \frac{\%\Delta C}{\%\Delta Y}\frac{C}{Y}$. This is higher than the estimate in Dupor et al. One way to reconcile these different estimates is to note that the ARRA was temporary, whereas DOD spending is strongly persistent.

that expansionary *demand* shocks also increase firm entry. The entry response that we estimate is relatively mild (compared, e.g., to the employment response) but nonetheless economically and statistically meaningful.

Assessment. Overall, we view Table 5 as presenting benchmark elasticities that can be used to evaluate macroeconomic models in an open-economy setting. These estimates are also useful for informing the microfoundations underpinning closed-economy models. For example, the small elasticity of hours with respect to output (large productivity response) implies restrictions on production functions that apply in open- or closed-economy settings. Similarly, the large decline in the household labor wedge implies that models (closed or open) should incorporate "wedges" large enough to accommodate large declines in response to an increase in demand. The household labor wedge may not ultimately fall as much in a closed-economy setting (if, for example, a goods-market-clearing condition or reactive monetary policy limits the extent to which consumption increases), but any limited response should not be due to a constraint implied by the household's first-order-condition with respect to hours and consumption.

Below we demonstrate how our estimates can guide us in determining the class of models that are consistent with the data.

F. Comparison of Empirical Evidence to Predictions of Workhorse Macro Models.

Here we evaluate our empirical evidence against the predictions of workhorse "medium-scale" macroeconomic models. Since our empirical setting is a small open economy, we begin by evaluating prominent open-economy models. Our baseline open-economy setting is the model in Nakamura and Steinsson (2012), which includes DOD spending across different regions bound by common monetary policy. They present a simple version of their model with only sticky prices, as well as extended versions that include GHH preferences. Our city-level empirical analysis is perhaps more analogous to a small-open-economy setting in that local taxes are nearly independent of local DOD spending and that a large share of local consumption is spent on imported goods. Therefore, we also examine the small-open-economy medium-scale NK model of Galí and Monacelli (2016). City-level DOD shocks are analogous to export shocks in a small-open-economy setting (in that local taxes do not finance the increased production), so we focus on the effects of export shocks in the Galí-Monacelli model.

Table 6 (columns 1 through 3) presents the effects of government spending shocks in these prominent open-economy macro models alongside our empirical estimates. Each model tends to do well by some metrics but poorly by others. For example, the Nakamura-Steinsson model with GHH preferences can match the large multiplier (row 1), but it also performs the worst in matching the increase in hours (row 2) and nominal wages (row 3). The Galí-Monacelli model performs well in matching the mild responses of wages to a shock but it underpredicts the multiplier and cannot match the strong decline in the household labor wedge.

A notable limitation of these open-economy models is that they do not include search frictions or other margins of adjustment that could improve their fit to the data. To assess the evidence against prominent existing models with these additional features, we must turn to closed-economy models. An important caveat when comparing our estimates to moments implied by closed-economy models is that the models are constrained by restrictions (e.g., market-clearing conditions) that do not apply to our empirical evidence. In evaluating the closed-economy models, therefore, we will focus on the metric that is based only on the production function and is independent of openness – the elasticity of hours with respect to output. We will then summarize other dimensions of the comparison between the model and our empirical evidence and discuss what can be learned from such a comparison.

Smets and Wouters (2007) is the basis for many of the medium-scale closed-economy models and therefore serves as our reference closed-economy model. It notably includes sticky wages, sticky prices, and variable capital utilization. We examine both the baseline medium-scale Smets-Wouters (henceforth SW) model, as well as a version of their model with flexible prices and wages and fixed capital utilization (the "neoclassical model"). Government spending in the model is similar in nature to DOD spending in that it does not enhance local productivity and it does not enter directly into the utility function.

Because the SW model does not have a well-defined notion of involuntary unemployment (and hence of "slack"), we also use a quantitative search-and-match model of business cycles developed in Christiano, Eichenbaum, and Trabandt (2016, henceforth CET). This medium-scale, closed-economy model has many frictions similar to those in earlier New Keynesian models (price stickiness, adjustment costs, habit in consumption, etc.). Finally, we use the FRB/US model developed by the staff of the Federal Reserve to examine whether this closed-economy model heavily employed for policymaking can rationalize the empirical patterns.

The SW and CET models nearly match the elasticity of hours with respect to output, which lends support to their underlying production functions. If anything, the hours response is too strong (0.63 and 0.65, respectively, versus 0.57 in the empirical evidence), suggesting that the models understate the firm-level margin of adjustment. The extensive margin of employment is also too strong in the CET model (0.66 versus 0.33).

Turning to moments for which the open-closed distinction is relevant, the starkest difference between the model and the results is in the response of consumption (and, relatedly, the household labor wedge). The empirically estimated consumption elasticity is 1.1 but -0.3 in both the SW and CET models. Chodorow-Reich (2019) provides a detailed discussion of the relationship between closed-economy and open-economy consumption effects of government spending. Among the forces that tend to pull down national effects relative to local (open-economy) effects are (a) factor reallocation across locations in response to localized government spending, (b) government-spending-induced interest rate increases at the national level, and (c) larger tax liabilities associated with national spending. Our evidence of an insignificant short-run population response to local spending suggests that (a) does not materially contribute to lower closed-economy multipliers.[17] With respect to (b), a large body of empirical evidence finds that interest rates in the United States are either unresponsive (or even negatively responsive) to government spending shocks (see Murphy and Walsh 2020 for a review of this evidence). This suggests that only (c) could in principle account for lower national consumption multipliers. However, this effect would need to be incredibly strong to account for the large estimated local consumption response and the negative consumption response implied by the closed-economy models.

Row 11 provides a summary measure of each model's fit relative to the data, equal to the simple sum of distance measures for individual metrics. Since different macro metrics (e.g., the labor share and the household labor wedge) share component metrics (e.g., hours and nominal wages), we focus only on subcomponent metrics in deriving our summary metric. These subcomponent metrics are hours, nominal wages, consumption, and the local CPI. We reiterate that the comparison between our evidence (based on an open-economy setting) and the closed-economy models should be interpreted with caution. Nonetheless, the arguments from Chodorow-Reich and the evidence from prior literature that interest rates do not rise in response to government

[17] Relatedly, AGM provide related evidence that increases in local DOD spending do not crowd out economic activity in nearby areas.

spending suggest that the open-economy consumption response should provide a closer estimate of the closed-economy consumption response than one might otherwise suspect. The summary measures imply that the medium-scale New Keynesian models and the search model offer the closest fit to the data, whereas the neoclassical model and the model with GHH preferences offer predictions that are substantially different from the data.

To summarize, each of these models can accommodate different aspects of the data but performs poorly along other dimensions. Variable capital utilization, rather than GHH preferences alone, seems important for capturing the increase in labor productivity alongside the large multiplier. Some variation on wage rigidity appears necessary to capture the decline in the household labor wedge. A common shortcoming of many models is that they tend to predict a fall in consumption, whereas empirically it increases. A standard rationalization of a positive consumption response is to have real worker wages increase (typically alongside credit-constrained households). But we find that real wages fall due to a large increase in the cost of housing.

More generally, the large increase in measured labor productivity and the decline in the household labor wedge imply that government spending increases production and firm entry without requiring large increases in worker wages or hours, which points to the existence of slack both within firms and/or in labor markets. Consistent with this notion, in prior work (AGM) we examine industry-level effects of DOD spending (in an otherwise identical empirical setting) to document that DOD spending in a particular industry in a city does not (on average) crowd out other industries in the city or in nearby locations. This prior evidence, in conjunction to our new broader set of estimates, motivates us to turn to models of economic slack as a way to reconcile theory with evidence.

5. Macroeconomic Implications

In this section, we sketch a framework that can accommodate key features of our empirical results, including i) an extensive margin of employment, ii) within-firm variation in labor productivity, iii) large fiscal multipliers, and iv) a decline in real worker wages along with increases in consumption and hours. Our objective is to outline a setting that is simple yet capable of capturing these key margins of adjustment in response to a local demand shock. Therefore, we include only the model elements that are necessary to capture the various adjustment margins. We evaluate the quantitative performance of the model and discuss further extensions that would improve its quantitative fit.

Our proposed setting builds on the notion of labor as a (quasi-)fixed factor (Oi 1962). Despite the treatment of labor as a marginal cost in contemporary macroeconomic models, the field has long acknowledged the potential relevance of fixed labor. For example, studies have documented that workers often supply their labor in increments (Card 1990; Faber 2005) and that firms often operate in regions of fixed-only costs (Brown 1992; Rotemberg and Summers 1990). Survey evidence also adds support to the notion of fixed costs and excess capacity. For example, the Census' Quarterly Survey of Plant Capacity reports utilization rates that are typically well below the full capacity. For example, even in the very low-unemployment period 2019Q4, the utilization rate across manufacturing industries was 71.4, which was similar to the rate for the industry that receives the majority of DOD spending (utilization in the Aerospace and transportation equipment industry was 76.8).

Murphy (2017) derives implications of "negligible marginal costs" (fixed factors of production) in a general equilibrium setting. Here we extend this framework to a small-open-economy setting and to include government spending, employment and firm entry margins, and land/housing (which is the main input into the local CPI). Agents inelastically supply labor, and firms hire workers as fixed costs. Employed workers can increase output costlessly (i.e., they do not demand higher wages for additional hours or effort).

A. Model

We extend the NMC framework of Murphy (2017) to explicitly include employment and firm entry margins. We also model different locations (cities), each of which is treated as a small open economy that exports goods to other locations and sells to the national government. In the extended model, local residents purchase tradable goods as well as a locally-endowed nontradable good that accounts for land and other immobile factors of production.

Households. The economy consists of locations ("islands") indexed by $\ell \in [0,1]$. In each location there is a representative household that consists of a mass of N workers, indexed by $\omega_\ell \in [0, N]$. The workers seek employment with local tradable-sector firms indexed by $j_\ell \in [0, J_\ell]$. Workers remit income to the household and the household consumes.

The household in location ℓ maximizes

$$U_\ell = \sum_{t=0}^{\infty} u_{\ell t}$$

where

$$u_{\ell t} = r\int_0^N 1(\omega_\ell)d\omega_\ell + \log\mathcal{L}_{\ell t} + \int_0^1\int_0^{J_m}\left[\theta_{jm\ell t}q_{jmlt} - \frac{1}{2}\gamma q_{jm\ell t}^2\right]dj\,dm, \quad (5)$$

$\mathcal{L}_{\ell t}$ is a locally endowed nontradable good ("land"), and $q_{jm\ell t}$ is location $\ell's$ consumption of the perishable tradable of variety j_m produced in location m. These Melitz and Ottaviano (2008) preferences over tradables give rise to demand curves with price-dependent demand elasticities, which is a necessary condition for equilibrium slack in the NMC framework. $1(\omega_l)$ indicates non-employment (e.g., household work), which is valued at r and captures the notion of indivisible labor as in Hansen (1985).

The household's within-period budget constraint is

$$\int_0^1\int_0^{J_m} p_{jmt}q_{jm\ell t}dj\,dm + \mathcal{L}_{\ell t}p_{\ell t}^{\mathcal{L}} + T_{\ell t} = \int_0^N w_{\omega\ell t}\,d\omega_\ell + \Pi_{\ell t} + I_{\ell t}, \quad (6)$$

where $w_{\omega\ell}$ is the total wage earnings of worker ω_ℓ, p_{jmt} is the price of variety j_m, $T_{\ell t}$ represents lump-sum taxes, and $p_{\ell t}^{\mathcal{L}}$ is the price of the local nontradable.[18] $I_{\ell t}$ represents other sources of income, including from ownership of non-labor local factors of production. $\Pi_{\ell t}$ is profits from owning firms on island ℓ and other islands. We assume that land and firm ownership is diversified so that households on island ℓ derive negligible income from owning land or firms on island ℓ. Let $\lambda_{\ell t}$ be the multiplier on location ℓ's budget constraint.

Tradable Sector Production and Demand. Each firm in the tradable sector hires workers as a fixed cost. To operate and produce, each firm requires a mass of n tasks (the input into the fixed labor cost is a Leontief technology over tasks). Each task requires an employee, and the total amount of perishable output that the n workers can produce is the capacity level $\bar{q}$, which for simplicity of aggregation is assumed to be constant across firms and is so high that it is not binding. At output levels below $\bar{q}$, output can be increased without additional costs to the firm, consistent with the notion of labor as a quasi-fixed factor (Oi 1962).

A firm's revenues depend on the demand curve it faces. Household optimization implies that demand from island ℓ for variety j_m (output produced by a firm located on island m) is

$$q_{jm\ell t}^d = \frac{1}{\gamma}\left(\theta_{jm\ell t} - \lambda_{\ell t}p_{jmt}\right). \quad (7)$$

[18] The budget constraint (6) implies that all firm profits are returned to the household as dividends each period.

Total private-sector demand for variety j_m is derived by integrating across locations:

$$q_{jmt}^{d} = \int_0^1 q_{jm\ell t}^{d} d\ell = \frac{1}{\gamma}\left(\theta_{jmt} - \lambda_t p_{jmt}\right),$$

where $\theta_{jmt} \equiv \int_0^1 \theta_{jm\ell t} d\ell$ and $\lambda_t \equiv \int_0^1 \lambda_{\ell t} d\ell$. When firms operate below the capacity level, a firm maximizes revenues by choosing a price $p_{jmt} = \theta_{jmt}/2\lambda_t$, which implies that the quantity sold to the private sector is

$$q_{jmt}^{p} = \frac{\theta_{jmt}}{2\gamma} \tag{8}$$

and revenues from the private-sector are

$$R_{jmt}^{P} = \frac{\theta_{jmt}^{2}}{4\gamma\lambda_t}.$$

Nontradable Sector. The nontradable goods in each location are produced competitively using a locally endowed commodity, which represents land or other factors of production that are immobile across locations and across sectors. Therefore, local consumption of the nontradables is invariant over time and independent of tradable sector output.

Government Spending. The government purchases tradable goods from the private sector. We assume that it spends ϕ_{mt} proportion of (potential) private-sector revenues on j_m across all firms on island m, which implies that demand from the government is

$$q_{jmt}^{G} = \phi_{mt} R_{jmt}^{P} \frac{1}{p_{jmt}} \tag{9}$$

This assumption captures the fact that the DOD spends more on large firms such as Boeing than on smaller firms.[19] Note that the government has a unit elasticity of demand and therefore does not influence the price, p_{jmt}, as variation in the firm's price has no impact on its profits from government sales. This is consistent with the fact that the government buys output at the price determined by the private market. Firms are willing to accept these extra purchases at the market price because they have spare capacity and can costlessly increase output.

Given equation (9), total firm revenues are

[19] We assume that the government purchases from all firms. This is a reduced-from way of capturing the fact that spending on a subset of firms/industries has strong effects on local upstream suppliers that are not the direct recipients of the contracts, as documented in AGM.

$$R_{jmt} = \frac{\theta_{jmt}^2}{4\gamma\lambda_t}(1 + \phi_{mt}). \tag{10}$$

Hiring. In each period, each of the tasks across firms is randomly matched with a worker. Firms employ labor locally, that is, a firm located on island ℓ can hire workers only from island ℓ. If a wage contract is agreed upon, the employment relationship lasts for the duration of the period. There is only one opportunity to match with a firm each period, so matched workers' opportunity cost of accepting a wage offer is the reservation utility r. The benefit to the firm of agreeing on an employment contract is a firm's revenue minus the wage,

$$V_{j\ell t}^F = \lambda_t(R_{j\ell t} - w_{j\ell t}),$$

where $w_{j\ell t}$ is the wage bill paid by firm j_ℓ to each worker with which it is matched (specifically, $w_{\omega\ell t}$ equals $w_{j\ell t}$ if worker ω_ℓ is matched with firm j_ℓ). We assume that ownership of firms is distributed across all islands and therefore firms value profits at the average marginal utility of income islands, $\lambda_t = \int_0^1 \lambda_{\ell t} d\ell$,.

The benefit to the worker of accepting a contract is

$$V_{j\ell t}^W = \lambda_{\ell t}(w_{j\ell t} - r).$$

where r is the value of not working. Workers value income at the local household's marginal utility of income. The household does not coordinate bargaining between firms and workers.

Workers and firms Nash bargain over the surplus. The equilibrium wage bill maximizes the product of the benefit to the worker and the benefit to the firm:

$$w_{j\ell t} = \underset{w_{j\ell t}}{\operatorname{argmax}} \left\{ \lambda_{\ell t}^{\psi}(w_{j\ell t} - r)^{\psi} \lambda_t^{1-\psi}(R_{j\ell t} - w_{j\ell t})^{1-\psi} \right\},$$

where ψ is the workers' bargaining power. At an interior optimum, the resulting wage income of a worker is

$$w_{j\ell t} = \psi[R_{j\ell t} + r] \tag{11}$$

and the total wage bill faced by the firm is $nw_{j\ell t}$.

Note that, because labor is hired as a fixed factor, workers bargain over the *wage bill* rather than the *wage rate* (compensation per hour). In other words, firms can ask employees to work more or less without adjusting the total payment to employees. Whether workers need to work more to accommodate demand shocks is unspecified in the model and likely varies according to the nature of the service provided. But, in any event, stated hours of work may not vary, to the extent that full-time work is regarded as a discrete outcome.

Firm Entry and Exit. Firms shut down if their revenues are not sufficient to cover the wage bill, $R_{j\ell t} < nw_{j\ell t}$. This occurs when

$$R_{j\ell t} < \frac{n\psi r}{(1-n\psi)} \equiv \underline{R}.$$

In that case, no wage contract is signed, the firm shuts down, and workers matched with the firm are unemployed for the period. Equation (10) implies that surviving firms are those that face a sufficiently strong private and/or public demand for their goods:

$$\theta_{j\ell t}^2 > \frac{4\gamma\lambda_t \underline{R}}{1+\phi_{\ell t}}.$$

Equilibrium. Aggregate local outcomes in the model depend on the distribution of revenues across firms (and hence the distribution of preference parameters). We assume $\theta_{j\ell t}^2$ is distributed Pareto with a shape parameter α. Because firms exit when profits are negative, the distribution of $\theta_{j\ell t}^2$ for existing firms has a lower support equal to the break-even level of demand shifter $\underline{\theta}_{\ell t}^2 \equiv \frac{4\gamma\lambda_t \underline{R}}{1+\phi_{\ell t}}$. For simplicity of aggregation, we assume that capacity levels are infinite. Then is it straightforward to show (see the Appendix) that the mass of surviving firms on island ℓ is

$$J_{\ell t} = \left[\frac{4\gamma\lambda_t \underline{R}}{1+\phi_{\ell t}}\right]^{-\alpha} = \left[\underline{\theta}_{\ell t}^2\right]^{-\alpha}, \tag{12}$$

the employment rate is

$$\frac{nJ_{\ell t}}{N} = \frac{n}{N}\left[\frac{4\gamma\lambda_t \underline{R}}{1+\phi_{\ell t}}\right]^{-\alpha} = \frac{n}{N}\left[\underline{\theta}_{\ell t}^2\right]^{-\alpha},$$

and total tradable sector revenue (which is equal to GDP) is

$$R_{\ell t} = (4\gamma\lambda_t)^{-\alpha}\frac{\alpha}{\alpha-1}\left[\frac{\underline{R}}{1+\phi_{\ell t}}\right]^{1-\alpha}(1+\phi_{\ell t}). \tag{13}$$

Note that λ_t is endogenous, although its value is pinned down by the household's first-order condition with respect to land, along with the fact that land is endowed (exogenous). The model's numeraire is the aggregate land price (average land prices across locations), so export prices and local land prices are relative to the aggregate land price.

B. Comparative Statics

Here we outline the various adjustment margins in response to an increase in local government spending $\phi_{\ell t}$ that is not financed with increased taxes in location ℓ. The expressions for these comparative statics are derived in the Appendix.

Entry and Employment Response. One can show that the reaction of employment to a change in ϕ is given by

$$d \log Emp_{\ell t} = d \log J_{\ell t} = \frac{\alpha \phi_\ell}{1 + \phi_\ell} d \log \phi_{\ell t}.$$

Government spending causes an increase in firm entry, as the additional revenue from the government causes more firms to produce with positive profits. Since employment increases with firm entry in the model, the increase in government spending also increases employment.

Note that the model assumes that employment is mediated through firm entry, which is an extreme assumption. An alternative and less restrictive setup would be to assume that some varieties represent worker task sets rather than establishments. In that case, some of the employment increase would occur through new task sets within incumbent firms. For example, a firm can open a new conveyer line or a new shift within an existing establishment.

GDP Multiplier. A dollar increase in government spending raises GDP (equal to revenue R) by

$$\frac{dR_{\ell t}}{dG_{\ell t}} = 1 + \frac{\alpha - 1}{\alpha \phi + 1}.$$

When the government spends on a tradable good produced by an incumbent firm, it increases revenues one-for-one with each dollar spent because there is no crowding out of private demand (firms have spare capacity and the marginal cost of producing extra output is negligible). The government also spends on new firms that can export both to the government and to the private sector. The additional private-sector exports from new firms imply that the local government spending multiplier in the model is strictly greater than 1. Note that the high fiscal multiplier is not driven by household income multipliers, since locally produced tradable goods are a negligible share of the consumption bundle.

Labor Share. The additional revenues generated by the government are allocated between firm owners and workers according to workers' bargaining power. In percentage terms, owners receive a slightly larger increase due to the fact that worker earnings include a reservation wage bill that does not adjust with government spending: $\frac{d \log PQ}{d \log WH} = \frac{R+r}{R}$. If the reservation wage r is small relative to firm revenues, then the share of labor income in GDP is approximately independent of demand shocks.

Prices. Firm-level export prices are independent of local government spending. Government spending does induce entry of some lower-value (low θ) products into the market. Land prices $p_{\ell t}^{\mathcal{L}}$ increase due to the increased demand associated with increased local income. The Appendix derives the local nontradable price response around a steady state in which households balance their budget.

Local Consumption. The increase in government spending increases worker earnings, which causes workers to import more tradable goods. Since the local economy is small, import prices do not increase to offset this increase in local consumption. Thus, the percent change in spending on imported goods is equal to the percent change in quantities of imported goods. We show in the Appendix that the change in the consumption of each imported good is equal to the change in spending on "land."

Household Labor Wedge. Households in the model do not experience any disutility from extra work hours or effort but rather only an opportunity cost of employment. Therefore, the model does not exhibit the traditional cost-benefit tradeoff that is captured in the household labor wedge in standard models. Nonetheless, the NMC model makes predictions about the variables (real worker wages, hours and consumption) that are typically used to infer the labor wedge from the data (e.g., Shimer 2009; Karabarbounis 2014). Each of these components of the household labor wedge adjusts in our model to contribute to a large fall in the measured household labor wedge.[20]

C. Calibration

To calibrate the model, we assign one value (α) from a previous study and we infer other parameters from average DOD spending shares, average labor shares, average housing expenditure shares, and an empirical estimate from Table 4. We then use the calibrated parameters to predict the response of various macro metrics (computed around a symmetric equilibrium in which government spending is equally distributed across locations). The model's numeraire is the average land price across locations, and we normalize the land quantity in each location to unity.

[20] The computation of the labor wedge requires elasticities of nominal wages and hours. We assume that there is no intensive margin adjustment of hours (although the model does not rule out an intensive margin adjustment) and therefore set the hours elasticity equal to the employment elasticity. The nominal wage is computed as the total local wage bill divided by local employment.

The parameter of the Pareto distribution for firm size is based on Axtell (2001): $\alpha = 1.05$. Table 7 shows the data moments that are used to calibrate other model parameters. In row 1, we pin down the government demand parameter ϕ by matching the model-implied value of the DOD contract spending share of GDP to the average share in the U.S. during the 2000s (approximately 0.01). In row 2, we use the estimate of $\frac{dWH}{dPQ}$ from column 2 of Table 4 to pin down workers' bargaining power $n\psi$. To pin down the remaining parameters r and γ, in rows 3 and 4, we set the model-implied values of the labor share and the housing expenditure share to their counterparts in the data (both approximately 0.4).

D. Assessment of the NMC Model

Table 8 compares estimated elasticities from the data with predictions from the calibrated model. Note that none of the moments in Table 8 are targeted. The predicted multiplier and elasticity of earnings with respect to GDP are nearly identical to their empirically estimated counterparts. Other predicted metrics are reasonably close to the empirical estimates. In particular, the elasticity of employment with respect to output is well below 1, consistent with the empirical evidence that labor productivity increases in response to an expansionary demand shock. The model predicts a large (relative to existing theories) response of local consumer prices and consumption of imports. While the predicted consumption response is less than the estimated response of consumption, it is in line with estimates from other studies that have examined the effect of other forms of government spending on local consumption (e.g., Dupor et al. 2019).

The model predictions that deviate the most from the data are the large response of establishments and the negative response of average wages. Both of these responses are based on the strong simplifying assumption in the model that all additional employment occurs through firm entry. These new employees work for marginal (low-revenue) firms and therefore have low earnings and low implied wages (earnings per worker), driving down average earnings and wages.

A straightforward modification to the model would temper the employment response and generate a positive average wage response while possibly enhancing the correspondence to reality: the establishment response would be muted and wage response increased if some export varieties represent worker task sets rather than establishments. In that case, some of the employment increase would occur through new task sets within incumbent firms that pay a larger wage bill. The higher earnings would also increase the consumption and land price responses, further improving the fit to the data.

Given the parsimony of the model, one should not expect the model to match all moments of the data (and indeed the model misses some moments quantitatively). In spite of these simplifications, the model offers a surprisingly strong overall fit to the data. In particular, the total distance measure is 18.5, considerably below the distance measures from the workhorse models in Table 6.[21] Thus, we view the NMC model as having sufficient potential to develop it further.

E. Aggregate Implications

Aggregate GDP is

$$R^A = \int_0^1 R_\ell d\ell.$$

In a symmetric equilibrium, $\phi_{\ell t} = \phi_{mt} = \phi_t \; \forall m \neq \ell$, which implies that $R^A = R_\ell$. This implies that a change in ϕ_t across islands has the same effect on national GDP as a change in local $\phi_{\ell t}$ on local GDP, and hence the national multiplier equals the local multiplier.

This result seems counterintuitive given that national DOD spending is financed with taxes across islands ($T_t = G_t$, where $G_t = \int_0^1 G_{\ell t} d\ell$ and an expression for local DOD spending $G_{\ell t}$ is provided in the Appendix), whereas a location's tax does not respond to local increases in government spending. In our model, the different tax responses imply different responses for land prices and consumption, but not for GDP.

Consider first an increase in national spending that is financed by taxes on the recipients of the DOD spending. The income side of the budget constraint increases from the rise in DOD spending. Income also rises as new firms enter and sell some of their output to the private sector. Incumbent firms' sales to the private sector remain fixed (as dictated by equation (8)). On the expenditure side of the budget constraint, taxes increase (the expenditure side of the budget constraint) by an amount equal to the increase in G. Expenditure on newly available tradable goods equals the increase in private-sector income generated by the new firms. On net, national income (and expenditure) increases from government spending and, in addition, from the production of new tradable goods.

Next consider an increase in local government spending that is financed externally (and thus does not require a local tax increase). As in the prior case, the income side of the budget constraint increases from the income from the DOD. Income also rises as new firms enter and sell some of their

[21] As columns 3-7 of Table 8 show, large variations in individual parameters have relatively small impacts on the model's overall fit, with the exception of the Pareto parameter, α.

output to the private sector in the form of exports. External demand for incumbent varieties is independent of local government spending, so there is no increase in exports of incumbent varieties. The expenditure side of the budget constraint differs from the prior case. Here, taxes do not offset the increase in income from the DOD. Instead, the local household spends more on housing and on imports of incumbent varieties. This is possible because, while average (across locations) land prices are pinned down as the numeraire, the land price in any given locale can deviate from the national average. And, the fact that local land prices can deviate implies that the local budget multiplier $\lambda_{\ell t}$ can also deviate from the average. The deviation of $\lambda_{\ell t}$ permits local consumption of an import variety to deviate from aggregate consumption of that variety (equation (7)).

Therefore, while the consumption and land price responses differ at the local level from the national level, these responses do not affect local GDP and hence the national multiplier equals the local multiplier in this model. Prior empirical work (e.g., AGM, DLM, Dupor and McCrory 2018) has found that multipliers increase with the size of the economic geography considered, reflecting positive spillovers across nearby highly localized economies. This suggests that national multiplier might exceed local multipliers. Extensions to the model would likely increase the aggregate multiplier relative to the local multiplier. For example, the model does not feature general equilibrium income multipliers that would, in a framework with negligible marginal costs, tend to push up national multipliers.

F. Discussion and Interpretation

Attempts to rationalize the various metrics examined here have a long tradition in macroeconomics. Based on challenges faced in doing so, prior researchers have called for new frameworks. For example, Hall (2009) calls for "new ideas outside the New Keynesian framework to explain the high value of the multiplier along with other mysteries of aggregate behavior." Likewise, Shimer (2009) encourages macroeconomists to "look beyond search models for an explanation of the labor wedge."

The NMC framework is a step in accounting for these and other macro metrics. A key feature of the NMC model is that additional firm-level production is costless (over some range of output). Therefore, labor productivity depends on firm demand. Likewise, there is no disutility from work hours or work effort (but rather a discreet opportunity cost of employment). This implies that the opportunity cost of employment does not vary with hours or with consumption; and, as a result, workers' labor supply does not contract during expansions.

The microfoundation underlying the NMC framework is that firms face fixed, rather than marginal, labor costs.[22] This treatment of the production process is consistent with observations of labor markets cited above and with lumpy adjustment costs. As an example, it could represent barbers who provide additional haircuts without incurring marginal costs, up until the point at which (s)he is working an eight-hour day (reaches capacity). The firm cannot choose a capacity level of less than a single barber because of the nature of the service provided. Because the barber supplies more labor than is demanded, the barber experiences no marginal disutility from serving additional customers. Indeed, psychology research (e.g., Hsee, Yang and Wang 2019) suggests that people may prefer to be busy rather than bored and so the opportunity cost of leisure may be lower.[23] Of course, if firms systematically overwork employees, employees will demand a higher wage bill, which will involve higher wages.

6. Conclusion

We exploit detailed data on local DOD spending to assess the effects of government purchases on a range of "macro metrics" that are used to distinguish among macroeconomic theories. Our results indicate that, in response to an expansionary demand shock, (a) the labor share is relatively constant, (b) measured labor productivity increases drastically, and (c) the increase in hours is primarily due to adjustment on the extensive margin (employment). Furthermore, (d) the real worker wage falls due to an increase in local rental prices. Accompanying the fall in real worker wages is (e) an increase in local consumption that, along with the increase in hours, contributes to a decline in the household labor wedge. Furthermore, (f) demand shocks induce firm entry. As a first step toward reconciling theory with the evidence, we expand a theory of negligible marginal costs to incorporate government spending and extensive margins of employment and firm entry. The model, while stylized in many dimensions, provides a strong fit with the data.

Our study suggests a number of fruitful avenues for future research. First, we have examined and sought to explain macro metrics in response to government expenditure shocks. It would be helpful to assess these metrics in response to other shocks, including supply-side shocks.

---

[22] The NMC framework is compatible with marginal costs in the form of intermediate inputs and can accommodate marginal costs for firms as long as workers provide their labor as a fixed (rather than marginal) cost. See Murphy (2017) for an extended discussion of the assumptions underlying the NMC framework.

[23] Even if a firm can choose its capacity level and it can perfectly forecast demand, it may choose a level of capacity such that it experiences slack the majority of the time if demand is variable. See Fine and Freund (1990) for a general formalization of optimal capacity investment under demand uncertainty.

Second, further theoretical extensions to the NMC framework or alternative frameworks such as Michaillat and Saez (2015) may prove useful for better matching the data and for welfare analysis. For example, the baseline framework presented here abstracts from income effects that can contribute to high national multipliers and large labor productivity responses.

References

Albouy, David, 2012. "Evaluating the Efficiency and Equity of Federal Fiscal Equalization," *Journal of Public Economics* 96(9-10): 824-839.

Auerbach, Alan J., and Yuriy Gorodnichenko. 2012. "Measuring the Output Responses to Fiscal Policy." *American Economic Journal: Economic Policy* 4(2): 1-27.

Auerbach, Alan J., Yuriy Gorodnichenko, and Daniel Murphy. 2020a. "Local Fiscal Multipliers and Fiscal Spillovers in the USA." *IMF Economic Review* 68(1): 195-229.

Auerbach, Alan J., Yuriy Gorodnichenko, and Daniel Murphy. 2020b. "Fiscal Policy and COVID19 Restrictions in a Demand-Determined Economy." NBER WP 27366.

Bilbiie, Florin O., Fabio Ghironi, and Marc J. Melitz, 2012. "Endogenous Entry, Product Variety, and Business Cycles," *Journal of Political Economy* 120(2): 304-345.

Bils, Mark, Peter J. Klenow, and Benjamin A. Malin. 2012. "Testing for Keynesian Labor Demand." National Bureau of Economic Research Working Paper 18149.

Blanchard, Olivier and Jordi Galí. 2010. "Labor Markets and Monetary Policy: A New Keynesian Model with Unemployment." *American Economic Journal: Macroeconomics* 2(2): 1-30.

Blanchard, Olivier and Lawrence F. Katz. 1992. "Regional Evolutions." *Brookings Papers on Economic Activity*. 3-75.

Boehm, Christopher, and Neil White. 2019. "Labor Market Implications of Fiscal Shocks." Mimeo.

Brown, Charles. 1992. "Wage Levels and Method of Pay." *The RAND Journal of Economics*. 23(3): 366-375.

Campbell, Jeffery R. and Beverly Lapham. 2004. "Real Exchange Rate Fluctuations and the Dynamics of Retail Trade Industries on the U.S.-Canada Border." *American Economic Review*. 94(4): 1194-1206.

Card, David. 1990. "Labor Supply with a Minimum Hours Threshold." *Carnegie-Rochester Conference Series n Public Policy*. 33:137-168.

Chodorow-Reich, Gabriel. 2019. "Geographic Cross-Sectional Fiscal Spending Multipliers: What Have We Learned?" *American Economic Journal: Economic Policy*. 11(2): 1-34.

Chodorow-Reich, Gabriel and Loukas Karabarbounis. 2016. "The Cyclicality of the Opportunity Cost of Employment." *Journal of Political Economy* 124(6): 1563-1618.

Christiano, Lawrence, Martin Eichenbaum, and Charles Evans. 2005. "Nominal Rigidities and the Dynamic Effecst of a Shock to Monetary Policy." *Journal of Political Economy* 113(1): 1-45.

Christiano, Lawrence J., Martin Eichenbaum, and Mathias Trabandt. 2016. "Unemployment and Business Cycles." *Econometrica* 84(4): 1523-1569.

DellaVigna, Stefano, and Matthew Gentzkow, 2019. "Uniform Pricing in US Retail Chains," forthcoming in *Quarterly Journal of Economics*.

Demyanyk, Yuliya, Elena Loutskina, and Daniel Murphy. 2019. "Fiscal Stimulus and Consumer Debt." *Review of Economics and Statistics* 101(4): 728-741.

Devereux, Michael B., Allen C. Head, and Beverly J. Lapham. 1996. "Monopolistic Competition, Increasing Returns, and the Effects of Government Spending." *Journal of Money, Credit and Banking* 28(2): 233-254.

Dupor, Bill, and Peter B. McCrory. 2018. "A Cup Runneth Over: Fiscal Policy Spillovers from the 2009 Recovery Act." *Economic Journal* 128: 1476-1508.

Dupor, Bill, Marios Karabarbounis, Marianna Kudlyak, and M. Saif Mehkari. 2019. "Regional Consumption Responses and the Aggregate Fiscal Multiplier." Mimeo.

Erceg, Christopher J and Andrew T. Levin. 2014. "Labor Force Participation and Monetary Policy in the Wake of the Great Recession." *Journal of Monetary Credit and Banking* 46(2): 3-49.

Faber, Henry S. 2005. "Is Tomorrow Another Day? The Labor Supply of New York City Cabdrivers." *Journal of Political Economy*. 113: 46-82.

Fine, Charles H. and Robert M. Freund. 1990. "Optimal Investment in Product-Flexible Manufacturing Capacity." *Management Science* 36(4): 449-466.

Foster, Lucia, John Haltiwanger, and Chad Syverson. 2008. "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?" *American Economic Review* 98(1): 394-425.

Galí, Jordi. 1999. Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations? *American Economic Review* 89(1): 249-271.

Galí, Jordi and Tommasso Monacelli. 2016. “Understanding the Gains from Wage Flexibility: The Exchange Rate Connection.” *American Economic Review* 106(12): 3829-3868.

Galí, Jordi, Mark Gertler, and J. David López-Salido, 2007. "Markups, Gaps, and the Welfare Costs of Business Fluctuations," *Review of Economics and Statistics* 89(1): 44-59.

Garriga, Carlos, Finn E. Kydland, and Roman Šustek. 2017. “Mortgages and Monetary Policy.” *Review of Financial Studies*. 30(10): 3337-3375.

Greenwood, Jeremy, Zvi Hercowitz, and Gregory W. Huffman (1988) "Investment, Capacity Utilization, and the Real Business Cycle," *American Economic Review* 78(3): 402-17.

Hall, Robert E. 2009. “By How Much Does GDP Rise if the Government Buys More Output?” *Brookings Papers on Economic Activity* (Fall 2009), 183-231.

Hansen, Gary D., 1985. "Indivisible Labor and the Business Cycle," *Journal of Monetary Economics* 16(3): 309-327.

Hsee, Christopher K., Adelle X. Yang, and Liangyan Wang (2010). “Idleness Aversion and the Need for Justifiable Busyness.” *Psychological Science* 21(7): 926–930.

Karabarbounis, Loukas. 2014. “The Labor Wedge: MRS vs. MPN.” *Review of Economic Dynamics*. 17:206-223.

Melitz, Marc J., and Gianmarco I. P. Ottaviano, 2008. "Market Size, Trade, and Productivity," *Review of Economic Studies* 75(1): 295-316.

Mian, Atif, Kamalesh Rao, and Amir Sufi. 2013. “Household Balance Sheets, Consumption, and the Economic Slump.” *Quarterly Journal of Economics* 128(4): 1687–1726.

Michaillat, Pascal and Emmanuel Saez. 2015. “Aggregate Demand, Idle Time, and Unemployment.” *Quarterly Journal of Economics* 130(2): 507-569.

Molloy, Raven, Christopher L. Smith, Riccardo Trezzi, and Abigail Wozniak. 2016. “Understanding Declining Fluidity in the U.S. Labor Market.” *Brookings Papers on Economic Activity*. 2016 (Spring): 183-237.

Moretti, Enrico. 2011. “Local Labor Markets.” In *Handbook of Labor Economics*, edited by David Card and Orley Ashenfelter, 1237-1313. Amsterdam: North Holland.

Mortensen, Dale T., and Christopher A. Pissarides, 1994. "Job Creation and Job Destruction in the Theory of Unemployment," *Review of Economic Studies* 61(3): 397-415.

Murphy, Daniel P. 2017. “Excess Capacity in a Fixed-Cost Economy.” *European Economic Review* 91: 245-260.

Murphy, Daniel P. 2018. “Home Production, Expenditure, and Economic Geography.” *Regional Science and Urban Economics* 70, 112–126.

Murphy, Daniel P. 2019. “Demand Complementarities and Cross-Country Price Differences.” *Canadian Journal of Economics* 52(1): 253-279.

Murphy, Daniel P. and Kieran Walsh. 2020. “Government Spending and Interest Rates.” Mimeo.

Nakamura, Emi and Jón Steinsson. 2014. “Fiscal Stimulus in a Monetary Union: Evidence from US Regions.” *American Economic Review* 104(3): 753-792.

Nekarda, Christopher J. and Valerie A. Ramey. 2011. “Industry Evidence on the Effects of Government Spending.” *American Economic Journal: Macroeconomics* 3(1): 36-59.

Oi, Walter. 1962. “Labor as a Quasi-Fixed Factor.” *Journal of Political Economy* 70(6): 538-555.

Peiser, Richard B., and Lawrence B. Smith, 1985. "Homeownership Returns, Tenure Choice and Inflation," *Real Estate Economics* 13(4): 343-360.

Suarez Serrato, Juan Carlos and Philippe Wingedener. 2016. “Estimating Local Fiscal Multipliers.” NBER Working Paper 22425.

Summers, Lawrence H. (1991) “The Scientific Illusion in Empirical Macroeconomics” *Scandinavian Journal of Economics* 93(2): 129-148.

Shimer, Robert. 2009. “Convergence in Macroeconomics: The Labor Wedge.” *American Economic Journal: Macroeconomics* 1(1): 280-297.

Smets, Frank and Rafael Wouters. 2007. “Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach.” *American Economic Review* 97(3): 586-606.

Solon Gary, Robert Barsky, and Jonathan A. Parker, 1994. “Measuring the Cyclicality of Real Wages: How Important is Composition Bias?” *Quarterly Journal of Economics* 109(1): 1-25.

Stroebel, Johannes and Joseph Vavra. 2019. “House Prices, Local Demand, and Retail Prices.” *Journal of Political Economy*, 127(3).

Zou, Ben. 2018. “The Local Economics Impacts of Military Personnel.” *Journal of Labor Economics.* 36(3):589-621.

Table 1. Variation in Department of Defense (DOD) spending and GDP by level of aggregation

	City	State	National
DOD Spending	0.0118	0.0052	0.0006
GDP	0.0461	0.0360	0.0217

Note: This table reports the standard deviations of GDP growth rates and DOD spending changes (normalized by lagged GDP) at various levels of economic geography.

Table 2. Data Sources

Data	Source	First year city-level data are available, from1998
Department of Defense spending	Federal Procurement Data System/USAspending.gov	1998
GDP (nominal and real)	Bureau of Economic Analysis	2001
Earnings	Quarterly Census of Employment and Wages	1998
Hours, employment, and labor force	American Community Survey	2005
Wages	American Community Survey	2005
Rental Prices	American Community Survey	2005
Establishments	County Business Patterns	2003
Auto Registrations	Polk	2002

Note: A smaller subset of cities have establishment data prior to 2003. We focus on post-2003 establishment data to maintain a balanced panel.

Table 3. Response of output (GDP) to government spending shocks

	Nominal GDP, $\Delta Y_{it}/Y_{i,t-1}$		Real GDP, $\Delta Q_{it}/Q_{i,t-1}$	
	(1)	(2)	(3)	(4)
$\Delta G_{it}/Y_{i,t-1}$	1.054**		1.103**	
	(0.508)		(0.469)	
$\widehat{\Delta G}_{it}/Y_{i,t-1}$		0.956***		0.996***
		(0.488)		(0.355)
Observations	5,605	5,610	5,605	5,610
R-squared	-0.077	0.004	-0.110	0.005
1st stage F-stat	10.18		10.18	

Notes: This table presents estimates from Specifications 3 and 3'. Estimates in columns 1 and 3 are for the IV specification in expression 3; those in columns 2 and 4 are for the reduced form specification in expression 3'. $\widehat{\Delta G}_{it}/Y_{i,t-1}$ is the Bartik instrument. Fixed effects for CBSA and year are included but not reported. Standard errors clustered by state are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 4. Labor Share Response to DOD Shocks

	Average Labor Share $\frac{WH}{PQ}$	Change in Earnings relative to change in GDP $\frac{dWH}{dPQ}$	Elasticity of Earnings w.r.t. GDP $\frac{d \log WH}{d \log PQ}$
	(1)	(2)	(3)
	0.41	0.360***	0.934***
		(0.068)	(0.178)
N	5,984	5,610	5,595
1st-stage F-Stat		20.41	19.52

Notes: Columns 2 and 3 are based on regressions of earnings on GDP, where GDP is in growth rates and is instrumented with the Bartik shock. In column 2, the change in earnings is normalized by lagged GDP (so that the coefficient captures relative changes). In column 3, the change in earnings is normalized by lagged earnings (so that the coefficient captures elasticities). In columns 2 and 3, time and city fixed effects are included but not reported. Standard errors clustered by state are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 5. Response to demand-driven changes in output

Row	Outcome variables	Coef. (s.e.)	1st stage F stat
1	$dlog\left(\frac{Earnings}{GDP}\right)$, labor share	-0.066 (0.144)	17.45
2	$dlog\ (H)$, hours	0.571*** (0.169)	9.98
3	$dlog(E)$, employment	0.409*** (0.143)	9.98
4	$dlog\left(\frac{H}{E}\right)$, hours per employee	0.152 (0.093)	9.98
5	$dlog\left(\frac{E}{L}\right)$, employment rate	0.332*** (0.134)	9.98
6	$dlog(L)$, labor force	0.087 (0.101)	9.98
7	$dlog\left(\frac{L}{Pop}\right)$, labor force participation rate	0.012 (0.084)	9.98
8	$dlog(Pop)$, population	0.075 (0.127)	9.98
9	$dlog(wage)$, wages	0.080 (0.236)	9.98
10	$dlog(resid\ wage)$, residualized wages	0.114 (0.134)	9.98
11	$0.4 \times dlog(resid\ rental\ price)$, CPI	0.261** (0.118)	9.98
12	$0.25 \times dlog(auto)$, consumption	1.101*** (0.251)	17.98
13	$dlog(establishments)$, firm establishments	0.151** (0.070)	10.43

Notes: This tables presents estimates based on specification (2). The regressor is growth in real GDP (instrumented by the Bartik shock). Time and city fixed effects are included but not reported. Standard errors clustered by state are in parentheses. The number of observations for row 1 is 5,610. The number of observations for rows 2-11 is 2,817. The number of observations for row 12 is 4,092. The number of observations for row 13 is 4,114. *** p<0.01, ** p<0.05, * p<0.1.

Table 6. Comparison of empirical and model-implied moments

			Open Economy			Closed Economy			
			Nakamura and Steinsson (2014)		Galí and Monacelli (2016)	Smets and Wouters (2007)		Cristiano et al. (2016)	FRB US
Row	Outcome	Our empirical results	Baseline (sticky prices)	Baseline with GHH preferences	Medium-scale NK with nontradables	Medium-scale NK	Neoclassical	Search and match	Mix
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1)	Output multiplier dQ/dG	1.10	0.79	1.50	0.90	0.88	0.44	0.92	0.97
	Elasticity with respect to output:								
(2)	Hours	0.57	1.49	1.49	1.35	0.63	1.25	0.65	0.54
(3)	Nominal wages	0.11	1.35	1.57	0.05	0.14	1.74	0.25	0.87
(4)	Consumer Price Index	0.26	0.07	0.08	0.00	0.09	1.99	0.11	0.01
(5)	Consumption	1.10	-0.77	0.87	-0.03	-0.30	-1.49	-0.27	0.06
(6)	Real worker wage	-0.15	1.28	1.49	0.05	0.05	-0.25	0.14	0.86
(7)	Household labor wedge	-1.82	0.00	0.00	0.05	-0.56	0.00	-0.24	-0.26
(8)	Employment Rate	0.33	NA	NA	NA	NA	NA	0.66	0.50
(9)	Firm Entry	0.15	NA	NA	NA	NA	NA	NA	NA
(10)	Capital stock	NA	0.00	0.00	0.00	-0.02	-0.12	-0.03	0.04
(11)	Fit: Distance Measure		179.0	158.67	46.72	33.7	503.0	33.1	56.2

Notes: The neoclassical version of the Smets-Wouters model is the version of the model with flexible wages and prices, no variable capital utilization, and no fixed output cost. In the Galí-Monacelli simulation, the multiplier is with respect to an export shock (dY/dX) rather than a government spending shock (dY/dG). In computing the empirical household labor wedge, we assume separable preferences ($I(C) = C$ in equation 1) and a Frisch elasticity of 1. The household labor wedge in the models is based each model's parameterization of the utility function. In the case of Christiano et al., labor is supplied inelastically, so there is no MRS. Therefore, we compute the "labor wedge" in that model as we do in the data (as if there were separable preferences). All responses measure the cumulative reaction of a given variable over one year after a government spending shock, which is equal to one percent of GDP. Row 11 reports a total distance measure for each model. The total distance measure is computed by summing the distance measures for the responses of hours, the nominal wage, the consumer price index, and consumption. For each of these individual measures, distance is computed as $\sum_i \left(Moment_{data}^{(i)} - Moment_{model}^{(i)}\right)^2 / \left[s.e.Moment_{data}^{(i)}\right]^2$. We do not include distance measures for the employment rate and firm entry, as most models do not include these margins of adjustment. We also do not include distance measures for composite metrics (the real worker wage and household labor wedge).

Table 7. Calibration

Parameter	Moment	Model-Implied Moment Value	Parameter Value
ϕ	Share of the Dept of Defense spending in GDP	$\frac{\phi}{1+\phi}$	0.01
$n\psi$	Change in labor earnings w.r.t change in GDP, $\frac{dWH}{dPQ}$	$n\psi$	0.34
r	Housing Expenditure share	$\frac{\mathcal{L}_{\ell t} p_{\ell t}^{\mathcal{L}}}{\mathcal{L}_{\ell t} p_{\ell t}^{\mathcal{L}} + R^P}$	2.7
γ	Labor Share, $\frac{WH}{PQ}$	$\frac{n\psi[R+r]}{R}$	3.4

Note: The table shows the implications of data moments and the empirical estimates for calibrated parameters. In rows 3 and 4, R and R^P are total revenues and private-sector revenues. The model parameters listed above are the following: α is the shape parameter from the firm size distribution. ϕ is the government demand parameter. $n\psi$ is workers' bargaining power. γ is a demand curve parameter. r is the value of non-employment. Parameters r and γ are jointly derived from moments in the third and fourth rows.

Table 8. Assessment of the NMC model

Metric	Estimate from data	Baseline calibration	50% Increase in the value of:				
			ϕ	$n\psi$	r	γ	α
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Local Multiplier $\frac{dR}{dG}$	1.05	1.05	1.05	1.05	1.05	1.05	1.57
Elasticity with respect to GDP							
Earnings, $\frac{d\log WH}{d\log PQ}$	0.93	0.85	0.85	0.85	0.79	0.85	0.35
Elasticity with respect to GDP quantity							
Nominal wage (Wage bill per employee)	0.11	-0.22	-0.22	-0.23	-0.31	-0.23	-0.86
Employment quantity, $\frac{d\log Emp}{d\log Q}$	0.41	0.68	0.68	0.68	0.68	0.68	0.76
Hours, $\frac{d\log Emp}{d\log Q}$	0.57	0.68	0.68	0.68	0.68	0.68	0.76
Establishments, $\frac{d\log J}{d\log Q}$	0.15	0.68	0.68	0.68	0.68	0.68	0.76
Consumer Price, $0.4 * \frac{d\log p^{\mathcal{L}}_{\ell t}}{d\log Q}$	0.26	0.12	0.12	0.18	0.12	0.12	0.02
Labor wedge $\frac{d\log WH-\left(1+\frac{1}{\xi}\right)d\log Emp-0.4*d\log p^{\mathcal{L}}-dlog\ q}{d\log Q}$	-1.82	-2.12	-2.12	-2.33	-2.21	-2.12	-2.25
Consumption of tradable goods $\frac{dlog\ q}{d\log Q}$	1.10	0.30	0.31	0.45	0.30	0.30	0.06
Fit: Distance Measure		18.5	18.4	14.4	22.7	18.9	78.3

Note: We report the elasticity of nominal variables (e.g., earning, land prices) with respect to nominal GDP (PQ), and we report the elasticity of real variables (e.g., employment) with respect to real GDP (Q). In the baseline calibration, $\phi = 0.01$, $n\psi = 0.34$, $r = 0.27$, $\gamma = 3.39$, $\alpha = 1.05$. ξ is calibrated at 1. The computation behind the distance measure is discussed in Section 4.F. The total distance measure is computed by summing the distance measures for the responses of hours, the nominal wage, the consumer price index, and consumption. For each of these individual measures, distance is computed as $\sum_i \left(Moment^{(i)}_{data} - Moment^{(i)}_{model}\right)^2 / [s.e.Moment^{(i)}_{data}]^2$. We also do not include distance measures for the employment rate and firm entry, as most models do not include these margins of adjustment. We also do not include distance measures for composite metrics (the real worker wage and household labor wedge).

Appendix (For Online Publication)

Here we derive expressions from the model's equilibrium as well as the response of macro metrics to increases in government spending.

Mass of Surviving Firms

Surviving firms are those for which

$$\theta_{j\ell t}^2 > \frac{4\gamma\lambda_t \underline{R}}{(1+\phi_{\ell t})}$$

Given our distributional assumption on θ^2, this implies that the mass of surviving firms is

$$J_{\ell t} = \int_{\frac{4\gamma\lambda_t \underline{R}}{(1+\phi_{\ell t})}}^{\infty} \alpha(\theta^2)^{-\alpha-1}\, d\theta^2 = -\infty^{-\alpha} + \left[\frac{4\gamma\lambda_t \underline{R}}{(1+\phi_{\ell t})}\right]^{-\alpha} = \left[\frac{4\gamma\lambda_t \underline{R}}{(1+\phi_{\ell t})}\right]^{-\alpha}$$

$$J_\ell = \left[\frac{4\gamma\lambda_t \underline{R}}{(1+\phi_{\ell t})}\right]^{-\alpha}$$

Revenues

Total local revenues from the private sector are

$$R_{\ell t}^P = \frac{1}{4\gamma\lambda_t}\int_{\frac{4\gamma\lambda_t \underline{R}}{(1+\phi_{\ell t})}}^{\infty} \theta^2 f(\theta^2) d\theta^2 = \frac{1}{4\gamma\lambda_t}\int_{\frac{4\gamma\lambda_t \underline{R}}{(1+\phi_{\ell t})}}^{\infty} \alpha(\theta^2)^{-\alpha} d\theta^2 = \frac{1}{4\gamma\lambda_t}\frac{\alpha}{1-\alpha}(\theta^2)^{1-\alpha}\Big|_{\frac{4\gamma\lambda_t \underline{R}}{(1+\phi_{\ell t})}}^{\infty}$$

$$= \frac{1}{4\gamma\lambda_t}\frac{\alpha}{1-\alpha}\left(-\left[\frac{4\gamma\lambda_t \underline{R}}{(1+\phi_{\ell t})}\right]^{1-\alpha}\right) = \frac{1}{4\gamma\lambda_t}\frac{\alpha}{\alpha-1}\left[\frac{4\gamma\lambda_t \underline{R}}{(1+\phi_{\ell t})}\right]^{1-\alpha}$$

$$= (4\gamma\lambda_t)^{-\alpha}\frac{\alpha}{\alpha-1}\left[\frac{\underline{R}}{(1+\phi_{\ell t})}\right]^{1-\alpha}$$

Revenues from the government in a location are

$$G_{\ell t} = \phi_{\ell t} R_{lt}^P.$$

Total local revenues are the sum of private-sector revenues and revenues from government spending:

$$R_{\ell t} = R_{\ell t}^P + G_{\ell t} = (4\gamma\lambda_t)^{-\alpha}\frac{\alpha}{\alpha-1}\left[\frac{\underline{R}}{(1+\phi_{\ell t})}\right]^{1-\alpha}(1+\phi_{lt}).$$

The government share of GDP in a location is $\frac{G_{\ell t}}{R_{\ell t}} = \frac{\phi_{\ell t}}{1+\phi_{\ell t}}$.

GDP Multiplier

The multiplier is the change in total revenues for every dollar of spending from the government:

$$\frac{dR_{\ell t}}{dG_{\ell t}} = \frac{d\{(1+\phi_{\ell t})R_{\ell t}^{p}\}}{d\{\phi_{\ell t}R_{\ell t}^{P}\}} = \frac{\left((1+\phi)d(R_{\ell t}^{P}) + R_{\ell t}^{P}d\phi\right)}{\phi d(R_{\ell t}^{P}) + R_{\ell t}^{P}d\phi}$$

Where $d(R_{\ell t}^{P}) = d(4\gamma\lambda_t)^{-\alpha}\frac{\alpha}{\alpha-1}\left[\frac{\underline{R}}{(1+\phi_{\ell t})}\right]^{1-\alpha} = (4\gamma\lambda_t)^{-\alpha}\underline{R}^{1-\alpha}\frac{\alpha}{\alpha-1}(\alpha-1)(1+\phi_{\ell t})^{\alpha-2}d\phi$

So

$$\frac{dR_{\ell t}}{dG_{\ell t}} = \frac{(1+\phi)(\alpha-1)(1+\phi_{\ell t})^{\alpha-2} + (1+\phi_{\ell t})^{\alpha-1}}{\phi(\alpha-1)(1+\phi_{\ell t})^{\alpha-2} + (1+\phi_{\ell t})^{\alpha-1}} = \frac{(\alpha-1)+1}{\phi(\alpha-1)(1+\phi_{\ell t})^{-1}+1}$$

$$= 1 + \frac{\alpha-1}{\alpha\phi+1}$$

(Inverse of) Labor Share

In the model, wage income w corresponds to earnings WH in the data, and revenues R correspond to PQ (GDP). Hence, the model analogue of the inverse of labor share is

$$\frac{PQ}{WH} = \frac{R_{\ell t}}{n\psi[R_{\ell t}+r]}.$$

We examine two measures of the response of the labor share to a demand shock: the elasticity of GDP with respect to earnings, and the change in GDP relative to the change in earnings.

We first derive the elasticity of GDP with respect to earnings, $\frac{d\log PQ}{d\log WH}$, driven by a change in local government spending $\phi_{\ell t}$.

$$\frac{d\log PQ}{d\phi_{\ell t}} = \frac{d\log R_{\ell t}}{d\phi_{\ell t}} = \frac{1}{R_{\ell t}}\frac{dR_{\ell t}}{d\phi_{\ell t}}$$

$$\frac{d\log WH}{d\phi_{\ell t}} = \frac{1}{n\psi[R_{\ell t}+r]}\frac{n\psi\, dR_{\ell t}}{d\phi_{\ell t}}$$

$$\frac{d\log PQ}{d\log WH} = \frac{R_{\ell t}+r}{R_{\ell t}}.$$

Next we derive the change in revenues as a ratio of the change in earnings $\frac{dPQ}{dWH}$:

$$\frac{dPQ}{dWH}=\frac{dR_{\ell t}}{d\{n\psi[R_{\ell t}+r]\}}=\frac{dR_{\ell t}}{n\psi\times dR_{\ell t}}=\frac{1}{n\psi}$$

Elasticity of Nontradable Prices with Respect to GDP

The household's first order condition relates expenditure on local nontradables to the local household's budget multiplier $\lambda_{\ell t}$:[24]

$$\mathcal{L}_{\ell t}p_{\ell t}^{\mathcal{L}}=\frac{1}{\lambda_{\ell t}}.$$

To determine how this responds to an increase in government spending, we examine deviations around a steady state in which the local household's expenditure equals its income (e.g., there is balanced trade):

$$\mathcal{L}_{\ell t}p_{\ell t}^{\mathcal{L}}+\int_0^1\int_{\underline{\theta}_{\ell t}^2}^{\infty}p_{jmt}q_{jm\ell t}dj\,dm+T_{\ell t}=n\psi(R_{\ell t}+r)+\Pi_{\ell t}+I_{\ell t}.$$

Totally differentiating this budget constraint with respect to locally-determined variables and dividing through by R (and assuming $\mathcal{L}$ is fixed by locally endowed production factors, $\Pi_{\ell t}$ and I_{lt} are independent of local conditions due to diversification, $T_{\ell t}$ is independent of local DOD spending, and prices p_{jmt} are independent of local conditions due to price setting at the aggregate level), this becomes:

$$\frac{\mathcal{L}_{\ell t}dp_{\ell t}^{\mathcal{L}}}{R_{\ell t}}+\frac{1}{R_{\ell t}}\int_0^1\int_{\underline{\theta}_{\ell t}^2}^{\infty}p_{jmt}dq_{jm\ell t}dj\,dm=n\psi\frac{dR_{\ell t}}{R_{\ell t}}.\text{[25]}\qquad(14)$$

Note that demand for $q_{jm\ell t}$ is given by

$$q_{jm\ell t}^d=\frac{1}{\gamma}\left(\theta_{jm\ell t}-\lambda_{\ell t}p_{jmt}\right)=\frac{1}{\gamma}\left(\theta_{jm\ell}-\frac{1}{\mathcal{L}_{\ell t}p_{\ell t}^{\mathcal{L}}}\frac{\theta_{jmt}}{2\lambda_t}\right)=\frac{\theta_{jmt}}{\gamma}\left(1-\frac{1}{2\lambda_t\mathcal{L}_{\ell t}p_{\ell t}^{\mathcal{L}}}\right).$$

---

[24] $\lambda_{\ell t}$ denotes the budget multiplier for the local household while λ_t is the average multiplier across locations.

[25] Note that the comparative statics at the national level would include changes in taxes. This implies that national land prices do not change in response to national spending.

Totally differentiating this expression yields

$$dq_{jm\ell t} = \frac{\theta_{jmt}}{\gamma 2\lambda_t}\left(\mathcal{L}_{\ell t} p_{\ell t}^{\mathcal{L}}\right)^{-2} d\left(\mathcal{L}_{\ell t} p_{\ell t}^{\mathcal{L}}\right). \tag{15}$$

$$= \frac{\theta_{jmt}}{\gamma 2\mathcal{L}_{\ell t} p_{\ell t}^{\mathcal{L}}} d\left(\mathcal{L}_{\ell t} p_{\ell t}^{\mathcal{L}}\right).$$

$$= \frac{\theta_{jmt}}{2\gamma} d\log\left(\mathcal{L}_{\ell t} p_{\ell t}^{\mathcal{L}}\right).$$

Substituting in for $dq_{jm\ell t}$ in (14) yields

$$\frac{\mathcal{L}_{\ell t} p_{\ell t}^{\mathcal{L}}}{R_{\ell t}} d\log \mathcal{L}_{\ell t} p_{\ell t}^{\mathcal{L}} + \frac{1}{R_{\ell t}}\int_0^1 \int_{\underline{\theta}_{\ell t}^2}^{\infty} \frac{\theta_{jmt}}{2\lambda_t}\frac{\theta_{jmt}}{2\gamma} d\log\left(\mathcal{L}_{\ell t} p_{\ell t}^{\mathcal{L}}\right) dj\, dm = n\psi \frac{dR_{\ell t}}{R_{\ell t}}$$

$$\frac{\mathcal{L}_{\ell t} p_{\ell t}^{\mathcal{L}}}{R_{\ell t}} d\log \mathcal{L}_{\ell t} p_{\ell t}^{\mathcal{L}} + \frac{\lambda_t}{\gamma R_{\ell t}}\int_0^1 \int_{\underline{\theta}_{\ell t}^2}^{\infty} \left(\frac{\theta_{jmt}}{2\lambda_t}\right)^2 d\log\left(\mathcal{L}_{\ell t} p_{\ell t}^{\mathcal{L}}\right) dj\, dm = n\psi \frac{dR_{\ell t}}{R_{\ell t}}$$

$$d\log \mathcal{L}_{\ell t} p_{\ell t}^{\mathcal{L}}\left[\frac{\mathcal{L}_{\ell t} p_{\ell t}^{\mathcal{L}}}{R_{\ell t}} + \frac{\lambda_t}{\gamma R_{\ell t}}\int_0^1 \int_{\underline{\theta}_{\ell t}^2}^{\infty} \left(\frac{\theta_{jmt}}{2\lambda_t}\right)^2 dj\, dm\right] = n\psi \frac{dR_{\ell t}}{R_{\ell t}}$$

Substitute in $\int_0^1 \int \left(\frac{\theta_{jmt}}{2\lambda_t}\right)^2 d\theta^2\, dm = R_{\ell t}^P \frac{4\gamma\lambda}{4\lambda^2} = R^P \frac{\gamma}{\lambda}$

$$d\log \mathcal{L}_{\ell t} p_{\ell t}^{\mathcal{L}}\left[\frac{\mathcal{L}_{\ell t} p_{\ell t}^{\mathcal{L}}}{R_{\ell t}} + \frac{\lambda_t}{\gamma R_{\ell t}} R_{\ell t}^P \frac{\gamma}{\lambda_t}\right] = n\psi \frac{dR_{\ell t}}{R_{\ell t}}$$

$$\frac{d\log \mathcal{L}_{\ell t} p_{\ell t}^{\mathcal{L}}}{d\log R_{\ell t}} = \frac{n\psi}{\frac{\mathcal{L}_{\ell t} p_{\ell t}^{\mathcal{L}}}{R_{\ell t}} + \frac{R_{\ell t}^P}{R_{\ell t}}}$$

$$\frac{d\log \mathcal{L}_{\ell t} p_{\ell t}^{\mathcal{L}}}{d\log R_{\ell t}} = \frac{R_{\ell t} n\psi}{\mathcal{L}_{\ell t} p_{\ell t}^{\mathcal{L}} + R_{\ell t}^P}$$

Because in a symmetric equilibrium $\lambda_{\ell t} = \lambda_t$, it follows that

$$\frac{d\log \mathcal{L}_{\ell t} p_{\ell t}^{\mathcal{L}}}{d\log R_{\ell t}} = \frac{\lambda_t R_{\ell t}}{1 + \lambda_t R_{\ell t}^P} n\psi$$

Elasticity of Consumption with Respect to Output

Appendix equation (15) gives the response of a variety of consumption to a change in spending on "land" $\mathcal{L}$. To turn this into an elasticity, note that $q_{jm\ell t} = \frac{\theta_{jmt}}{2\gamma}$. Then equation (15) can be written as

$$d \log q_{jm\ell t} = d \log\left(\mathcal{L}_{\ell t} p_{\ell t}^{\mathcal{L}}\right).$$

Note that the right-hand-side of this equation is the same for all j_m and therefore consumption bundle of tradable goods increases by $d \log\left(\mathcal{L}_{\ell t} p_{\ell t}^{\mathcal{L}}\right)$. It follows that the response of consumption to $d \log R_{\ell t}$ is given by

$$\frac{d \log q_{jm\ell t}}{d \log R_{\ell t}} = \frac{d \log q_{jm\ell t}}{d \log \mathcal{L}_{\ell t} p_{\ell t}^{\mathcal{L}}} \frac{d \log \mathcal{L}_{\ell t} p_{\ell t}^{\mathcal{L}}}{d \log R_{\ell t}} = 1 \times \frac{\lambda_t R_{\ell t}}{1 + \lambda_t R_{\ell t}^P} n\psi = \frac{\lambda_t R_{\ell t}}{1 + \lambda_t R_{\ell t}^P} n\psi$$

Elasticity of Employment with Respect to Output

Employment in the model is proportional to the number of firms: $Emp_{\ell t} = nJ_{\ell t}$. Total output is the sum of firm-level output in a location, the private-sector component of which is given by equation (6). Total private-sector output is

$$Q_{\ell t}^P = \frac{1}{2\gamma} \int_{\frac{4\gamma\lambda_t \underline{R}}{(1+\phi_{mt})}}^{\infty} \theta f(\theta^2) d\theta^2 = \frac{1}{2\gamma} \int_{\frac{4\gamma\lambda_t \underline{R}}{(1+\phi_{mt})}}^{\infty} \alpha \, (\theta^2)^{-\frac{1}{2}-\alpha} d\theta^2$$

$$= \frac{1}{2\gamma} \frac{\alpha}{.5 - \alpha} (\theta^2)^{\frac{1}{2}-\alpha} \Big|_{\frac{4\gamma\lambda_t \underline{R}}{(1+\phi_{mt})}}^{\infty}$$

$$= \frac{1}{2\gamma} \frac{\alpha}{\alpha - \frac{1}{2}} \left(\frac{4\gamma\lambda_t \underline{R}}{(1 + \phi_{mt})}\right)^{-(\alpha - .5)}.$$

Total output is the sum of Q^P and Q^G:

$$Q = (1 + \phi_{mt}) \frac{1}{2\gamma} \frac{\alpha}{\alpha - \frac{1}{2}} \left(\frac{4\gamma\lambda_t \underline{R}}{(1 + \phi_{mt})}\right)^{-(\alpha - .5)} = \frac{1}{2\gamma} \frac{\alpha}{\alpha - \frac{1}{2}} \left(4\gamma\lambda_t \underline{R}\right)^{-(\alpha - .5)} (1 + \phi_{mt})^{(1+\alpha - .5)}$$

Hence

$$d\log Q = \left(\frac{1}{2} + \alpha\right)\frac{\phi}{1+\phi}d\log\phi$$

Employment is $\frac{n}{N}\left[\frac{4\gamma\lambda_t R}{(1+\phi_{mt})}\right]^{-\alpha}$, so

$$d\log Emp = \alpha\frac{\phi}{1+\phi}d\log\phi$$

Hence,

$$\frac{d\log Emp}{d\log Q} = \frac{\alpha}{.5+\alpha}$$

Since $d\log J = d\log Emp$, it follows that $\frac{d\log J}{d\log Q} = \frac{\alpha}{.5+\alpha}$.

We can also derive

$$d\log R = \alpha\frac{\phi}{1+\phi}d\log\phi$$

Hence

$$\frac{d\log Q}{d\log R} = \frac{0.5+\alpha}{\alpha},$$

which we use when converting elasticities with respect to nominal GDP to elasticities with respect to real GDP.

Household Labor Wedge

We can write the labor wedge in growth rates as:

$$\tau^H = w - p_c - \frac{1}{\xi}h - c$$

In our model, the wage is the same as the wage bill $n\psi(R_{\ell t} + r)$ per employee. Therefore,

$$w = d\log n\psi(R_{\ell t}+r) - d\log Emp = d\log(R_{\ell t}+r) - d\log Emp = \frac{1}{R_{\ell t}+r}dR_{\ell t}$$

$$= \frac{R_{\ell t}}{R_{\ell t}+r}\,d\log R_{\ell t} - d\log Emp.$$

In our model, let the consumption price be

$$\mathrm{dlog}\, p_{\ell t}^{c} = s^{\mathcal{L}}\ \mathrm{dlog}\, p_{\ell t}^{\mathcal{L}} + (1 - s^{\mathcal{L}})\, \mathrm{dlog}\, p_{jmt},$$

where $s^{\mathcal{L}} = 0.4$ is the share of land expenditure in total household spending. In our model, p_{jmt} is invariant to local demand shocks. Recall also that

$$\frac{d \log \mathcal{L}_{\ell t} p_{\ell t}^{\mathcal{L}}}{d \log R_{\ell t}} = \frac{d \log p_{\ell t}^{\mathcal{L}}}{d \log R_{\ell t}} = \frac{\lambda_t R_{\ell t}}{1\ + \lambda_t R_{\ell t}^{P}} n\psi$$

Hence,

$$\mathrm{dlog}\, p_{\ell t}^{c} = s^{\mathcal{L}} \times \mathrm{dlog}\, p_{\ell t}^{\mathcal{L}} = s^{\mathcal{L}} \times \frac{\lambda_t R_{\ell t}}{1\ + \lambda_t R_{\ell t}^{P}} n\psi \times d \log R_{\ell t}$$

Finally, we have the response of consumption of tradable goods:

$$d \log q_{\ell t} = d \log\left(\mathcal{L}_{\ell t} p_{\ell t}^{\mathcal{L}}\right) = \frac{\lambda_t R_{\ell t}}{1\ + \lambda_t R_{\ell t}^{P}} n\psi \times d \log R_{\ell t}$$

It follows that the response of labor wedge is

$$\tau^{H} = (average\ wage) - p_c - \frac{1}{\xi} hours - consumption$$

$$= \left(\frac{R_{\ell t}}{R_{\ell t} + r}\ d \log R_{\ell t} - d \log Emp_{\ell t}\right) - \mathrm{dlog}\, p_{\ell t}^{c} - \frac{1}{\xi} d \log Emp_{\ell t} - d \log q_{\ell t}$$

$$= \frac{R_{\ell t}}{R_{\ell t} + r}\ d \log R_{\ell t} - \mathrm{dlog}\, p_{\ell t}^{c} - \left(1 + \frac{1}{\xi}\right) d \log Emp_{\ell t} - d \log q_{\ell t}$$

$$= \frac{R_{\ell t}}{R_{\ell t} + r}\ d \log R_{\ell t} - s^{\mathcal{L}} \times \frac{\lambda_t R_{\ell t}}{1\ + \lambda_t R_{\ell t}^{P}} n\psi \times d \log R_{\ell t} - \left(1 + \frac{1}{\xi}\right) d \log Emp_{\ell t}$$
$$- \frac{\lambda_t R_{\ell t}}{1\ + \lambda_t R_{\ell t}^{P}} n\psi \times d \log R_{\ell t}$$

$$= \frac{R_{\ell t}}{R_{\ell t} + r}\ d \log R_{\ell t} - (1 + s^{\mathcal{L}}) \times \frac{\lambda_t R_{\ell t}}{1\ + \lambda_t R_{\ell t}^{P}} n\psi \times d \log R_{\ell t} - \left(1 + \frac{1}{\xi}\right) d \log Emp_{\ell t}$$

Substituting in $d \log R_{\ell t}$ for $d \log Emp_{\ell t}$, we have

$$\tau^{H} = \left\{\frac{R_{\ell t}}{R_{\ell t} + r} - (1 + s^{\mathcal{L}}) \times \frac{\lambda_t R_{\ell t}}{1\ + \lambda_t R_{\ell t}^{P}} n\psi - \left(1 + \frac{1}{\xi}\right)\right\}\ d \log R_{\ell t}.$$

Appendix Table 1. Long-horizon fiscal multipliers

	2-year Multiplier		3-year Multiplier	
	Nominal GDP	Real GDP	Nominal GDP	Real GDP
	(1)	(2)	(3)	(4)
$(G_{it} - G_{it-h})/Y_{i,t-h}$	1.116**	1.056***	1.099**	1.006***
	(0.450)	(0.383)	(0.445)	(0.367)
Observations	5,605	5,610	5,605	5,610
R-squared	-0.077	0.004	-0.110	0.005
1^{st} stage F-stat	10.18		10.18	

Notes: This table presents estimates from a modified version of Specification 3 in which changes in government spending are over a horizon of $h \in \{2,3\}$ years, using the Bartik instrument, $(\widehat{G_{it} - G_{it-h}})/Y_{it-h}$ for the scaled change in government spending. Fixed effects for CBSA and year are included but not reported. Standard errors clustered by state are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Appendix Table 2. Long-term (3-year) response to demand-driven changes in output

Row	Outcome variables (3-year-horizon):	Coef. (s.e.)	1^{st} stage F stat
1	$dlog\left(\frac{Earnings}{GDP}\right)$, labor share	-0.096 (0.140)	20.01
2	$dlog\ (H)$, hours	0.642*** (0.100)	13.93
3	$dlog(E)$, employment	0.504*** (0.098)	13.93
4	$dlog\left(\frac{H}{E}\right)$, hours per employee	0.146*** (0.042)	13.93
5	$dlog\left(\frac{E}{L}\right)$, employment rate	0.293*** (0.083)	13.93
6	$dlog(L)$, labor force	0.203*** (0.078)	13.93
7	$dlog\left(\frac{L}{Pop}\right)$, labor force participation rate	-0.027 (0.044)	14.01
8	$dlog(Pop)$, population	0.223*** (0.086)	13.93
9	$dlog(wage)$, wages	0.324* (0.194)	13.93
10	$dlog(resid\ wage)$, residualized wages	0.215* (0.115)	13.93
11	$0.4 \times dlog(resid\ rent\ price)$, CPI	0.288*** (0.115)	13.93
12	$0.25 \times dlog(auto)$, consumption	1.958*** (0.744)	6.81
13	$dlog(establishments)$, firm establishments	0.210** (0.103)	15.36

Notes: This tables presents estimates based on a modified version of specification (2) in which changes in government spending are over a 3-year horizon, using the Bartik instrument, $(\widehat{G_{it} - G_{it-3}})/Y_{it-3}$ for the scaled change in government spending. The regressor is growth in real GDP (instrumented by the Bartik shock). Time and city fixed effects are included but not reported. Standard errors clustered by state are in parentheses. The number of observations for row 1 is 4,488. The number of observations for rows 2-11 is 1,938. The number of observations for row 12 is 2,976. The number of observations for row 13 is 3,732. *** p<0.01, ** p<0.05, * p<0.1.