# Wage Inequality and the Spatial Expansion of Firms*†

Benny Kleinman‡
Stanford University

September 2025

**Abstract**

The U.S. economy is increasingly dominated by multi-location service firms. These firms have undergone substantial geographic expansion in recent decades and account for most of the rise in wage inequality. I develop a theory that links these trends, in which firms open branches to serve local markets and employ headquarters workers whose output is non-rival across locations. An increase in the number of firm branches is akin to a technical change that augments the non-rival headquarters labor. Through this mechanism, the model can rationalize a range of labor market dynamics, including endogenous skill-biased technical change, widening spatial disparities, and rising wage dispersion both across and within firms. I provide supporting evidence for these predictions and for the assumption of within-firm non-rivalry. Quantitatively, a decline in spatial frictions to firm expansion goes a long way toward explaining key labor market trends from recent decades.

*I am grateful to my advisors Gene Grossman, Ezra Oberfield, Stephen Redding, and Richard Rogerson for their continued guidance and support. This paper has also benefited from discussions and comments from Mark Aguiar, Juanma Castro-Vincenzi, Oren Danieli, Xiang Ding, Pablo D. Fajgelbaum, Teresa Fort, Manuel García-Santana, Esteban Rossi-Hansberg, Erik Hurst, Gregor Jarosch, Marta Lachowska, Hugo Lhuillier, Ernest Liu, Isabela Manelici, Simon Mongey, Joan Monras, Eduardo Morales, Pascual Restrepo, Hannah Rubinton, Jose P. Vasquez, Gianluca Violante, Jonathan Vogel, and Conor Walsh, as well as numerous seminar participants. I thank Shahin Davoudpour, Shirley Liu, and Hannah Rubinton for help with the data. This research benefited from financial support from the International Economics Section (IES) and the Simpson Center for the Study of Macroeconomics in Princeton University, the Washington Center for Equitable Growth, and the Becker Friedman Institute for Economics at University of Chicago.

†Any views expressed are those of the authors and not those of the U.S. Census Bureau. The Census Bureau's Disclosure Review Board and Disclosure Avoidance Officers have reviewed this information product for unauthorized disclosure of confidential information and have approved the disclosure avoidance practices applied to this release. This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 2542. (CBDRB-FY22-P2542-R9788, CBDRB-FY22-P2542-R10063, CBDRB-FY24-P2542-R11261)

‡Stanford University, bennyk@stanford.edu.

# 1 Introduction

The U.S. economy is increasingly dominated by multi-location firms that operate establishments across multiple local markets. This trend is evident across a wide range of service sectors, from retail and food services to finance and healthcare. In this paper, I study the labor market implications of this trend and provide a new theory that links changes in firms' geographical scope to the distribution of wages in the economy. The resulting framework is able to rationalize multiple labor market trends from recent decades, providing new micro-foundations for skill-biased technical change, increasing spatial disparities, and increasing inequality across firms and establishments.

I conceptualize multi-location service firms as networks of local branches linked by national headquarters, in which branch workers' output is *non-tradable*, and headquarters workers' output is *non-rival* across the firm's locations.[1] Examples include retail and restaurant chains, banks, healthcare networks, real-estate management companies, and telecommunications firms. In such an environment, an increase in firms' spatial scope is similar to a technical change that augments the non-rival factors of production, yielding rich distributional implications. I investigate this idea in several steps. First, I provide motivational evidence for the increased prevalence of this class of firms and their central role in the rise of inequality. Second, I develop a model of these firms in a spatial general equilibrium setting, and use it to show how shocks that raise firms' ability to expand in space can lead to endogenous skill-biased technical change, to increases in residual wage inequality, and to increases in spatial disparities. Third, I provide reduced-form evidence for the model's key mechanisms and implications. Finally, I quantify the model to address: (a) To what extent can recent labor market trends be explained by declining spatial frictions to firm expansion? (b) What are the implications of policies that shape firms' expansion opportunities, such as deregulation of cross-state activity?

I begin by documenting key motivating facts about multi-establishment service firms in the U.S. economy since the 1980s. First, these firms have undergone substantial spatial expansion. Second, they account for most of the increase in wage variance across the universe of U.S. establishments. Third, their sizable role in rising inequality reflects both greater variance across multi-establishment service firms and greater variance across establishments within these firms, underscoring the importance of their internal structure for this trend.[2]

I develop a general equilibrium model with this form of organization at its core. The

[1] Throughout the paper, I use a broad notion of headquarters, encompassing all workers whose output is non-rival across space. I later show that such workers are indeed highly spatially concentrated within firms, and that actual headquarters locations provide a good proxy for this concept.

[2] Moreover, the increase in variance across establishments within multi-establishment firms is substantially larger than that across stand-alone establishments. Thus, frameworks that focus solely on inequality across firms or solely on inequality across establishments are ill-suited to capture these patterns, making a model of the internal structure of multi-establishment firms essential.

model has four main components. First, firms open branches and hire local workers to provide a non-tradable product across local labor markets (e.g., a bank branch, a restaurant, a maintenance facility, a warehouse). Second, firms hire headquarters workers to raise the productivity (or quality) of their branches. Importantly, the output of headquarters workers is *non-rival* across the firm's locations, reflecting their role in developing and designing products and processes that can be deployed across all branches.[3] Third, firms choose how many locations to serve, subject to a cost of spatial expansion that rises with the number and distance of branches. This cost also serves as a key exogenous determinant of the extent of spatial expansion in the economy. Lastly, I allow firms to set wages, facing upward-sloping, firm-specific labor supply curves in each location. This feature generates additional predictions for inequality both across and within firms, and enables direct tests of the model's mechanisms using firm-level data.[4] The allocation of workers, headquarters, and firm branches across space is determined in equilibrium.

I emphasize three main takeaways from the model. The first is a new microfoundation for skill-biased technical change (SBTC). As in standard models in this literature following Katz and Murphy (1992) and Krusell et al. (2000), firms' relative expenditure on skilled (headquarters) workers depends on relative factor prices and on exogenous biased technical change. In the present setting, however, firms' expenditure on skilled workers also varies with their chosen geographical scope. Expanding geographical scope operates like a technical change that augments factors of production used intensively in the firm's non-rival component. Intuitively, because of the non-rivalry assumption, the marginal product of headquarters workers scales with the number of locations in which their output can be deployed, while the marginal product of branch workers does not. This effect arises even in simple versions of the model with symmetric space and perfectly competitive labor markets, relying only on the assumption that the output of a subset of workers is non-rival across the firm's locations.

Second, allowing for firm heterogeneity and wage-setting power, the same mechanism generates changes in inequality both across and within firms. As in standard models with firm wage-setting power (e.g., Card et al. (2018), Berger et al. (2022)), larger firms pay higher wages on average. Unlike those models, however, larger firms here also exhibit greater within-firm inequality, particularly between headquarters and branches. This result follows from the interaction between non-rival headquarters output and firm-specific labor supply curves. Headquarters wages scale with the firm's total output across all branches, while branch wages scale only with local output. Thus, as firms expand geo-

[3]Consider, for example, the output of a firm's designers, software engineers, marketing specialists, and researchers. These are among the top headquarters occupations and perform tasks that can be classified as non-rival, as I demonstrate later.

[4]I model firm-specific wages following the recent monopsony literature in labor economics, given its tractability and growing empirical support. That said, the model's main implications also arise under alternative microfoundations for wage heterogeneity across firms and establishments.

graphically, within-firm inequality rises. Moreover, if it is disproportionately the already large firms that expand – as in the data and in my quantification – then between-firm inequality rises in tandem with within-firm inequality. Notably, firm expansion does not necessarily raise branch wages: while higher productivity increases the firm's labor demand across locations, it also makes local production more dependent on the firm's non-rival component. As a result, the wage–size gradient is steep for headquarters workers but can be flat for branch workers, and firm-level shocks pass through more strongly to headquarters wages than to branch wages.

Third, allowing for spatial heterogeneity, the same mechanism generates changes in inequality across local labor markets. As firms' ability to expand geographically increases, markets with a comparative advantage in headquarters activity experience a relative rise in demand for skilled labor. First, their locally headquartered firms become relatively larger, resulting in what appears as stronger skill-biased technical change in these markets. Second, these markets increasingly specialize in exporting headquarters services to other regions, further boosting demand for skilled labor.[5] These predictions align with the rising spatial disparities across U.S. local labor markets in recent decades, and with the transformation of key cities into service-providing centers and away from production of tradable goods.

I provide reduced-form empirical evidence for the model's main mechanisms. To this end, I use data on the near-universe of firms' job postings from Lightcast, matched to information on firms' spatial structure and headquarters locations from Dun & Bradstreet. This combination allows me to analyze how firms' job composition and pay evolve as they expand geographically, distinguishing between headquarters and branch jobs. A key advantage of these data relative to the Census is the ability to measure headquarters activity through occupational classifications. As a complementary analysis, I also use establishments' industry classifications to proxy headquarters activity in the Census data.

I begin by providing evidence for the non-rival nature of headquarters jobs. While non-rivalry is inherently unobservable, I test several predictions implied by this assumption. First, using O*NET task-level data, I show that headquarters jobs are characterized by highly non-rival task content. Second, I find that, consistent with the implications of non-rivalry, relative compensation for headquarters jobs rises as firms expand geographically. This prediction holds in particular when expansion is driven by demand shocks (rather than changes in firm technology) constructed from differential firm exposure to regional growth. Third, I confirm another key implication of non-rivalary: a positive demand shock in a subset of a firm's locations spills over to higher employment and wages in its other unaffected locations.

[5] A striking example is Bentonville, Arkansas, home to Walmart's headquarters (see Appendix A for more details). In the early 1970s, its labor market resembled the rest of Arkansas, but as Walmart expanded, it transformed into a high-skill service hub, with average income double the rest of the state.

I provide additional evidence for the model's predictions along other dimensions. Across local labor markets, those with greater expansion of locally headquartered firms experienced stronger reallocation toward headquarters-intensive occupations, greater shifts away from manufacturing, and faster wage and employment growth. Across industries and occupations, headquarters-intensive jobs saw larger increases in their economy-wide employment share and relative wages, particularly in industries that underwent greater spatial expansion.

In the final part of the paper, I use a quantitative version of the model to assess the aggregate relevance of the new mechanisms. I estimate the model for 200 local labor markets in the contiguous U.S. in 1980. Frictions to firm spatial expansion are estimated from the universe of headquarters–branch linkages in the data. To identify the main production parameters and the cost of expansion, I employ a Simulated Method of Moments (SMM) approach, targeting key moments on firm structure and inequality from the empirical analysis, and inverting region-level fundamentals from the model's equilibrium conditions.

The main thought experiment I conduct is a realistic reduction in the cost of opening distant branches. I infer this shock from changes in the network of headquarters-branch linkages over time. Between 1980 and 2017, these linkages became much less sensitive to distance, consistent with the idea that improvements in communication and transportation infrastructure lowered the barriers to operating far-away branches. Accordingly, I use only information on changes in the relative number of far-away versus nearby branches, not their absolute levels. In the model, I implement this by reducing the same distance elasticity while holding all other primitives constant. This experiment has several advantages. First, all labor-market and welfare implications operate purely through the novel mechanisms of the paper, since the shock affects only the cost of spatial expansion. Second, although many primitives in the model could generate greater geographic scope, this shock is tightly disciplined by data on the relative prevalence of far-away branches within firms. Lastly, it has a natural interpretation as a reduction in spatial frictions to operating establishments across locations.

Quantitatively, I find that a decline in the cost of spatial expansion generates labor market implications that mirror much of the U.S. experience since the 1980s. First, wage inequality (as measured by the variance of log wages) rises by roughly one-third of the within-industry increase observed in the data. Second, as in the data, inequality grows both within firms (across establishments) and between firms. Third, a central driver of the rise in within-firm inequality is greater wage dispersion between headquarters and branch workers. Fourth, spatial inequality and skill segregation increase, accounting for about one-third of the observed rise in spatial disparities. Fifth, wage gains are concentrated in markets that were already specialized in providing headquarters services – typically dense urban markets – thereby raising the urban wage premium. At the same time, the shock

delivers welfare gains of about 4%, reflecting both variety gains from broader firm scope and productivity gains from the non-rivalry in production.

As a final exercise, I use the model to evaluate the aggregate and distributional consequences of policies that affect firms' ability to expand geographically. A prominent example from recent decades is the deregulation of cross-state firm activity during the 1980s and 1990s in key service sectors, notably financial services, transportation, and utilities. This deregulation was likely an important driver of the rise in firms' geographical scope over the period. I replicate this experiment in the model by reducing the role of state border effects, measured using a gravity equation for headquarters – branch linkages in the data. The exercise yields average welfare gains of about 0.8%, accompanied by a modest increase in firm scope and wage inequality.

This paper connects to several strands of related literature. First, it relates to recent work documenting the spatial expansion of firms, including Cao et al. (2017), Jiang (2021), Rossi-Hansberg et al. (2021), Aghion et al. (2023), and Hsieh and Rossi-Hansberg (2023). Building on this literature, I focus on the labor market implications of firm expansion. I argue that changes in firms' geographical scope can have non-trivial distributional consequences when combined with within-firm non-rivalries, particularly in an environment with firm wage-setting.

Second, I relate to multiple strands of the vast literature on rising wage inequality. Relative to the literature on skill-biased technical change (henceforth SBTC) – following Katz and Murphy (1992) and Krusell et al. (2000) – I contribute a new micro-foundation for endogenous SBTC that is based on the combination of within-firm non-rivalries and changes in firms' scope. Relative to the literature on "the great divergence" in spatial economics – e.g. Berry and Glaeser (2005), Moretti (2012), Diamond (2016), Giannone (2017), Eckert (2019), Rubinton (2020), Eckert et al. (2022) and Card et al. (2025) – I contribute a new mechanism that relies on spatial firm expansion and connects to other trends in the broad inequality literature. I also provide theory and evidence on the importance of cross-region firm linkages in determining the spatial allocation of economic activity. Relative to the empirical literature that documents rising inequality across firms in the U.S. – Barth et al. (2016), Song et al. (2019) and Haltiwanger et al. (2022) – I provide new evidence on the central role of multi-establishment service firms, emphasizing a new important margin of inequality across different establishments within these firms, and provide a theory that rationalizes these patterns. Finally, relative to papers that study wage inequality in an environment with monopsonistic competition – as in Card et al. (2018) and Berger et al. (2022) – I provide an environment in which firm size matters differently for skilled and unskilled workers and which allows for heterogeneous pass-through of firm-level shocks to different workers.

From a theoretical perspective, I relate to studies that model the expansion of firms through space, including Jia (2008) and Holmes (2011) in the industrial organization

literature; Oberfield et al. (2024), Giroud et al. (2024), and Argente et al. (2025) in the macroeconomics literature; and Helpman (1984), Ramondo and Rodríguez-Clare (2013), Tintelnot (2017), and Arkolakis et al. (2018) in the literature on trade and multinational firms. The key deviation from this literature is to model the firm's non-rival component as a variable labor cost (and not just as fixed costs). This has important distributional implications when firms change their geographical scope, even when space is symmetric. I also deviate from most of this literature by allowing for imperfectly-competitive labor markets, and highlight important interactions of this margin with firms' geographical scope. Finally, I integrate the expansion of these firms into a spatial general equilibrium in which the allocation of headquarters, branches, and workers are all determined in equilibrium. I also relate to the theoretical literature that connects within-firm wage inequality to firms' hierarchical organization, e.g. Garicano and Rossi-Hansberg (2004) and Garicano and Rossi-Hansberg (2006). My notion of headquarters is not a hierarchical layer, but rather it captures tasks for which the output is non-rival across space, which allows me to encompass broader patterns of inequality; and it provides a natural link between firms' organization to the literature on rising spatial disparities.

I also relate to the large literature that documents and models the flows of intangible knowledge and know-how within firms, including Atalay et al. (2014), Fort (2017), Alviarez et al. (2023), Ding et al. (2022), and Ding (2023). I explicitly model this knowledge as the output of firms' (headquarters) workers, and highlight the distributional implications that arise when there is a change in firms' ability to spread knowledge across space due to spatial expansion. I also use the structure of the model to overcome the challenge of measuring within-firm knowledge flows and recover frictions to within-firm communication using the empirical network of HQ-branch linkages.

The remainder of the paper is structured as follows. I start with some simple motivational facts in Section 2, highlighting the spatial expansion of firms in recent decades and the centrality of multi-location service firms in the growth of wage inequality. In Section 3, I lay out a model of these firms in spatial general equilibrium, and characterize the link between changes in firms' scope to skill-biased technical change, firm specific wages, and spatial disparities. In Section 4, I provide empirical evidence for the key predictions of the model and for the presence of within-firm non-rivalries. I estimate the model in Section 5 and use the quantitative model to analyze the aggregate implications of lower frictions to firm expansion in Section 6. Section 7 concludes.

# 2 Data and motivating facts

## 2.1 Data

I start by briefly describing the data used in the paper, both for the motivating facts in this Section, and for the tests of the model's mechanisms in Section 4. More details on each dataset and on the construction of key variables are given in Appendix D.

**U.S. Census Bureau establishment-level data.** I use establishment-level data from the U.S. Census Bureau's Longitudinal Business Database (LBD) to study historical firm expansion patterns and their relationship to wage dispersion across and within firms. I focus on firms in the services sector,[6] which account for most of the economy.[7] A key limitation of these data is the absence of clear headquarters classifications and occupation-level details. I therefore supplement the analysis with additional commercial firm-level datasets that have such information. Nonetheless, I also exploit variation in industry classifications across establishments within firms as an alternative approach to measuring headquarters versus branch activity (see details below).

**Lightcast job-posting data.** To analyze the distinction between headquarters and branch-level jobs, and how wages in these jobs evolve with firm expansion, I use online job-posting data from Lightcast.[8] These data cover the near-universe of U.S. job postings since 2010. They provide information on employer, job location, occupation, and education requirements, and report salary information for roughly one-fifth of postings. In Section 4, I exploit the detailed firm-by-occupation-by-location structure of the data to test key model implications, in particular how the wage gap between headquarters and branch jobs changes within firms as they expand geographically. I also use the occupational detail to connect headquarters activity to alternative measures of non-rivalry derived from task-level sources. I merge the Lightcast data with headquarters location information from Dun & Bradstreet, as detailed below.

**Dun & Bradstreet establishment-level data.** In addition to the Census data, I use commercial firm- and establishment-level data from Dun & Bradstreet, which cover the near-universe of U.S. firms since 1969. These data provide information on firm structure, establishment locations, headquarters identifiers, industry classifications, and other key attributes. A firm is defined as the set of establishments sharing the same headquarters D-U-N-S number. I primarily use these data to obtain the geographic locations of firms' headquarters and branches – most notably to identify headquarters locations in the Lightcast data and to compute changes in headquarters-branch distance elasticities

[6] I classify a firm as belonging to the services sector if over half of its payroll across all establishments accrues in services-sector establishments. An establishment is classified as services-sector if its 2-digit NAICS code falls under the Bureau of Economic Analysis (BEA) classification of services-producing sectors (i.e., 1-digit NAICS code of 4, 5, 6, 7, or 8).

[7] As shown later, goods-producing firms exhibit very different patterns.

[8] See https://lightcast.io/.

over time. A key advantage relative to the Census data is the consistent availability of headquarters identifiers for all firms. The geographical data I use have been shown to be of high quality, whereas the employment data are less reliable than the Census and therefore I try to avoid it.[9]

**Supplemental publicly available datasets.** I obtain additional county, industry, and aggregate data from standard publicly-available datasets, including the U.S. Census Bureau's Decennial Census and American Community Survey (ACS), the Business Dynamics Statistics (BDS), the Bureau of Labor Statistics' Quarterly Census of Employment and Wages (QCEW), and the Bureau of Economic Analysis (BEA) Regional Accounts. I also use occupational task-level data from O*NET – the U.S. Department of Labor's Occupational Information Network.

**Measuring headquarters and branch activity.** Some of the model's distinctive predictions are for wage and payroll dispersion between headquarters and branch activities within firms. In the model, headquarters are defined as the set of jobs with non-rival output across a firm's locations, and not necessarily the administrative center. Therefore, I view the preferred mapping of the model to the data as identifying occupations within each firm whose output is non-rival across branches, and I use several approaches to capture this. I now briefly describe the data I use to this end, and return to this distinction in more detail in Section 4 after I introduce the model.

My main measure of headquarters vs. branch activity is the propensity of occupations to be hired in firms' headquarters locations. To construct this measure, I merge data from Lightcast and Dun & Bradstreet and compute, for each occupation in each firm, the probability that jobs in that occupation are posted in the firm's headquarters location.[10] I then classify the most headquarters-intensive occupations as headquarters jobs, and the least headquarters-intensive occupations as branch-level jobs.[11] Second, using the same data, I measure the within-firm spatial concentration of each occupation, capturing the idea that non-rival output can be produced anywhere in the firm, whereas branch activity must be performed near consumers. Third, using task-level data from O*NET and a large language model (LLM), I compute the non-rival task content of each occupation as the share of tasks classified as non-rival, in the sense that their output can be applied across branches without diminishing value or effectiveness. I provide further details on these measures in Section 4 after introducing the model's notion of headquarters, and show that all three measures of occupational headquarters intensity are highly correlated.

[9] See Barnatchez et al. (2017) for more details.

[10] To ensure this measure does not reflect compositional differences between sectors, locations, and firms, I take the occupational fixed effect from a regression of headquarters probability on occupation, firm, and location fixed effects.

[11] For each firm, I classify headquarters jobs as the top 10% of employment-weighted occupations by headquarters intensity, and the bottom 90% as branch-level jobs. I set the 10% cutoff because it approximates the employment share of headquarters workers in my final calibration (discussed below), though results are similar with alternative cutoffs.

An alternative way to measure headquarters activity is to use the industry classifications of establishments within a firm. This approach has two important drawbacks. First, the connection to my theoretical notion – jobs whose output is non-rival across branches – is less direct. Second, industry classifications are prone to substantial reporting error, particularly in services, where many firms assign the same industry code to all establishments. The main advantage of this approach, however, is that it enables me to test some of the model's implications in the Census firm-level data, which lack occupational information. In practice, I map headquarters activity to establishments classified as business services (NAICS 51–55), and branch activity to establishments in the firm's primary subsector, defined as the NAICS subsector with the largest payroll share. These two classifications together capture the majority of payroll for most firms.[12,13] Given these limitations, in the LBD exercises that require a definition of headquarters activity, I restrict attention to firms that display a clear differentiation across establishment classifications.

I now turn to document key motivational facts on multi-location service firms in the U.S. economy since the 1980s. I demonstrate the substantial expansion of these firms and their key role in the increase of U.S. wage inequality.

## 2.2 The spatial expansion of service firms

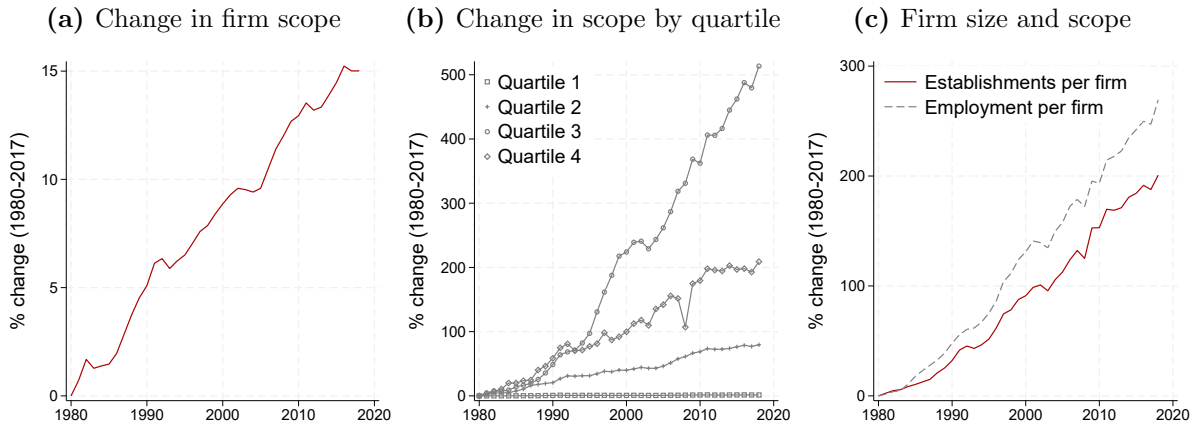
First, I revisit key facts on the expansion of service firms in recent decades.[14] Figure 1 depicts the increase in firms' scope between 1980-2017, as measured by the number of establishments per firm. Subplot (a) presents a simple unweighted average of this metric. Evidently, the average firm operates today around 15% more establishments than in 1980. Note, however, that when we are interested in aggregate labor market outcomes, the employment-weighted expansion might be a more relevant metric, as later demonstrated in the model. Accordingly, subplot (b) divides all firms into four *employment-weighted* bins for each year, and computes the average number of establishments per firm in each bin. This exercise reveals a much more substantial expansion pattern. The majority of U.S. service workers are now employed by firms that are, on average, more than 100%

[12]This definition restricts attention to service firms whose primary activity is not business services (i.e., firms with the majority of payroll in NAICS 42, 44, 48, 49, 61, 62, 71, 72, or 81). While the model also applies to firms specializing in business services (e.g., telecommunications, banking), I cannot cleanly distinguish headquarters from branch activity in such firms.

[13]Another option is to identify headquarters activity using only NAICS 55, which is explicitly defined as headquarters services. I find this problematic for several reasons. Many firms do not report any NAICS 55 activity despite clearly having a headquarters in other sources; conditional on reporting NAICS 55 activity, the payroll share is unusually high. This issue is especially pronounced in business services sectors (e.g., finance, communications), further complicating the headquarters–branch distinction. In addition, the definition is narrower than in my model, which refers to non-rival firm-level tasks that may also be classified under other NAICS codes. That said, my results are robust to using this narrower definition.

[14]Related patterns have been documented also in Cao et al. (2017) and Hsieh and Rossi-Hansberg (2023).

**Figure 1: The increase in the number of establishments for service firms (1980-2017)**




Note: Subplot (a) shows the percentage change relative to 1980 in the (unweighted) average number of establishments per firm in the U.S. Census BDS data for service sectors. Subplot (b) shows the same metric when dividing the universe of firms in each year to four employment-weighted quartiles. Subplot (c) shows changes in the average across these four quartiles for the number of establishments per firm (solid red line) and employment per firm (dashed gray line).

larger than in 1980. Furthermore, about a quarter of workers are employed by firms that are, on average, 500% larger than in 1980. Averaging across these four quartiles, the solid red line of subplot (c) demonstrates that the typical worker is now employed by firms with around three times more establishments than in 1980. Finally, subplot (c) contrasts the change in the average number of establishments per firm with changes in firm size, as measured by total employment per firm (dashed gray line). Evidently, the increase in firms' scope accounts for most of the increase in the size of service firms over this period; firms are now larger because they operate more establishments. More evidence for this point is provided in Appendix E.1.

Some additional features of this trend are worth noting. First, it is widespread across the different service sectors. Second, over half of the increase in establishments per firm is due to expansion across local labor markets (across U.S. commuting zones), though there is non-trivial expansion also within local markets. Third, unlike services, average firm size and scope in the goods sector have both declined.[15] Further details on all these points are provided in Appendix E.

Finally, this trend is associated with a substantial reallocation of workers towards multi-establishment service firms. This class of firms has accounted for 29% of the total U.S. workforce in 1980. By 2017, this share increased to 48%. This reallocation reflects both the general transition from goods to services ("structural transformation") and the reallocation from single-establishment to multi-establishment firms within services.

[15] Holmes and Stevens (2014) provide one potential explanation for this trend, based on the idea that rising international trade has shifted the production of large-scale standardized goods to other countries. In this paper I do not analyze the decline in the size and scope of manufacturing firms, and focus mostly on the expansion of firms in services.

## 2.3 Multi-location service firms and wage inequality

I now show that the same class of firms also accounts for most of the increase in U.S. wage inequality. To this end, I present a simple decomposition of the rise in the variance of log wages across the universe of U.S. establishments. This increase in wage dispersion across establishments has been shown to account for the vast majority of the increase in wage inequality across U.S. workers (see e.g. Barth et al. (2016)), with minimal role for changes in within-establishment variance. I also confirm these findings in Appendix E.2.[16]

I investigate which firms drive the increase in the variance of log wages across establishments in the LBD, starting with a visual representation of the importance of multi-establishment firms. Figure 2 shows the change in within-industry variance of log wages across all establishments (solid black line), and separately for establishments in multi-establishment firms (dashed red line) and single-establishment firms (dotted gray line). The variance rises sharply for multi-establishment firms, while for single-establishment firms there is only a mild increase in the 1980s and a decline thereafter. I focus on within-industry inequality since this is the main moment of interest in the quantitative model, but the same pattern holds without removing industry fixed effects, as shown in Appendix E.2.

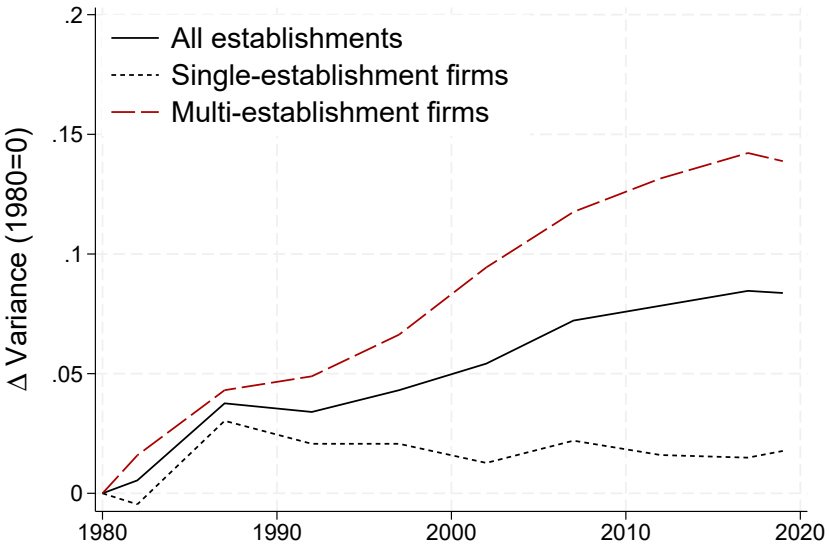
I formally quantify the contribution of multi-establishment firms to the rise in inequality using a variance decomposition of log wages. To this end, consider a partition of the universe of establishments into $G$ distinct groups. The total change in variance equals to the sum of several components:

$$\underbrace{\Delta\sigma_t^2}_{\text{Change in variance}} = \underbrace{s_{g',0}\left(\Delta\sigma_{g',t}^2\right)}_{\text{Variance within group } g'} + \underbrace{\sum_{g\in G}\left(\Delta s_{g,t}\right)\sigma_{g,0}^2}_{\text{Emp. reallocation across groups}}$$
$$+ \underbrace{\sum_{g\in G}\left(\Delta s_{g,t}\right)\left(\Delta\sigma_{g,t}^2\right)}_{\text{Comovement of variance and emp.}} + \underbrace{\sum_{g\in G/g'} s_{g,0}\left(\Delta\sigma_{g,t}^2\right)}_{\text{Variance within other groups}} + \underbrace{\sum_{g\in G}\Delta s_{g,t}\left(\mu_{g,t}^2-\mu_t^2\right)}_{\text{Variance between groups}} \quad (1)$$
$$\text{Residual}$$

where $s_{g,t}$ is the employment-share of group $g$ in period $t$; $\mu_{g,t}$ is the employment-weighted mean of log-wages in group $g$, period $t$; $\sigma_{g,t}^2$ is the employment-weighted variance in group $g$, period $t$; $\mu_t$ is the aggregate mean; and 0 and 1 mark the initial and the final period. In this equation, the first term captures the rise in variance for a particular group $g' \in G$, e.g. establishments belonging to multi-establishment firms, multiplied by its initial share

[16] As detailed below, it is important to distinguish between wage dispersion across firms and wage dispersion across establishments. In particular, most of the increase in inequality within firms is also across different establishments. Therefore, the fact that almost all of the increase in inequality is across establishments does not contradict the sizable role that has been found for rising wage dispersion within firms, e.g. in Song et al. (2019).

**Figure 2: The role of multi-establishment firms in the rise of wage inequality**




Note: this figure shows changes in the employment-weighted variance of log average payroll across establishments in the Longitudinal Business Dataset (LBD) in selected years relative to 1980, after demeaning industry fixed effects (4 digits NAICS code) from establishment log wages. The change in variance for the universe of all establishments is given by the solid-black line. The dashed-red line shows this change for establishments in multi-establishment firms, and the dotted-gray line for single-establishment firms.

of total employment. The second term captures reallocation of employment across groups, keeping each group's variance constant at its base level. The third term captures cross-changes: it adds a positive value to total variance when a group with rising variance also sees an increase in its employment share. Finally, I include all other terms in a residual, comprising the rising variance within all other groups $g \in G/g'$ and rising variance between groups.

Table 1 reports the decomposition of the rise in within-industry wage inequality in the U.S. economy over 1980–2017. Column (1) isolates multi-establishment firms as the group $g'$, showing that over 80% of the overall increase is due to rising variance across establishments belonging to this groups – consistent with the pattern in Figure 2. Column (2) focuses on service-sector firms and shows that 70% of the increase is due to rising variance within this group, with an additional 16% due to reallocation toward services. Finally, Column (3) narrows the focus to multi-establishment service firms. In this case, 45% of the overall increase in inequality is due to rising inequality across establishments that belong to this group, and 22% is due to the cross effect arising from the fact that this group became both more unequal and a much larger part of the economy. The first term is smaller relative to Column (1) since multi-unit service firms accounted for only 29% of total employment in 1980. Still, even holding this share constant, the rising variance in this group accounts for 45% of the overall change. Together with the reallocation of employment to this group, it accounts for 67% of the total rise in inequality.

Table 1 shows the results from this decomposition for the overall rise in within-industry wage inequality in the U.S. economy over 1980-2017. Column (1) considers this decomposition when singling out multi-establishment firms as the group $g'$. Consistent with the pattern in Figure 2, Over 80% of the rise in overall (within-industry) wage inequality is due to rising variance across establishments that belong to multi-establishment firms. I now employ the same decomposition to highlight two additional points about the role of multi-establishment firms in the rise of inequality. First, I show that most of the increase in inequality is driven by firms in services-producing sectors. To see this, Column (2) repeats the decomposition in 1 when singling out the group of service-sector firms. In this case, 70% of the increase is accounted by rising variance for services firms, and 16% due to reallocation to services. Second, I highlight the particular importance of multi-establishment service firms. To see this, Column (3) repeats the decomposition in 1 when narrowing down the focus group to only multi-establishment service firms. In this case, 45% of the overall increase in inequality is due to rising inequality across establishments that belong to this group, and 22% is due to the cross effect arising from the fact that this group became both more unequal and a much larger part of the economy. The first term is smaller relative to Column (1) since as mentioned above, multi-unit service firms accounted for only 29% of total employment in 1980. Still, even holding this share constant, the rising variance in this group accounts for 45% of the overall change. Together with the reallocation of employment to this group, it accounts for 67% of the total rise in inequality.

**The role of wage dispersion within firms.** Having established the importance of multi-establishment service firms, a natural question is whether their role stems from rising dispersion across firms or across establishments within them. The answer is both. To see this, I decompose the increase in variance for multi-unit service firms, $\Delta\sigma^2_{\text{MU-serv},t}$, into within- and between-firm components:

$$\Delta\sigma^2_{\text{MU-serv},t} = \underbrace{\Delta_t \mathbb{V}\left[\ln \bar{w}_f\right]}_{\text{Between-firms - 55\%}} + \underbrace{\Delta_t \mathbb{V}\left[\ln w_{fj} - \ln \bar{w}_f\right]}_{\text{Within-firms, between establishments - 45\%}}$$

where $w_{fj}$ is the average wage in establishment $j$ of firm $f$, and $\ln \bar{w}_f$ is the firm's employment-weighted average log wage. Differences across establishments within firms explain about 45% of the overall increase $\Delta\sigma^2_{\text{MU-serv},t}$.

Although the model below speaks to both components, establishing this fact is important for two reasons. First, it shows that single-unit firm models miss a key margin of rising inequality, highlighting the need to model firms' internal structure. Second, it distinguishes this paper's mechanism from other firm-based theories. For example, the rise of domestic outsourcing should lead to higher wage dispersion across firms, but lower wage dispersion *within* firms, as firms increasingly focus on their core activities. The theory presented in this paper is well suited to address the simultaneous increase in both

**Table 1: Decomposition of the rise in wage inequality**

| | (1) | (2) | (3) |
|---|---|---|---|
| **Group of firms:** | **Multi-unit firms** | **Service firms** | **Multi-unit service firms** |
| **Share out of the overall rise in variance of log wages:** | | | |
| Rising variance within the group of firms | 83% | 70% | 45% |
| Changes in employment shares b/w groups (reallocation) | -4% | 6% | 0% |
| Comovement of variance and employment shares | 9% | 10% | 22% |
| Residual | 12% | 13% | 33% |
| **Total change across all firms in the economy** | **100%** | **100%** | **100%** |

Note: Data from the Census Bureau Longitudinal Business Database. Each column is a separate decomposition of the total increase in within-industry wage inequality across all establishments between 1980 and 2017. The first row shows the share of total increase in variance due to rising variance in the group of firms that is mentioned at the top of each column, matching the first RHS term in Equation (1). The second row shows the share due to changes in employment between that group and the other firms in the sample (employment reallocation), holding constant the change in variance in each group, matching the second RHS term in Equation (1). The third row shows the share that is due to the cross-product of rising variance and rising employment share, matching the third RHS term in Equation (1), and the fourth row is a residual so that the sum for each column is 100%. See Section 2.3 for additional details.

margins.

To recap, I have shown that: (a) multi-establishment service firms have experienced substantial spatial expansion between 1980-2017; (b) they account for most of the increase in wage dispersion in the economy over this period; (c) a big part of their role in due to rising dispersion across different establishments within these firms. Taken together, these evidence provide the motivation to model multi-location firms and their expansion in the analysis of growing inequality.[17]

# 3 Model

In this section I develop a model to formalize the distributional consequences of firm expansion in the presence of within-firm non-rivalries. I start by setting up the firm's problem in partial equilibrium to highlight the novel components of the model. I then add the households side of the economy – which takes a standard form – and complete setting up the spatial general equilibrium.

[17]Note that in contrast to other decompositions in the literature on wage inequality, the goal of the above decomposition is not to separate between workers' characteristics and residual inequality. In fact, the model below will accommodate both. Rather, it is to demonstrate *which class of firms* accounts for most of the increase in wage dispersion in the economy, and to motivate my focus on the expansion of multi-location firms and on their internal structure.

## 3.1 General setting

The economy consists of $N$ heterogenous regions (e.g., commuting-zones), indexed by $i$ and $j$. Each region consists of a unit-mass of identical locations. There are two sectors: services and tradable-goods. In the services sector, households consume a continuum of varieties in a setting of monopolistic competition. These varieties are partially non-tradable in the sense that part of the value added is generated near the consumer, by hiring workers in local branches; while another part is tradable and generated by hiring workers at a potentially distant headquarters.

## 3.2 The problem a service firm

Firms in the service sector differ by their headquarters location $i$ and baseline productivity $z$. I use subscripts to denote locations, and refer to the firm's productivity in parentheses. I use $t_{ij}(z)$ to denote the value of any variable $t$ for a $i$-headquartered firm with productivity $z$ in its region $j$ branches, and $t_i(z)$ to denote variables at the level of the firm.

**Demand.** Each firm $(i, z)$ faces a downward-sloping isoelastic demand curve in each location with demand elasticity $\sigma$, implying that the firm's sales in region $j$ are given by

$$r_{ij}(z) = \Upsilon_j \left(q_{ij}(z)\right)^{\frac{\sigma-1}{\sigma}}, \tag{2}$$

where $\Upsilon_j$ captures a regional demand shifter, $r_{ij}(z)$ is the firm's sales in each of region $j$'s locations, and $q_{ij}(z)$ is the quantity supplied by the firm.

**Production.** A firm $(i, z)$ can supply $q_{ij}(z)$ units of its output in region $j$ using the production function

$$A_i z \left(\gamma h_i(z)^{\frac{\eta-1}{\eta}} + \ell_{ij}(z)^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta}{\eta-1}}, \tag{3}$$

where $z$ is the productivity of the firm across all its markets; $A_i$ is a productivity term common to all firms with headquarters in location $i$; $\gamma$ is a parameter capturing the relative productivity of headquarters-level workers; and $\eta$ is the elasticity of substitution between the headquarters-level and branch-level workers. Importantly, the firm faces this production function in each location $j$, and the bundle of headquarters workers $h_i(z)$ is common and non-rival across locations.

For example, consider the case of Starbucks branches in San Francisco. In this case, $z$ captures the productivity of Starbucks across all its locations; $j$ stands for San Francisco; $i$ stands for Seattle, the headquarters location of Starbucks; $A_i$ captures the productivity of all Seattle-based firms; $\ell_{ij}(z)$ stands for Starbucks workers in its San Francisco branches; and $h_i(z)$ stands for Starbucks headquarters workers in Seattle, such as Starbucks designers, food-scientists, and programers.

**Labor supply.** In each location, the firm faces a convex labor supply curve, such

that in order to employ $\ell_{ij}(z)$ units of branch-level labor in a region $j$ location, it must pay a wage given by

$$w_{\ell,ij}(z) = W_{\ell j}\left(\ell_{ij}(z)\right)^{\frac{1}{\epsilon}}, \tag{4}$$

where $W_{\ell j}$ is a labor-supply shifter for branch-level workers in market $j$; $\epsilon$ is the firm-level labor supply elasticity; and $\ell_{ij}(z)$ is the amount of labor employed by the firm. When $\epsilon \to \infty$, the firm's payroll in market $j$ is linear in the number of workers ($W_{\ell j}\ell_{ij}(z)$), in which case $W_{\ell j}$ captures the competitive wage level in market $j$ for branch-level ($\ell$)workers. More generally, $W_{\ell j}$ is a local wage index capturing the degree of labor-market competition in market $j$.

The firm also faces a similar labor supply curve of headquarters-level workers in market $i$. To hire $h_i(z)$ of headquarters-level workers, it pays a wage $w_{h,i}(z) = W_{hi}\left(h_i(z)\right)^{\frac{1}{\epsilon}}$, where $W_{hi}$ is the labor-supply shifter for headquarters-level workers in market $i$.

Below, I micro-found the case of a finite $\epsilon$ using a monopsonistic labor market environment in which workers have idiosyncratic preferences for firms, as in Card et al. (2018). However, other micro-foundations for such convex labor supply curves, such as screening of higher ability workers as in Helpman et al. (2010), efficiency wages as in Davis and Harrigan (2011), or fair wages as in Amiti and Davis (2012) would also be compatible with the main results.

**Market penetration.** I allow firms to decide on the degree of market penetration in each region $j$, in the spirit of Arkolakis (2010) and Oberfield et al. (2024). Specifically, firms choose to serve a random subset $x_{ij}(z) \in [0,1]$ of the identical locations in each region $j$. Firms do not serve all locations since expansion is costly, with the cost of expansion given by a generic function $C\left(\{x_{ij}(z)\}_{j=1}^{N}\right)$, which depends on the vector of the firm's market penetration decisions in all regions. The firm's total sales in region $j$ is thus given by $x_{ij}(z)r_{ij}(z)$, and its wage bill in region $j$ is given by $x_{ij}(z)w_{\ell,ij}(z)\ell_{ij}(z)$.

A special case of this formulation that follows much of the literature on trade with heterogenous firms is when firms pay a fixed cost to enter each region. In this case, $C(\cdot)$ is the sum of these fixed costs, and firms either fully serve a region or do not serve it at all ($x_{ij}(z) \in \{0,1\}$). More generally, I allow $x_{ij}(z)$ to vary continuously between 0 and 1. This formulation has a few advantages. First, it features spatial expansion both across and within regions – an important feature of the data – while still preserving cross-region heterogeneity. This allows me to later fit the model to changes in the number of establishments per firm. Second, it allows me to consider marginal changes in firms' geographical scope, which is useful when characterizing the link to wage inequality. Third, it aids with the quantitative computation of the model, since under suitable functional form for $C(\cdot)$, it is possible to employ an iterative algorithm to compute $\{x_{ij}(z)\}$.

For tractability, I assume that $C(\cdot)$ is paid with freely tradable goods, such that the cost function faced by a firm is independent from its headquarters location. This allows

me to avoid introducing another source of spatial heterogeneity, although pricing it in terms of the local price index is a straightforward extension. In addition, note that if $C(\cdot)$ includes fixed costs and/or negative externalities across regions (e.g. a span of control in the total number of branches, as I allow in the quantitative section), then firms might choose $x_{ij}(z) = 0$ for some regions. Finally, as I elaborate in Section 5 below, I include spatial frictions to firm expansion in this cost function, to match the gravity pattern of headquarters-branch linkages in the data.

**The firm's problem.** Combining the above elements, the firm's problem becomes

$$\max_{h_i(z),\{\ell_{ij}(z)\},\{x_{ij}(z)\}} \sum_{j=1}^{N} \underbrace{x_{ij}(z)\,\Upsilon_j \left( A_i z \left( \gamma h_i(z)^{\frac{\eta-1}{\eta}} + \ell_{ij}(z)^{\frac{\eta-1}{\eta}} \right)^{\frac{\eta}{\eta-1}} \right)^{\frac{\sigma-1}{\sigma}}}_{\text{Sales in region } j} - \sum_{j=1}^{N} \underbrace{x_{ij}(z)\, W_{\ell j} \ell_{ij}(z)^{\frac{\epsilon+1}{\epsilon}}}_{\text{Labor costs in region } j} - \underbrace{W_{hi} h_i(z)^{\frac{\epsilon+1}{\epsilon}} - C\left( \{x_{ij}(z)\}_{j=1}^{N} \right)}_{\text{Headquarters labor and expansion costs}}. \tag{5}$$

Firms maximize the sum of sales across all regions, net of labor costs at all branches, labor costs at their headquarters, and the cost of expansion. They choose the mass of branch-level workers $\ell_{ij}(z)$, headquarters workers $h_i(z)$, and market penetration $x_{ij}(z)$, internalizing that they face a downward sloping demand curve (captured by $\sigma$) and an upward sloping labor supply curve (captured by $\epsilon$) in each location. They take as given labor supply shifters $W_{\ell j}, W_{hi}$ and demand shifters $\Upsilon_j$ in all locations – all of which are determined in equilibrium – as well as the exogenous productivity shifter $A_i$ and the cost function $C(\cdot)$.

## 3.3 Households and other model components

I now layout the households block and the rest of the model's components, which are rather standard in the literature. In doing so, I also provide one particular set of micro-foundations for the labor supply shifters $W_{\ell j}, W_{hi}$ and the demand shifters $\Upsilon_j$ from the firm's problem.

**Households.** There are $S$ types of workers who differ in skill and are indexed by $s$, with an aggregate supply of $\bar{L}_s$ workers of type $s$. For now I set $S = 2$ and assume that headquarters workers correspond to high-skilled workers, and that branch-level workers correspond to low-skilled workers. In the quantification of the model below, I generalize this and allow for more skill-groups and for mixes across skill groups in both branch and headquarters production.

All households consume a bundle of services, $Q$, and a bundle of tradable-goods, $G$. I add the tradable-goods sector to allow for regional specialization in headquarters

services.[18] Households also choose their region $i \in \{1, ..., N\}$ and employer $\nu \in \mathcal{V}_i$, where $\mathcal{V}_i$ is the set of employers in region $i$, i.e., the set of headquarters and branches of locally-active firms. The bundle of services $Q$ is a constant elasticity of substitution aggregator across locally provided varieties, i.e., across the set of firms that open branches in region $i$, taking into account each firm's market penetration in region $i$. Each household of type $s$ solves the following problem, where to avoid overburdening the notation, I omit indexes for specific households:

$$\max_{i\in\{1,...,N\},\nu\in\mathcal{V}_i,G,\{q(\omega)\}_{\omega\in\Omega_i}} b_{i\nu}Q_{si\nu}^{\beta}G_{si\nu}^{1-\beta}, \quad Q_{si\nu} = \left[\int_{\omega\in\Omega_i}\left(q_{si\nu}\left(\omega\right)\right)^{\frac{\sigma-1}{\sigma}}d\omega\right]^{\frac{\sigma}{\sigma-1}}. \tag{6}$$

A household of type $s$ chooses a region $i$, employer $\nu$, quantity of tradable goods $G_{si\nu}$, and a quantity of services $Q_{si\nu}$ to maximize its idiosyncratic preference shock for region $i$ and employer $\nu$, $b_{i\nu}$, and a Cobb–Douglas aggregator of goods and services with an expenditure share $\beta \in (0, 1)$ on services. In turn, $Q_{si\nu}$ aggregates across locally-provided varieties $\omega \in \Omega_i$ with an elasticity of substitution $\sigma$, where $\Omega_i$ is the set of varieties provided in region $i$.[19]

**Regional price indices and demand shifters.** I denote the price index for local services by $P_i$, and the price index that aggregates goods and services by $\bar{P}_i$. The services price index is the standard ideal CES price index, given by $P_i = \left(\int_{\omega\in\Omega_i} P_i\left(\omega\right)^{1-\sigma}d\omega\right)^{\frac{1}{1-\sigma}}$. I choose the price of goods as the numeraire, such that $\bar{P}_i = P_i^{\beta}$. In the quantification of the model in Section 5 I add local housing, which also enters into the formula for the regional final price index $\bar{P}_i$. The demand shifters that firms face in each market from Equation 2 are then given by $\Upsilon_j = E_j^{\frac{1}{\sigma}}P_j^{\frac{\sigma-1}{\sigma}}$, where $E_i$ is regional expenditure on services (a share $\beta$ of total regional income).

**Labor supply.** Households draw the set of idiosyncratic shocks $b_{i\nu}$ from a nested Fréchet distribution, which guides their sorting decisions across space and across employers.[20] The upper nest reflects preferences across locations with dispersion given by $\xi$, capturing a regional labor supply elasticity. The lower nest reflects preferences across employers with dispersion given by $\epsilon$, capturing an employer-level labor supply elasticity.[21]

[18] Such specialization is not possible under balanced trade without an additional sector due to the assumption that the output of branches is not-tradable.

[19] Note that $\Omega_i$ is the mass of varieties faced by each household, taking into account the market penetration decisions of firms. Since each region is a unit-mass of identical locations, each household in region $i$ has a probability $x_{oi}(z)$ to access varieties of an $o$-based firm with productivity $z$ that chooses market penetration $x_{oi}(z)$ in region $i$. The quantity index in region $i$ can thus also be written as $Q_i = \left[\sum_{n=1}^{N}\int_z x_{ni}\left(z\right)\left(q_{ni}\left(z\right)\right)^{\frac{\sigma-1}{\sigma}}dG_n\left(z\right)\right]^{\frac{\sigma}{\sigma-1}}$, where $G_n\left(z\right)$ is the measure of $z$-type firms that are headquartered in market $n$.

[20] Specifically, I assume that $F\left(\boldsymbol{b}\right) = \exp\left(-\sum_{i=1}^{N}\left(\int_{\nu\in\mathcal{V}_i} b_{i\nu}^{-\epsilon}d\nu\right)^{\frac{\xi}{\epsilon}}\right)$.

[21] An employer in the model is a combination of a firm (i.e. a variety $\omega$) and a local market. For example, agents draw different shocks $b_{i\nu}$ for a Starbucks branch in New-York and a Starbucks branch in Seattle. In addition, I assume that headquarters jobs and branch-level jobs for the same firm are distinct

Consequently, the probability that a household of type $s$ chooses region $i$ is increasing in the effective regional real income for type $s$, and given by

$$\frac{L_{is}}{\sum_{j=1}^{N} L_{js}} = \frac{\left(W_{si}/\bar{P}_i\right)^{\xi}}{\sum_{j=1}^{N} \left(W_{sj}/\bar{P}_j\right)^{\xi}}, \tag{7}$$

where $W_{si}$ is the ideal wage index for workers of type $s$ in region $i$:

$$W_{si} = \left( L_{is}^{-1} \int_{\nu \in \mathcal{V}_i} w_{is}\left(\nu\right)^{\epsilon} d\nu \right)^{\frac{1}{\epsilon}}. \tag{8}$$

This wage index reflects both the level of wages and the variety of jobs in that region. When the employer-specific labor supply elasticity is infinite ($\epsilon \to \infty$), all local employers pay a common wage for type-$s$ workers, equal to the index $W_{si}$. Otherwise, firms need to pay higher wages to attract more workers, giving rise to upward-sloping firm-specific labor supply curves as given in Equation 4.

**Entry.** Firms can enter freely in each market by paying $f_e$ units of the tradable good. The entry location determines the headquarters location of a firm. I choose to price entry at the common numeraire (i.e. in units of tradable-goods) to avoid heterogeneity in entry costs across space, though allowing for this margin of heterogeneity (e.g. by using regional price indices) is straightforward.

**Production of tradable goods.** For simplicity, I assume that tradable-goods are freely traded and produced using constant returns to scale technology. Firms produce using a Cobb-Douglas production function that combines high and low skilled labor, with constant expenditure shares $\alpha$ and $1-\alpha$, respectively. Similar to the services sector, firms can enter freely after paying $f_e$ units of the local final good.[22]

This completes the set-up of the model. I now turn to characterize the link between firms' geographical scope and the wage structure in the economy.

## 3.4 Firm spatial expansion and skill-biased technical change

The first takeaway from the model is a new micro-foundation for skill-biased technical-change (SBTC). Specifically, I show that changes in firms' scope are isomorphic to technical change that augments the productivity of headquarters workers relative to branch workers. To see this, I study the share of labor costs accrued to headquarters workers in

jobs: working for the Starbucks HQ in Seattle is not the same as working for one of its branches there. Therefore, firms face multiple labor supply curves: one for each market in which they have a branch, and another one for their headquarters.

[22]Note that since firms in this sector operate in the same labor markets as service firms, they also face the same firm-specific upward-sloping labor supply curves and earn rents due to labor market power. The free entry condition ensures that net profits are zero in equilibrium.

service firms. Recall that this cost-share pins down skill intensity in the services sector, while in the tradable-goods sector skill intensity is exogenous by construction. I denote this cost-share by $\Gamma_i(z)$, and characterize it in the following proposition:

**Proposition 1.** *Geographical scope as a micro-foundation for skill-biased technical change.*

*(a) The share of costs accrued to headquarters workers in firm $(i, z)$ is given by*

$$\Gamma_i(z) = \frac{\gamma^{\bar{\eta}} \left( \frac{W_{hi}}{\sum_{j=1}^{N} x_{ij}(z)} \right)^{1-\bar{\eta}}}{\bar{W}_{\ell i}(z)^{1-\bar{\eta}} + \gamma^{\bar{\eta}} \left( \frac{W_{hi}}{\sum_{j=1}^{N} x_{ij}(z)} \right)^{1-\bar{\eta}}}, \quad \bar{\eta} \equiv \frac{\eta}{\eta - \frac{\epsilon}{\epsilon+1}(\eta - 1)}$$

*where $\bar{W}_{\ell i}(z) \equiv \left( \sum_{j=1}^{N} \Omega_{ij}(z) W_{\ell j}^{\bar{\eta}-1} \right)^{\frac{1}{\bar{\eta}-1}}$ is a weighted power-mean of the labor supply shifters $W_{\ell j}$ across all of the firm's markets. The loadings $\Omega_{ij}(z)$ capture the share of region $j$ in the firm's total branch-level payroll, as well as its dependency on the headquarters input relative to the firm's average.*

*(b) Consider the case that $\epsilon \to \infty$, all locations are symmetric, the aggregate endowments of headquarters labor and branch-level labor are given by $\bar{H}$ and $\bar{L}$, respectively, and $\beta = 1$. Then, $W_h = W_{hi}$ and $W_\ell = W_{\ell j}$ are the economy-wide prices of headquarters and branch-level labor, and their ratio is given by*

$$\frac{W_h}{W_\ell} = \gamma \bar{x}^{\frac{\eta-1}{\eta}} \left( \frac{\bar{H}}{\bar{L}} \right)^{-\frac{1}{\eta}},$$

*where $\bar{x}$ is a weighted power-mean of firms' geographical scope in the economy, with the weight of each firm determined by its total branch-level employment.*

**Proof**: see Appendix B.

The first part of Proposition 1 characterizes the cost-share of headquarters workers at the firm level in terms of four key objects. Similar to the SBTC literature, this cost-share depends on: (1) relative factor prices, as captured by the ideal wage index for headquarters workers in the firm's headquarters region ($W_{hi}$), and a power-mean of the ideal wage indices for branch-level workers across all markets, ($W_{\ell j}$); (2) an exogenous technological shifter, as captured by the productivity of headquarters labor $\gamma$; and (3) the elasticity of substitution between factors, which in this case depends on the raw elasticity between headquarters and branch-level workers in the production function, $\eta$, and on the labor supply elasticity $\epsilon$ (since the latter affects the firm's ability to substitute across workers in different locations). Departing from the SBTC literature, the firm's geographical scope – captured by the sum of the firm's market penetration terms $\sum_{j=1}^{N} x_{ij}(z)$ – also enters this expression, in the same way as a reduction in the cost of headquarters labor. Therefore,

an increase in firm scope acts as a factor-biased technical change for the factors that are used intensively in the production of the firm's non-rival component.

To see most clearly the implications for aggregate inequality across skill-groups, the second part of Proposition 1 considers a special case of the model with symmetric space, competitive labor markets ($\epsilon \to \infty$), and exogenous aggregate factor endowments of headquarters and branch workers ($\bar{H}$ and $\bar{L}$). In this case, the economy-wide wage differential between headquarters and branch workers is decreasing with the relative factor endowment of headquarters workers ($\bar{H}/\bar{L}$) and increasing with their relative productivity $\gamma$, as is standard in the literature. In addition, when $\eta > 1$, it also increases with the employment-weighted average of firms' geographical scope in the economy, captured by $\bar{x}$. Note that in the simple case of symmetric space, market penetration is common across locations $x_{ij}(z) = x(z)$, the firm's geographical scope is given by $x(z)N$, and $\bar{x}$ is the employment-weighted geographical scope in the economy.

Intuitively, the non-rivalry assumption implies that the marginal product of headquarters workers scales with the number of locations in which their output can be applied, whereas the marginal product of branch workers does not. The above result does not take a stance on what drives changes in firms' geographical scope ($\bar{x}$ in the second part of the proposition). Rather, it highlights that conditional on factor endowments and relative productivity $\gamma$, shocks that change firms' geographical scope (such as a reduction in the cost of expansion $C$) would act as a technological change that augments the productivity of headquarters workers.

## 3.5 Firm spatial expansion and firm-specific wages

The above result highlights the role of firms' spatial scope for inequality across skill-groups, and holds also for the version of the model with $\epsilon \to \infty$. In practice, residual wage inequality is pervasive, and accounts for most of the increase in wage dispersion over time. The model allows for such residual inequality when $\epsilon$ is finite. In this case, different firms and establishments pay different wages, conditional on skill group and location. The following Proposition characterizes firm-specific wages in the model:

**Proposition 2.** *The wages that firm $(i, z)$ pays to its headquarters workers and branch-level workers can be expressed as*

$$\underbrace{\ln w_{\ell,ij}(z)}_{\text{Wages to local labor}} = const_\ell + \underbrace{\frac{\epsilon}{\epsilon+1}\ln W_{\ell,j}}_{\text{Local labor market}} + \underbrace{\frac{1}{\epsilon+1}\ln r_{ij}(z)}_{\text{Local sales}} + \underbrace{\frac{1}{\epsilon+1}\ln\left(1-\Gamma_{ij}(z)\right)}_{\text{Local labor intensity}}$$

$$\underbrace{\ln w_{h,i}(z)}_{\text{Wages to headquarters labor}} = const_h + \underbrace{\frac{\epsilon}{\epsilon+1}\ln W_{h,i}}_{\text{Local labor market}} + \underbrace{\frac{1}{\epsilon+1}\ln\sum_j x_{ij(z)}r_{ij}(z)}_{\text{Total firm sales}} + \underbrace{\frac{1}{\epsilon+1}\ln\Gamma_i(z)}_{\text{HQ intensity}}$$

**Proof**: see Appendix B.

When $\epsilon \to \infty$, all firms pay the same market-level wages, given by the wage indices $W_{\ell,j}$ and $W_{h,i}$. Otherwise, branch level wages increase with local output (as captured by local sales $r_{ij}(z)$) and decrease with the intensity of local output in the non-rival headquarters production, which I label $\Gamma_{ij}(z)$.[23] In contrast to branch-level wages, headquarters wages rise with the firm's *total* sales across all markets, as well as with the firm-level cost-share of headquarters from Proposition 1, $\Gamma_i(z)$.

These expressions help to build-up intuition for the effects of spatial expansion on firm-specific wages and within-firm inequality. For a given set of regional wage-indices, spatial expansion at the firm level pushes up wage dispersion between its headquarters and branches, by raising the ratio of total sales to local sales. This is true in particular when $\eta = 1$, in which case $\Gamma_{ij}(z) = \Gamma_i(z) = \frac{\gamma}{1+\gamma}$ is constant and does not vary with firm scope. If in addition $\eta > 1$, spatial expansion also reallocates production from the branches to the headquarters (a rise in $\Gamma_{ij}(z)$ and $\Gamma_i(z)$, as in Proposition 1), adding another force that pushes up within-firm wage dispersion.[24]

More formally, the following proposition describes how firm-level wages respond when expansion is driven by a small reduction in the cost of expansion $C$.

**Proposition 3.** *Consider a firm that experiences a small decline in its cost-of expansion $dC(\cdot)$ that results in non-negative expansion in all markets ($dx_{ij}(z) \geq 0$). Consequently:*

*(a) The firm always raises headquarters wages ($w_{h,i}(z)$).*

*(b) The firm raises branch-level wages ($w_{\ell,ij}(z)$) only if $\sigma > \eta$.*

*(c) The ratio of headquarters wages to branch-level wages ($w_{h,i}(z)/w_{\ell,ij}(z)$) rises, irrespective of the values for $\sigma$ and $\eta$.*

**Proof**: see Appendix B.

According to Proposition 3, an increase in the geographical scope of a firm due to lower cost of expansion leads to higher headquarters wages and higher dispersion between headquarters and branches. Interestingly, the effect on branch-level wages is ambiguous,

[23] $\Gamma_{ij}(z)$ is the local equivalent of the firm-level cost-share of headquarters from Proposition 1. It is defined as the share of local branch sales that is used to pay headquarters workers, or one minus the ratio of local wage bill to local sales. See Appendix B for more details.

[24] On top of these effects, an increase in firms' geographical scope raises wage-dispersion across branches due to a mechanical effect, as the firm becomes active in more markets with different regional wage indices. This same mechanical effect pushes down inequality between firms.

as can be seen also in the two opposing terms in the expression for $w_{\ell,ij}(z)$ in Proposition 2. On the one hand, spatial expansion generates a positive *productivity effect* across all branches, following an increase in the firm's non-rival inputs. This is captured by an increase in local sales in Proposition 2. On the other hand, it generates an increase in *headquarters intensity* when $\eta > 1$, which lowers demand and wages for local branch workers. On net, branch-level wages go up if $\sigma > \eta$.

It is worth benchmarking these results relative to the literature on firms in monopsonistic labor markets, as in Card et al. (2018) and Berger et al. (2022). The typical outcome from such models is that larger firms pay higher wages. In the current framework, this is still true on average, though it is also associated with larger within-firm dispersion. Moreover, despite the existence of upward-sloping labor supply curves, firm growth does not necessarily lead to higher wages for branch workers, as demonstrated in Proposition 3. Consequently, while there is a clear positive firm size gradient for headquarters wages, it can be very flat (or even negative if $\sigma < \eta$) for branch wages. In addition, it highlights that firm-level shocks pass-through more to headquarters wages than to branch wages, in line with evidence on differential pass-through in Kline et al. (2019).

A similar result to Proposition 3 also applies following a small increase in the firm-level productivity $z$. Therefore, it is also a characterization of wage dispersion between firms with different productivities in the cross-section. Conditional on local wage indices, higher productivity firms pay higher wages on average, but not necessarily in their branches; and they are characterized by higher within-firm dispersion.

The above propositions also highlight the special role of space for wage inequality across firms and establishments. First, unsurprisingly, part of overall wage dispersion is due to differences in regional wage indices ($W_{\ell j}$ and $W_{hi}$ in Proposition 2). Second, firm wages in each market $j$ are also a function of the firm's headquarters location $i$.[25] This is reflected in the dependency of regional sales and local labor intensity ($r_{ij}(z)$ and $\Gamma_{ij}(z)$ in Proposition 2) on the firm's headquarters location $i$, on top of the firm's productivity $z$ and the local market $j$. Therefore, wages are characterized by a *headquarters market effect*, in line with evidence on wage setting in affiliates of multinational firms, as in Setzler and Tintelnot (2021).[26] An implication of this headquarters market effect is that changes in the spatial distribution of economic activity can affect not only inequality across regions, but also inequality within any region: two branches can pay increasingly different wages from each-other even if they are in the same market $j$ and have the same firm productivity $z$, if their headquarters market $i$ is not the same.

[25] The dependency on the headquarters location $i$ is mediated through choices of the firm's non-rival component $h$, which are affected by headquarters-market characteristics such as the productivity $A_i$ and the cost of headquarters labor $W_{hi}$.

[26] Through the lens of simple wage decompositions as in Section 2, I find that this effect is quantitatively meaningful. Firms' headquarters locations account for between 30%-50% of the between-firm component of overall wage inequality for multi-region firms.

## 3.6 Firm spatial expansion and spatial disparities

The model also yields predictions for the connection between firm scope and spatial disparities in skill intensity, yielding a new mechanism for what Moretti (2012) described as "the great divergence".

**Proposition 4.** *Spatial disparities in skill-intensity.*

*(a) The share of skilled-labor in aggregate income is given by*

$$\left(1-\tilde{\beta}\right)\alpha+\tilde{\beta}\bar{\Gamma},$$

*where we define* $\tilde{\beta}\equiv\beta\frac{\epsilon}{\epsilon+1}\frac{\sigma-1}{\sigma}$, *and* $\bar{\Gamma}$ *is the sales-weighted average of the headquarters intensity* $\Gamma_i(z)$ *from Proposition 1 across all firms and regions in the economy.*

*(b) The share of skilled-labor in region i's income is given by*

$$\left(1-\tilde{\beta}\right)\alpha+\tilde{\beta}\left[(1-\alpha)\frac{R_i}{E_i}\Gamma_{out,i}+\alpha\Gamma_{in,i}\right],$$

*where* $R_i$ *is total (national) sales by locally-headquartered firms in region i;* $E_i$ *is total expenditure on services in region i; and* $\Gamma_{out,i}$ *and* $\Gamma_{in,i}$ *are the sales-weighted averages of the headquarters intensity for locally-headquartered firms and for locally-active branches, respectively.*

*(c) When* $\eta=1$, *then* $\bar{\Gamma}=\Gamma_{out,i}=\Gamma_{in,i}=\frac{\gamma}{1+\gamma}$. *In this case, a sufficient statistic for region i's high-skilled income share is its specialization in providing headquarters services, as summarized by the ratio of total sales by locally-headquartered firms,* $R_i$, *to local expenditure on services,* $E_i$.

*(d) In the limit economy with no multi-location firms (*$C(\cdot)\to\infty$ *when* $x_{ij}(z)>0$ *for* $i\neq j$*), the share of region i's income accrued to high-skilled labor is summarized the sales-weighted average of the headquarters intensity of its firms,* $\Gamma_i=\Gamma_{out,i}=\Gamma_{in,i}$. *If in addition* $\eta=1$, *there are no spatial differences in high-skilled income shares.*

**Proof**: see Appendix B.

Proposition 4 starts with characterizing aggregate skill intensity in the economy, as a benchmark for regional skill intensity. The aggregate income-share of skilled labor is a weighted average of the (exogenous) skill intensity in the tradable-goods sector, $\alpha$, and the average intensity of the non-rival component of firms in services, $\bar{\Gamma}$. The weight $\tilde{\beta}$ reflects the importance of services in the economy. In turn, regional skill intensity depends in addition on two region-specific objects. The first object, captured by $R_i/E_i$, is regional specialization in headquarters services (as opposed to specialization in production of tradable goods). When the non-rival component of services is more skill-intensive than

production of tradables, greater specialization in headquarters services leads to higher relative demand for skill. Note that in the above baseline model this is always the case, since I assumed that only skilled labor is used in headquarters production, though in the quantification below I relax this. When $\eta = 1$, this measure of specialization is also a *sufficient statistic* for regional skill intensity (part (c) of Proposition 4). The scope for such specialization is tightly linked to firms' geographical scope. For example, when cross-region firm activity is prohibited (part (d) of Proposition 4), all headquarters services are supplied locally, and there is no room for regions to specialize in this activity: $R_i = E_i$ for all $i$, and there is no spatial variation in $R_i/E_i$.

The second endogenous object that shapes regional skill intensity is the share of non-rival factors in the cost structure of local firms ($\Gamma_{out,i}$ and $\Gamma_{in,i}$). This object itself is a function of firms' spatial scope, as summarized in Proposition 1. Therefore, through these two channels, an increase in firms' spatial scope can lead to greater spatial disparities in the economy. An illustrative case is the limit economy with no multi-region firms ($C(\cdot) \to \infty$ when $x_{ij}(z) > 0$ for $i \neq j$) and $\eta \to 1$. In this case, there are no spatial disparities in high-skilled income shares, demonstrating the importance of firms' scope for spatial disparities in an extreme case.

# 4 Empirical evidence for the main mechanisms

I now turn to provide reduced-form evidence for the key assumptions and implications of the above model.

## 4.1 Within-firm non-rivalries

A key assumption in the model is that the output of headquarters workers is non-rival across the firm's locations. A central challenge is that non-rivalry is not directly observable: we cannot directly measure whether the use of an idea by one establishment reduces the effectiveness of this idea elsewhere. However, we can empirically test implications of this assumption, which I do in the following sections.

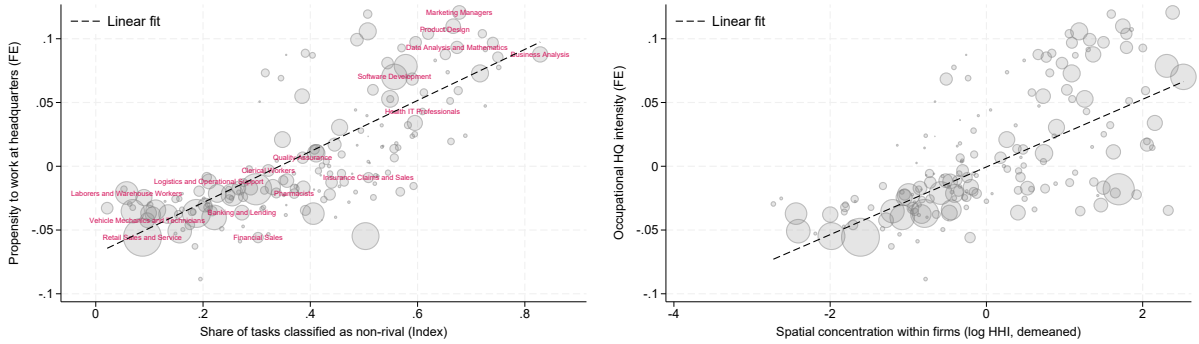
### 4.1.1 Qualitative evidence for headquarters non-rivalries from task-level data

First, I examine whether the tasks that headquarters workers perform in practice are consistent with the theoretical concept of non-rivalry. To this end, I first compute a measure of the non-rivalry of tasks across occupations, and then relate it to the propensity of each occupation to be employed in firms' headquarters. Using task-level data from O*NET and a large language model (LLM), I classify each of the 19,530 O*NET tasks as

either non-rival, rival, or ambiguous.[27] For example, the task "Develop pricing strategies, balancing firm objectives and customer satisfaction" is classified as non-rival, while the task "Answer telephone calls to assist computer users encountering problems" is classified as rival. Subsequently, for each occupation, I construct a measure of output non-rivalry based on the proportion of tasks classified as non-rival out of the total number of non-ambiguous tasks.

**Figure 3: The non-rival nature of headquarters jobs**

**(a)** Task non-rivalry and headquarters intensity **(b)** Spatial concentration and headquarters intensity




Note: Panel (a) plots a measure of occupational headquarters propensity (y-axis) against the non-rival content of each occupation's tasks (x-axis). Headquarters propensity is constructed from job-postings data (Lightcast) and headquarters location data (from Dun & Bradstreet), and is defined as the occupational fixed effect after controlling for firm and location in each posting (see Section 2.1). The non-rival index is measured as the share of an occupation's tasks in O*NET that a large language model classifies as non-rival (see Section 4.1). Circle size reflects the total number of jobs in each occupation. Panel (b) relates the same headquarters propensity measure to the geographic concentration of occupations within firms. The x-axis reports the average across firms of the log within-firm normalized Herfindahl–Hirschman index (HHI) for the distribution of jobs across firms' locations.

The results reveal a strong relationship between the non-rival content of tasks and the propensity of occupations to be performed at firms' headquarters, as can be seen in panel (a) of Figure 3. The plot shows occupational headquarters propensity on the y-axis (see Section 2.1 for more details) against the above measure of task non-rivalry on the x-axis, across all Lightcast 4-digit occupation groups. Occupation groups such as Product Design, Business Analysis, and Software Development are classified as having highly non-rival output and are frequently performed at firms' headquarters. In contrast, occupations like Financial Sales, Food and Beverage Preparation, and Healthcare Therapy tend to produce rival output and are rarely performed at headquarters.[28]

[27] See Appendix E.3 for the complete prompt used with the LLM. Specifically, non-rivalry is defined as: "Non-rivalry occurs when a task output can be used simultaneously across multiple firm locations without diminishing its value, availability, or functionality at any single location."

[28] Naturally, both measures are closely tied to the skill intensity of occupations, which is desirable since my aim is to provide a micro-foundation for skill-biased technical change. Importantly, however, the link between non-rivalry and headquarters activity remains strong even after controlling for standard measures of skill, such as college intensity by occupation. For example, physicians are highly skilled, yet they are classified as having low headquarters intensity and a low non-rivalry index.

Another implication of non-rivalry is tradability: while rival tasks often need to be performed near consumers — especially in the case of the non-tradable services I focus on — firms can, in principle, produce the non-rival components of output anywhere. This suggests that spatial concentration within firms offers an alternative proxy for non-rivalry. I examine this in Panel (b) of Figure 3, which plots the headquarters propensity of each occupation against a measure of its within-firm spatial concentration. Specifically, I compute the Herfindahl–Hirschman Index (HHI) for the distribution of job postings across locations within each firm and occupation, then average across firms. The x-axis shows the (demeaned) log of this measure. Evidently, occupations more likely to be performed at headquarters are also more spatially concentrated within firms.

#### 4.1.2 Testing non-rivalry through the compensation of headquarters workers

A second piece of evidence for headquarters non-rivalry concerns the compensation of headquarters workers. The non-rivalry assumption implies that the marginal product of headquarters workers scales with the number of firm locations, other things equal. To the extent that a worker's salary is related to their marginal product, a natural test is whether a change in the number of firm locations leads to an increase in the compensation to headquarters workers relative to branch-level workers in that firm.[29]

This test aligns with the model's predictions as demonstrated in Proposition 3. In that proposition, the firm's expansion is induced by a decrease in the cost of expansion ($C$). Similar results are obtained when expansion is induced by an increase in firm-level productivity ($z$) and an increase in demand ($\Upsilon_j$). All of these shocks are factor-neutral conditional on the number of locations, but raise the marginal product of headquarters workers (relative to branch workers) when the firm can adjust its number of locations. As I discuss later, a potential threat to this test is when firm growth is driven by a firm-specific non-neutral productivity shock. In the model, this can be clearly seen in the wage formulas from Proposition 2, in which a firm-specific change in $\Gamma_i(z)$ can shift relative wages without a change in the number of locations. I address this concern in the empirical analysis below.

To study changes in relative headquarters wages within firms, I use data on salaries from firms' job postings in the Lightcast database. Although posted wages do not perfectly capture actual compensation, they offer two advantages for my analysis. First, they provide firm-level measures of compensation by occupation that are not available in U.S. administrative data. Second, they reflect changes in firms' labor demand, abstracting from concerns about wage stickiness for incumbent workers and hiring success.

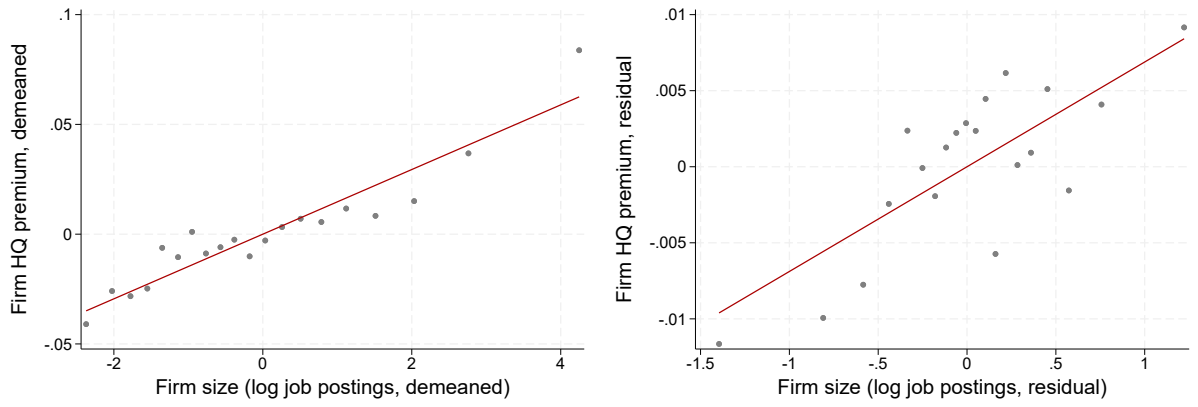
[29]Note that while the model has clear predictions for the response of relative headquarters wages to firm size, it does not have clear predictions for relative headquarters quantities. For example, if $\eta = 1$, the headquarters payroll share is constant, implying that as firms expand, relative headquarters wages rise, while relative headquarters employment falls. However, if $\eta$ is particularly high, then relative headquarters employment can rise with firm size.

To measure the headquarters premium in the data—the analogue of $w_{h,i}(z)/\bar{w}_{\ell,i}(z)$ in the model, where $\bar{w}_{\ell,i}(z)$ denotes the average of $w_{\ell,ij}(z)$ across $j$—I use my occupational headquarters-intensity measure (see Section 2.1), computing the ratio of salaries in headquarters jobs relative to branch-level jobs. I then examine how this ratio changes when firms expand. I proxy growth in the number of firm branches with the total number of firm job postings, or with the number of counties in which the firm posts job openings.

The baseline relationship between firm size and the headquarters premium appears in the left panel of Figure 4. This panel presents a binscatter of the log headquarters premium (y-axis) against the log number of job postings across all occupations (x-axis). The plot reveals a clear positive relationship: larger firms tend to pay higher headquarters wages relative to branch-level wages. While consistent with the model's predictions, this finding raises several concerns. First, firms may differ systematically in their factor composition beyond what is captured by size. Second, firm growth could be driven by non-neutral technical change. To address these issues, the right panel introduces a set of controls: firm and year fixed effects to absorb baseline firm heterogeneity, and sector×year fixed effects to capture trends in sectoral technical change. The adjusted plot continues to display a strong positive relationship between firm size and the relative compensation of headquarters jobs. As firms expand, they raise posted salaries for headquarters occupations relative to branch-level occupations within the same firm. Quantitatively, according to these plots, doubling firm size is associated with an increase of relative headquarters wages of approximately 1%.

**Figure 4: As firms expand, relative headquarters wages rise**

**(a)** Cross-sectional relationship **(b)** Controlling for firm effects and sectoral trends




Note: the left panel shows a binscatter plot of the (log) relative headquarters wages (y-axis) against the log number of job postings (x-axis) across firms. Relative headquarters wages are computed as the average posted salary for the top 10% of the most headquarters-intensive job-postings in the firm, divided by the average posted salary for the remaining 90% of postings. The right panel shows the same relationship after controlling for firm, year, and sector×year fixed effects. See Section 4.1 for additional details on the construction of all variables.

I now turn to investigate this relationship using an instrumental variable (IV) approach to address two additional concerns. First, as discussed above, non-neutral productivity growth might be firm-specific, and not accounted for by the sector×year fixed effects. This can bias the above result in either direction, depending on the direction of firm-specific technical change. Second, a standard attenuation bias is very natural in this setting, given the imperfect mapping between firms' scope (number of branches) and its number of job postings or number of counties with job postings.

To this end, I instrument changes in firm size using exposure to regional economic growth. The intuition is simple: consider two otherwise identical coffee shop chains, one concentrated in Austin, Texas (e.g., *Summer Moon Coffee*), and the other in Indianapolis, Indiana (e.g., *Indie Coffee Roasters*). The former is likely to benefit from Austin's booming economy, while the latter faces a more stagnant environment. Comparing differences in firm growth between such firms isolates expansion driven by factor-neutral demand shifts (i.e., a change in $\Upsilon_j$ in the model). In practice, I employ a "shift-share" design, such that for a firm $f$ in year $t$, the IV is defined as

$$IV_{f,t} = \sum_{i=1}^{N} \frac{\mathbb{I}_{i,f,t-1}}{N_{i,f,t-1}} \times \ln \frac{GDP_{it}}{GDP_{i,t-1}} \tag{9}$$

where $\mathbb{I}_{i,f,t-1}$ is an indicator for whether firm $f$ had job postings in region $i$ in year $t-1$, and $N_{i,f,t-1}$ is the number of regions in which $f$ had postings in $t-1$. The shifter $\ln\left(\frac{GDP_{it}}{GDP_{i,t-1}}\right)$ captures regional GDP growth between $t-1$ and $t$, obtained from the BEA's regional economic accounts. The identifying assumption is that, conditional on sectoral trends, firm-specific non-neutral technical change is orthogonal to regional GDP growth.[30] As an alternative, I construct an instrument that fixes firm weights at their initial year of appearance in the data rather than updating them each period. This version has lower statistical power, since it does not track time-varying exposure to regional demand shocks, but it avoids introducing identifying variation from changes in a firm's market weights over time.[31]

Results from this analysis are reported in Table 2. Column (1) revisits the OLS relationship from the right panel of Figure 4. It shows the coefficient from regressing log relative headquarters wages on log firm size, controlling for firm, year, and sector-by-year fixed effects. Panel A shows result when measuring firm size as the log number of job postings, while Panel B shows results when measuring firm size as the log number of counties. In both cases, the coefficient is positive and statistically significant. Column (2) presents the baseline 2SLS estimate. The coefficient remains positive and statistically

[30] Using regional population growth instead of GDP growth yields qualitatively similar results, though—unsurprisingly—the instrument is weaker in terms of statistical power.

[31] More specifically, $IV^B_{f,t} = \sum_{i=1}^{N} \frac{\mathbb{I}_{i,f,t_0(f)}}{N_{i,f,t_0(f)}} \times \ln \frac{GDP_{it}}{GDP_{i,t_0(f)}}$, where $t_0(f)$ is the first year in which firm $f$ appears in the data.

significant, with first-stage F-statistics above conventional thresholds. The point estimates are larger than the corresponding OLS estimates, suggesting either strong attenuation bias in OLS or that firm-specific growth is biased towards branch workers. Despite different point estimates across both panels (0.14 when measuring size with # of postings, and 0.33 when measuring it with # of counties), there is substantial overlap in their confidence intervals, with the range [0.04,0.24] covered by both CIs. Column (4) considers the same regression with the alternative IV that uses weights from the initial year in which each firm appears in the data. The point estimates remain similar to my baseline IV, although the standard errors are larger and the first-stage F-stats are lower (particularly when measuring size with the number of firm counties).

While one of the goals of this paper is to explain differences across skill groups, a natural question is whether the results also hold within skill groups. In principle, even conditional on skill, the non-rivalry assumption implies that more headquarters-intensive occupations should experience greater wage growth when firms expand. Column (3) of Table 2 provides such a test by replacing the outcome variable with relative headquarters wages after residualizing out differences associated with educational requirements, in particular whether a job requires a college degree. The results remain very similar, suggesting that the results are not driven purely by differences in educational requirements across occupations.[32]

The estimated effect cannot generally be mapped to a single structural parameter. However, in the special case with $\eta = 1$ (a constant headquarters payroll share) and symmetric locations, the elasticity of the relative headquarters wage with respect to firm size equals $\frac{1}{\epsilon+1}$ (see Proposition 2). In this knife-edge case, the coefficient identifies the firm-specific labor supply elasticity: the number of locations pins down the relative change in headquarters versus branch output, and the labor supply elasticity translates relative output into relative wages. Taking my average 2SLS estimate, we obtain an implied firm-specific labor supply elasticity of roughly 6, consistent with the literature.[33] Note, however, that unlike most of the existing literature, the relevant elasticity here governs the transmission from firm size to within-firm wage dispersion, in particular the headquarters premium, and not only to firm-average wages. More generally, with asymmetric space and $\eta \neq 1$, the estimated coefficient does not map to a single parameter, but remains tightly linked to the firm's labor supply elasticity, the substitutability between headquarters and branch workers ($\eta$), and the degree of non-rivalry ($\gamma$). As discussed below, the full quantitative model delivers headquarters-premium elasticities with respect to firm size that are similar to those estimated in the data.

More generally, beyond the specific classification into headquarters and branches, a central implication of the theory is that pass-through of firm shocks to wages depends

[32] Similar patterns also emerge when controlling for geographic variation.

[33] See, for example, Lamadon et al. (2022) and Berger et al. (2022).

**Table 2:** Firm expansion and relative headquarters wages

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Outcome: log relative headquarters wages | OLS | IV (A) | IV (A) | IV (B) |
| **Panel A: size measured by # of jobs** | | | | |
| Log # of job postings | 0.007*** | 0.137** | 0.114** | 0.116* |
| | (0.002) | (0.054) | (0.052) | (0.063) |
| First-Stage F-stat (KP-Wald) | – | 47.600 | 47.600 | 34.890 |
| **Panel B: size measured by # of counties** | | | | |
| Log # of counties | 0.012*** | 0.330** | 0.274** | 0.394 |
| | (0.003) | (0.146) | (0.138) | (0.262) |
| First-Stage F-stat (KP-Wald) | – | 14.460 | 14.460 | 5.185 |
| Firm and year×sector fixed effects | Y | Y | Y | Y |
| Residualize educational requirement | N | N | Y | N |
| Observations | 65,055 | 65,055 | 65,055 | 65,055 |

*Notes:* This table shows how firm expansion affects relative headquarters wages in the firm. Firm size is measured either as the log number of job postings (Panel A) or as the log number of counties with job postings (Panel B). The outcome variable is the log ratio of average posted salaries in the top 10% most headquarters-intensive occupations to those in the remaining 90% of occupations within each firm. Column (1) shows OLS results, controlling for firm, year, and sector-by-year fixed effects. Column (2) shows the baseline 2SLS estimates, using a firm's exposure to regional economic growth, as given in Equation 9 in Section 4.1. Column (3) replaces the outcome variable with relative headquarters wages after residualizing out differences associated with educational requirements, in particular whether a job requires a college degree. Column (4) presents 2SLS estimates using an alternative IV that fixes firm weights at their initial year of appearance in the data rather than updating them each period - see addition details in Section 4.1. The sample is restricted to firms in services sectors with at least 5 job-postings classified as branches and at least 1 job-posting classified as headquarters, and exluding outlier firms in which the headquarters share is below 1% or above 99%. Standard errors are clustered at the firm level. $^{*}$ $p < 0.1$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$.

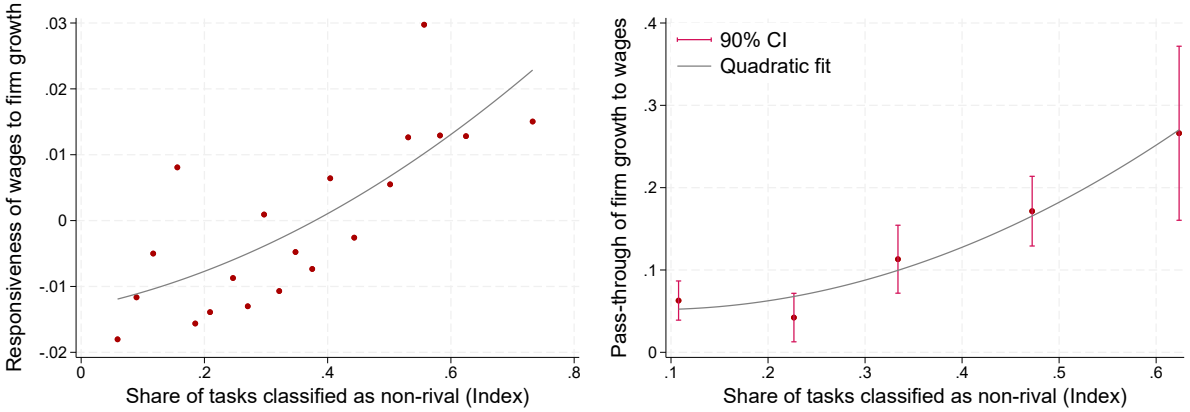
systematically on the non-rival nature of a job. In particular, ranking occupations by the non-rivalry of their output within the firm, the model predicts that wages in more non-rival occupations should be more responsive to firm growth. I test this prediction in Figure 5, which plots pass-through of firm growth to wages for each occupation against the task-based non-rivalry index for each occupation. To construct occupation-level pass-through, I project log wages in each occupation on total firm size, following the same approach used above for the headquarters premium.

The left panel shows a binscatter of OLS estimates, controlling for firm, year, and sector-by-year fixed effects, and normalized relative to the mean pass-through level. Although these OLS estimates are biased, they reveal a clear positive relationship: wages in more non-rival occupations respond more strongly to firm growth. The right panel shows a plot for 2SLS estimates, instrumenting firm growth with the IV defined in Equation

9. Since my IV approach lacks the statistical power to estimate pass-through for each occupation individually, I group occupations into five bins by their non-rivalry index and estimate pass-through for each bin. The overall average pass-through is around 0.15, consistent with existing estimates in the literature. More importantly, pass-through is relatively low for occupations with limited non-rivalry (below 0.1), but rises sharply for highly non-rival occupations (around 0.2-0.3), matching the theory's predictions.

**Figure 5: Stronger non-rivalry implies higher passthrough of firm growth to wages**

**(a)** Passthrough across all occupation groups (OLS) **(b)** Passthrough across occupation bins (IV)




Note: the figure plots occupation-level pass-through of firm growth to wages against a measure of the non-rivalary of the occupation output. Panel (a) shows a binscatter of OLS estimates of pass-through against the measure of non-rivalry from Section 4.1. The pass-through is computed as the estimate a regression of log firm wages on log firm size for a given occupation, controlling for firm, year, and sector-by-year fixed effects, and normalized relative to the mean pass-through level. Panel (b) shows a scatterplot of 2SLS estimates of pass-through for 5 equally-sized bins of occupations, where firm growth is instrumented using the IV defined in Equation 9. Firm size and firm-by-occupation wages obtained from Lightcast.

Finally, I examine whether similar patterns arise in the U.S. Census LBD data. In this setting, the measurement of firm scope is improved by administrative records on firms' establishments. A key limitation, however, is the absence of occupational information and reliable measures of headquarters activity. To address this, I proxy relative headquarters wages using average wages across establishments within a firm, classifying certain establishments as headquarters based on their industry codes (see more details in Section 2.1). The full results are reported in Appendix E.4. Although this measure is more restrictive and available only for a subset of firms, the findings are reassuringly similar. In particular, regressing log headquarters premium on the log number of firm establishments—controlling for firm, year, and sector-by-year fixed effects—yields a coefficient of about 0.04, which falls between the OLS and 2SLS estimates from the Lightcast data in Table 2.

### 4.1.3 Testing non-rivalry through spillovers across firm locations

Another implication and test of the non-rivalry assumption is that demand shocks propagate across firms' branches. Rising demand in a subset of the firm's locations raises the market size of the firm, inducing it to increase the firm's non-rival inputs. Consequentially, branches in unaffected markets experience a productivity boost, leading to greater firm activity in those uneffected locations. This effect declines with the strength of non-rivalries. The model also includes two forces that can work in the opposing direction. First, if the cost of expansion $C$ exhibits convexities (e.g., due a span-of-control cost, as I allow in the quantitative version of the model below), then expansion in one market would raise the marginal cost of expansion in other markets, potentially leading to less firm activity there. Second, if the elasticity of substitution between headquarters and branch workers is very high, more expansion opportunities in one market would reallocate production away from all markets to the headquarters.

I investigate whether such spillovers exist in practice by studying how regional demand shocks affect firm branch activity in other regions. Since this exercise does not require to measure headquarters activity, I implement it with data from the U.S. Census LBD, since its measurement of firm activity across space is more comprehensive relative to commercial datasets. However, I find that similar results hold in the commercial datasets that I use in the paper, the lightcast data and the Dun and Bradstreet data.

More specifically, let $y_{fjt}$ denote an outcome of interest for firm $f$ in market $j$ in period $t$, and let $Y_{jt}$ denote the GDP of market $j$ in period $t$. The key question is whether $y_{fjt}$ responds not only to local demand, as captured by $Y_{jt}$, but also to changes in the other markets in which $f$ operates.[34] I focus on two outcomes: the log of total payroll of firm $f$ in market $j$, and the log of average wages of firm $f$ in market $j$ (with their difference corresponding to the log of $f$'s employment in $j$). Importantly, the responsiveness of total payroll does not depend on the firm-level labor supply elasticity $\epsilon$, whereas the division of this response between wages and employment does. I define a region as a U.S. labor market area (LMA, a slight aggregation of commuting zones) and examine 5-year changes in these outcomes across Economic Census years. All specifications include firm-by-year fixed effects, and standard errors are clustered at the firm level. The analysis is restricted to service-producing sectors, excluding tradable-goods industries, thereby alleviating concerns that results may be confounded by input trade across branches or by output trade from region $j$ to consumers in other firm markets.

My baseline specification regresses 5-year changes in $\log y_{fjt}$ on 5-year changes in $\log Y_{jt}$ and on a weighted average of changes in $\log Y_{kt}$ across all regions $k$, where the weights reflect firm $f$'s ex-ante activity in each market $k$. Results are reported in Columns (1) and (4) of Table 3 for total payroll and average wages, respectively. As expected, both

[34] See Giroud and Rauh (2019) for an exercise of a similar spirit.

outcomes respond positively to local demand shocks in market $j$. Moreover, consistent with within-firm non-rivalries, both also respond positively to regional growth in $f$'s other locations. The magnitudes are in line with the assumption of firm-specific labor supply curves (finite $\epsilon$). Columns (2) and (5) replicate these regressions while flexibly controlling for region-by-time fixed effects for the focal region $j$ and sector-by-time fixed effects for the sector of firm $f$. In this specification, the coefficient on regional growth in $j$ is absorbed by the time trend, but the coefficient on growth in other markets remains positive and statistically significant for both total payroll and average wages. Controlling for this rich set of fixed effects rules out mechanisms related to contemporaneous regional developments in $j$ (including spillovers into this market from others) as well as sector-specific shocks to firm $f$.

A remaining concern is that firm $f$'s activity in market $j$ rises because of a firm-specific productivity shock (an increase in $z$ in the model) that is correlated with regional GDP growth in $f$'s other locations in ways not fully absorbed by sector-by-time fixed effects. For instance, if $f$ is large relative to some of its markets, a shock common to all of $f$'s locations could itself raise regional growth there; conversely, if the shock is specific to market $j$ and the expansion cost $C$ embeds a span-of-control constraint, activity might contract in other markets. In addition, in the model, demand shocks operate through both regional income and the regional price index, so using only income-based measures introduces a measurement error in the regressor. To address these issues, I implement an additional specification that instruments regional growth in all other markets with each region's exposure to national sectoral growth trends (Bartik, 1991). Because I also control for the sectoral trend relevant for firm $f$, identification comes from regional growth induced by national shocks in other sectors. Columns (3) and (6) report the IV estimates using this instrument. For both total payroll and average wages, the effects remain positive and statistically significant, with F-stats way above conventional levels. The IV point estimates exceed the OLS estimates, consistent with a bias due to firm-by-region productivity shocks in the presence of span-of-control constraints and/or attenuation from measurement error.[35]

As with the pass-through to the headquarters premium, the propagation of demand shocks across a firm's locations does not map to a single structural parameter. Nevertheless, in the model quantification below, I verify that the framework reproduces similar patterns.

**Taking stock.** While non-rivalry is traditionally a challenging object to measure, I provide three pieces of evidence that the output of headquarters workers is non-rival

[35]The large coefficient for total payroll – implying roughly one-for-one transmission from growth in $f$'s other markets to payroll growth in a typical market – is natural given that the typical market is very small relative to the firm. Consider for example the response of a national chain's payroll in a small market to a common demand increase across all its other markets.

**Table 3: Within-firm spillovers across markets**

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Outcome: 5-year changes firm $f$, market $j$: | $\Delta$ ln payroll | $\Delta$ ln payroll | $\Delta$ ln payroll | $\Delta$ ln wage | $\Delta$ ln wage | $\Delta$ ln wage |
| Regional income growth in market $j$ | 0.254*** | | | 0.0685*** | | |
| | (0.0131) | | | (0.00733) | | |
| Average income growth in all $f$'s markets | 0.395*** | 0.268*** | 1.024*** | 0.178*** | 0.134*** | 0.280*** |
| | (0.0234) | (0.0247) | (0.110) | (0.0133) | (0.0140) | (0.0667) |
| Observations (rounded) | 856000 | 856000 | 856000 | 856000 | 856000 | 856000 |
| First-stage F stat | . | . | 3283 | . | . | 3283 |
| Year and firm FE | Y | Y | Y | Y | Y | Y |
| Year × market FE | N | Y | Y | N | Y | Y |
| Year × sector FE | N | Y | Y | N | Y | Y |

Note: This table shows how firm outcomes in a particular local market $j$ respond to demand shocks in other markets. Column (1) shows the results from a regression of 5-year changes in the (log) total payroll of firm $f$ in market $j$ against regional GDP growth in market $j$ over the same time period; and against a weighted-average of regional GDP growth in all of firm $f$'s markets. Column (2) repeats Column (1) with time-varying fixed effects for the firm's sector and for market $j$, which also absorbs regional GDP growth in the focal market $j$. Column (3) repeats Column (2) by instrumenting regional GDP growth in all markets using ex-ante exposure to national sectoral trends, as described in the text. Columns (4)-(6) repeat Columns (1)-(3), replacing changes in (log) total payroll with (log) average payroll per employee as the outcome variable. All columns include time and firm fixed effects. Standard errors are clustered at the firm level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

across a firm's locations. First, headquarters jobs are strongly associated with non-rival task content as captured in task-based data. Second, when firms expand in response to firm-level demand shocks, they raise relative wages for headquarters jobs, consistent with marginal productivity in these roles depending on the number of firm locations. Third, local demand shocks in a subset of a firm's markets spill over into higher activity in its other markets, again consistent with within-firm non-rivalry.

## 4.2 Additional evidence

I now briefly discuss additional evidence for the model's predictions, leaving more details to Appendix E.4.

**Firm expansion and spatial disparities.** The model yields several key implications for the spatial distribution of economic activity. First, Proposition 4 predicts that markets experiencing greater expansion of locally headquartered firms should see a relative increase in expenditure on headquarters jobs within the services sector (a higher $\Gamma_i$ in the model). Second, such markets should exhibit increased specialization in services relative to tradable goods production (reflected in a higher $R_i/E_i$ in the model), implying a decline in the employment share of manufacturing and other goods-producing sectors. Third, since this shift amounts to a positive regional labor demand shock, both regional wages and employment should rise.

I examine these predictions in the data across U.S. commuting zones, focusing on

decadal changes over the period 1980-2017. To construct a region measure for the expansion of locally headquartered firms, I use Dun & Bradstreet data on firm geography.[36] Specifically, for each market and year, I compute the average number of commuting zones served by firms that were headquartered in that market in 1980. For instance, Walmart Corporation, headquartered in Benton County, AR in 1980, substantially expanded the number of its locations during 1980–2017, contributing significantly to Benton's average expansion measure.[37]

Figure 6 shows how key regional outcomes comove with this measure of spatial expansion for local firms. Panel (a) plots the log relative payroll share in headquarters-intensive occupations on the y-axis against the log number of commuting zones served by locally headquartered firms on the x-axis, controlling for year and regional fixed effects. The figure reveals a strong, positive, and statistically significant relationship: as locally headquartered firms expand into more markets, regional services payroll is increasingly devoted to headquarters tasks. Second, panel (b) replaces the y-axis with the log of the regional manufacturing employment share. The strong negative relationship reflects increased specialization in providing headquarters services to other regions relative to producing tradable goods. Third, panels (c) and (d) replace the y-axis with regional log average wages and regional log total employment, respectively. Since expansion of locally headquartered firms amounts to a positive local demand shock, both regional wages and employment rise.

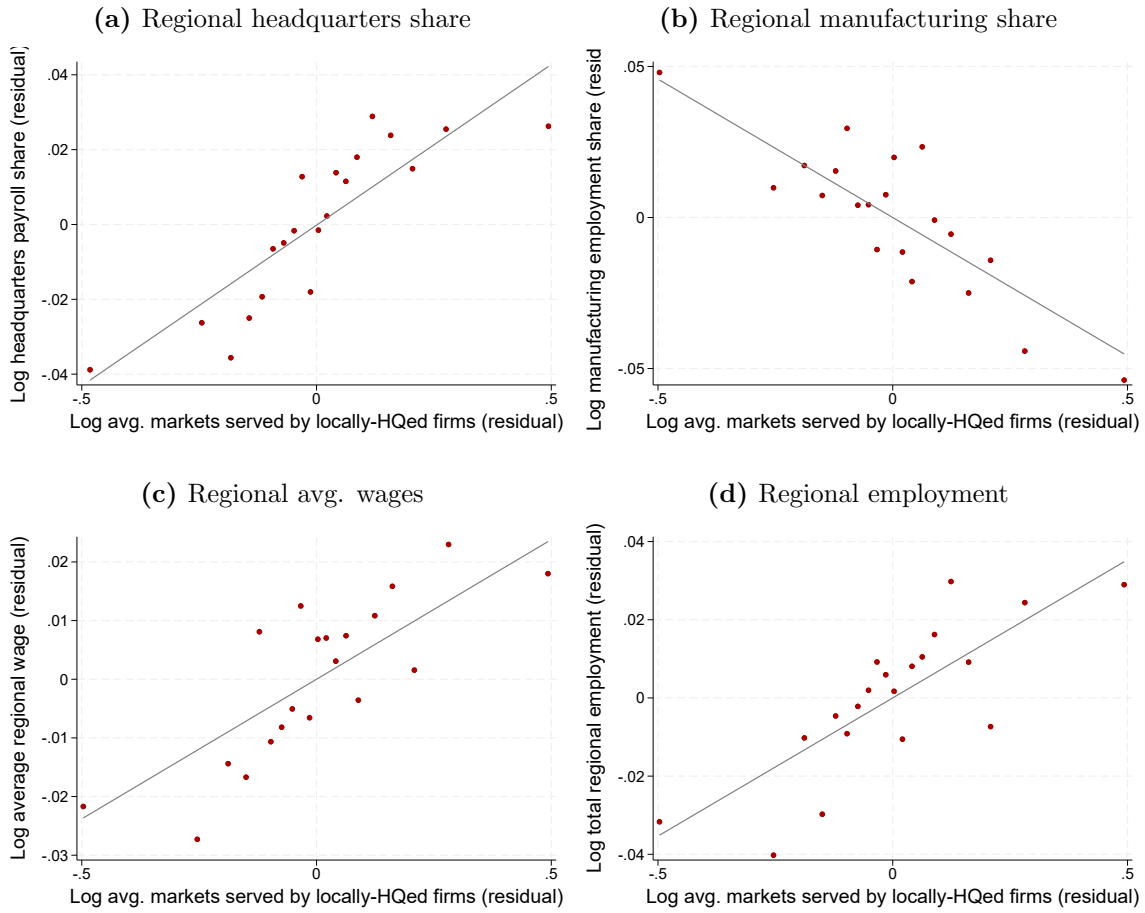
In Appendix E.4, I present the corresponding regression results and show that the correlations reflect a causal relationship. To establish this, I exploit regional exogenous variation in firm expansion arising from ex-ante exposure to national expansion trends, following the approach of Bartik (1991). The idea is that if a market such as Benton County, AR, was already specialized in hosting headquarters of productive retail firms in 1980, then the substantial nationwide spatial expansion of retail sector firms over 1980–2017 generated an "expansion shock" for Benton that was not driven by other local factors. Using this source of variation, I show that the results remain consistent with those in Figure 6. Finally, I demonstrate that the findings are robust to controlling for regional employment shares in the information and financial sectors, confirming they are not driven by regional exposure to the exceptional growth of these two sectors.

**Firm expansion and headquarters payroll share.** Another prediction of the model is that if $\eta \neq 1$, then the payroll share on headquarters tasks changes as firms expand in space. The job posting data I use to test the implications for headquarters wages is ill-suited to study this prediction since I observe job postings and not employment

[36] Recall from Section 2.1 that this source is particularly well suited for the exercise, as it reports the headquarters location of every firm in every year (unlike the Census LBD). Note that for this exercise, I use only information on firms' locations, and I do not use data such as employment or payroll which is lower quality in D&B.

[37] See Appendix A for more details on this specific example.

**Figure 6: Firm spatial expansion and local labor market outcomes**




Note: Binscatter plots of local labor market outcomes against the expansion of locally headquartered firms. Subplot (a) shows the log payroll share in headquarters-intensive occupations within the services sector (y-axis) against the log average number of markets served by locally headquartered firms (x-axis). See Section 4.2 for details on variable construction. Each observation is a commuting zone by year, covering the years 1980, 1990, 2000, 2010, and 2017, and commuting zones with at least 10,000 employees in 1980. All variables are residualized with respect to commuting zone and year fixed effects. Subplot (b) replaces the y-axis with the log regional employment share in manufacturing; subplot (c) with the log average regional wages; and subplot (d) with the log total regional employment. The black dashed line in all figures shows the linear fit. Regression results corresponding to these relationships appear in Appendix E.4.

quantity. However, I can study this relationship using the Census LBD data, leveraging the alternative measurement of headquarters activity using the variation in establishments' industry classifications within firms (see Section 2.1 for details). I investigate this in Appendix E.4, in which I study how the ratio between total payroll to headquarters establishments and total payroll to branch-level establishments changes when the number of firm establishments increases. The results show that controlling for firm, year, and sector-by-year fixed effects, doubling the number of firm establishments is associated with a 15% increase in this payroll ratio, consistent with the predictions of the model for $\eta > 1$.

**Long-run occupational and industry dynamics.** Up until now I have studied firm-level and region-level evidence. The model also predicts that given the substantial firm spatial expansion in the data, we should observe rising relative wages of headquarters occupations at the aggregate level in the long run. When headquarters and branch tasks are substitutes ($\eta > 1$), it further predicts an increase in the payroll share of headquarters-intensive occupation at the aggregate level in the long run. These effects should hold within sectors, not just across them, and be especially pronounced in industries that experienced greater spatial expansion. I test these predictions in Appendix Section E.4, using data from the Decennial Census and the American Community Survey to study changes between 1980 and 2017. The findings support the above predictions.

# 5 Model Quantification

I now turn to estimate the model. Using the quantitative model, I demonstrate the importance of the above mechanisms for the overall increase in U.S. wage inequality, and evaluate policies that shape firms' ability to span multiple local markets. I match key features of the U.S. economy in 1980, and study counterfactuals that reflect changes in the economic environment between 1980-2017. I map the model to 200 local labor markets in the U.S., which I construct by combining small commuting zones with neighboring large ones.

## 5.1 Additional model components

First, I enrich the model with additional components that help to discipline the baseline equilibrium and which are common in spatial equilibrium models.

**Multiple skill groups in production.** I now assume that households belong to $S$ different skill groups. The headquarters labor bundle $h_i(z)$ and the branch-level bundle $\ell_{ij}(z)$ are now Cobb-Douglas aggregators of all skill groups with skill intensities $\alpha_{hs}$ and $\alpha_{\ell s}$, respectively, such that $\sum_{s=1}^{S} \alpha_{hs} = \sum_{s=1}^{S} \alpha_{\ell s} = 1$. Similarly, the tradables sector also uses a Cobb-Douglas production function with skill-intensities $\alpha_{gs}$, where $\sum_{s=1}^{S} \alpha_{gs} = 1$.

**Regional amenities.** Agents of type $s$ choosing location $i$ now enjoy an exoge-

nous amenity component $B_{is}$ that enters multiplicatively in their utility function. Consequently, the probability that an agent of type $s$ chooses location $i$ is given by

$$\frac{L_{is}}{\sum_{j=1}^{N} L_{js}} = \frac{\left(B_{is}\bar{W}_{is}/\bar{P}_i\right)^{\xi}}{\sum_{j=1}^{N}\left(B_{js}\bar{W}_{js}/\bar{P}_j\right)^{\xi}}, \tag{10}$$

where $\bar{P}_i$ is the price index of the local final good, and $\bar{W}_{is} \equiv L_{is}^{\frac{1}{\epsilon}} W_{is}$ is the ideal wage index for type $s$ workers in location $i$. Households are more likely to choose locations with higher effective real wages $\bar{W}_{js}/\bar{P}_j$ and higher amenities $B_{js}$.

**Housing.** I add local housing, which introduces a second dispersion force on top of the idiosyncratic preference shocks. Agents spend a share $\zeta$ of their income on housing and a share $1-\zeta$ on the composite of goods and services. I assume a constant housing supply elasticity $\varrho$ and that the rights to housing rents are owned by immobile absentee landlords who have a similar structure of preferences for consumption as workers. The aggregate price index, previously given by $P_i^{\beta}$, is now equal to $\bar{P}_i = \bar{\beta} P_i^{\beta\left(1-\frac{\zeta}{(1+\varrho)-\zeta\varrho}\right)} Y_i^{\frac{\zeta}{(1+\varrho)-\zeta\varrho}}$, where $Y_i$ is total regional income and $\bar{\beta}$ is a constant that subsumes various parameters. Conditional on the price index of services $P_i$, local cost of living is increasing in regional income $Y_i$ due to higher cost of housing. The exponent on regional income is higher when agents spend more on housing (higher $\zeta$) or when housing supply is relatively inelastic (lower $\varrho$).

**Regional productivity in tradables.** I allow locations to differ in their total factor productivity in the production of tradable goods, given by $A_{i,g}$.

**Firm productivity distribution.** In line with the literature, I parameterize the firm productivity distribution $G(z)$ as a Pareto distribution, with a shape parameter $\theta$.

## 5.2 Parameterizing the cost of expansion

I consider a particular structure for market penetration and the cost of expansion $C(\cdot)$ for easy mapping to the data. I assume that market penetration is an increasing function of the number of establishments that a firm opens in a given market, denoted by $n_{ij}(z)$: $x_{ij}(z) = \frac{1}{1+n_{ij}(z)^{-1}}$. Market penetration is zero when $n_{ij}(z) = 0$, and approaches 1 when $n_{ij}(z) \to \infty$. The concavity of $x_{ij}$ in $n_{ij}$ reflects potential cannibalization effects across the firm's establishments in a given market: the marginal establishment attracts fewer consumers than the average of existing establishments.[38]

The cost of expansion $C(\cdot)$ captures the cost of setting up establishments $n_{ij}(z)$ in all

[38] See Arkolakis (2010), Oberfield et al. (2024), and Wenning (2024) for similar specifications.

markets, and is given by

$$C\left(\{n_{ij}(z)\}_{j=1}^{N}\right) = \sum_{j=1}^{N} \tau_{ij} n_{ij}(z) + \delta \left(\sum_{j=1}^{N} n_{ij}(z)\right)^2.$$

The first term in the cost of expansion captures a constant cost per each establishment: a firm based in location $i$ with presence in location $j$ pays a cost of $\tau_{ij}$ for every establishment that it opens there. The second term allows for a potential span-of-control constraint (SoC) for the firm as a whole, reflected in a quadratic cost in the total number of establishments in all markets. The importance of this SoC effect is captured by the parameter $\delta$. Note that this SoC effect works in the opposite direction to the non-rivalry assumption in the production function: if $\delta$ is high and the productivity of headquarters inputs ($\gamma$) is low, then expansion in one market (e.g. following a positive demand shock there) could lead to contraction in other markets.

With an eye toward my main quantitative exercise, I assume a simple parameterization of the bilateral cost per-establishment $\tau_{ij}$, given by

$$\tau_{ij} = \tau_0 \times \text{dist}_{ij}^{\rho} \times \kappa^{\mathbb{I}\{\text{st}_i \neq \text{st}_j\}},$$

where $\tau_0$ is a scaling constant; $\text{dist}_{ij}$ is the geographical distance between $i$ and $j$; $\rho$ is the distance-elasticity in the cost of opening new establishments; and $\kappa$ is a cost premium when $i$ and $j$ are in different states. The cost of setting up a new establishment is then higher in markets that are far away from the firm's headquarters.

The motivation for this specification is the strong gravity pattern in bilateral firm linkages, detailed in Appendix E.5. As shown in Table 16, the number of headquarters-branch relationships between market pairs is well described by a gravity equation with headquarters-location fixed effects, branch-location fixed effects, and distance. The estimated distance elasticity of headquarters-branch linkages is about -1 and has declined over time – a pattern I return to in the counterfactual analysis below.

In addition to distance, I allow for a higher cost per-establishment $\kappa > 1$ when operating establishments in a different state than the firm's headquarters. Indeed, in the data, cross-region firm linkages are lower when operating establishments in different state from the firm's headquarters, as can also be seen in Appendix Table 16. This state border effect could arise due to the role of state-level regulations and compliance requirements. An extreme example of such regulations is the restriction on cross-state banking activity that characterized the U.S. banking system before a vast deregulation in the 1980s-1990s. In such an extreme case, $\kappa$ would be set to infinity.

It is worth noting that I allow for spatial frictions in the cost of expansion $C(\cdot)$, but not in the firm's production function, as often assumed in the literature. The reason for this assumption is that while gravity is a very strong feature of firms' location decisions, it

is not clearly evident in firm regional wages. Thus, in models with firm-specific wages, one would like to avoid introducing spatial frictions that distort the marginal product of labor across the firm's branches, since it would create a counterfactual systematic relationship between wages and distance.

## 5.3 Calibration of the households block

I divide the model's parameters into three groups. The first group includes parameters that govern household preferences and the housing market. This group consists of commonly-used building blocks from the literature and I rely on existing estimates to calibrate it. The second group includes regional productivity and amenity fundamentals, which I invert from the model's equilibrium conditions. The final group includes parameters from the production block, in particular the services production function and the cost of expansion. This block includes new features relative to the existing literature, and so I estimate it using Simulated Method of Moments (henceforth SMM).

I begin with the external calibration of the household block. I calibrate the expenditure share on services ($\beta$) to match the share of services value-added in national accounts. I set the dispersion of idiosyncratic shocks across regions ($\xi$) to 2.8, in line with the range of values in the trade and spatial literature, e.g. Galle et al. (2023). I set the dispersion of idiosyncratic shocks across employers ($\epsilon$) to 5.0, matching recent estimates of the average wage markdown in Lamadon et al. (2022), Berger et al. (2022) and Azar et al. (2022), which is given in the model by $\epsilon/(1+\epsilon)$. The elasticity of substitution across varieties $\sigma$ is set to 5.0, to match a price markup of 25% over marginal cost. This is also in line with the range of existing estimates on substitution between varieties in the trade literature, e.g. Costinot and Rodríguez-Clare (2014). The expenditure share on housing is set to 0.24, following Davis and Ortalo-Magné (2011). Finally, I set the local housing supply elasticity to the population-weighted estimate of 1.75 in Saiz (2010). Table 4 summarizes the calibration of this part of the model.

**Table 4:** Calibration of the households block of the model

| Parameter | Interpretation | Source | Value |
|---|---|---|---|
| $\xi$ | Dispersion of location preference shocks | Galle et al. (2023) | 2.8 |
| $\epsilon$ | Dispersion of employer preference shocks | Lamadon et al. (2022), Berger et al. (2022) | 5.0 |
| $\sigma$ | EoS between varieties | Costinot and Rodríguez-Clare (2014) | 5.0 |
| $\beta_k$ | Sectoral expenditure shares | Direct computation - BEA NIPA (1980) | [0.64,0.36] |
| $\delta$ | Housing expenditure shares | Davis and Ortalo-Magne (2011) | 0.24 |
| $\varrho$ | Housing supply elasticity | Saiz (2010) | 1.75 |

I define four skill groups ($S = 4$) in the quantified model, and compute the skill elasticities in production $\alpha_{hs}, \alpha_{\ell s}, \alpha_{gs}$ directly from the data, by computing the payroll-

share of each group in all three activities. For additional details, see Appendix E.5.

## 5.4 Model inversion

The set of location fundamentals, $\mathcal{A} \equiv A_i, B_{is}, A_{g,i}i, s$, comprises service-sector productivity shifters, regional amenities, and tradable-goods productivity. I recover $\mathcal{A}$ by iterating on the model's equilibrium conditions, given a guess for the parameters. The inversion matches three empirical moments: total regional labor income ($Y_i$), employment by skill group ($Lis$), and the relative payroll of skilled to unskilled workers (the ratio of total payroll for the top two skill groups to that of the bottom two). Intuitively, $B_{is}$ follows from inverting workers' location-choice, while $A_i$ and $A_{g,i}$ are pinned down jointly from regional income and skill-intensity. Thus, in the baseline equilibrium, I match employment, income, and skill-intensity across all 200 regions, replicating key features of the U.S. economy. Appendix C provides further details.

## 5.5 Simulated method of moments for the production block

The remaining parameters to estimate are from the production block of the model: the relative productivity of headquarters inputs ($\gamma$); the elasticity of substitution between headquarters and branch labor ($\eta$); the dispersion of firm productivity draws ($\theta$); the distance-elasticity of the cost of expansion ($\rho$); the state-border effect in the cost of expansion ($\kappa$); the scale of the cost of new establishments ($\tau_0$); the span of control component of the cost function ($\delta$); and the entry cost ($f_e$). I collect these parameters in the vector $\Theta$, and estimate them by matching a set of empirical moments. Below I provide brief intuition for which empirical moments helps to identify each of these parameters, though recall that in practice all moments are jointly determined in equilibrium by all of these parameters.

The parameter $\eta$ is directly related to how the expenditure-share on headquarters varies with firm scope, as given in Proposition 1. In particular, when $\eta = 1$, this expenditure share does not commove with firm scope. I therefore target the estimate for this co-movement Section 4.2. The productivity shifter of headquarters inputs ($\gamma$) affects directly the level of the expenditure share on headquarters workers, and thus it shapes the amount of within-firm inequality. I therefore use it to discipline the importance of within-firm inequality in the baseline equilibrium, targeting the within-firm component of the overall variance of log wages across establishments in the economy (as measured in 1980).[39] These two parameters ($\eta$ and $\gamma$) also affect the strength of within-firm nonrivalries in the economy, as manifested in the response of the headquarters premium to

[39] Within-firm inequality is easier to measure than the expenditure share on headquarters workers (see discussion in Section 4.2), and especially how it changed over time. Nevertheless, I show that the model also yields reasonable values for the expenditure share on headquarters workers as an over-identification check below.

firm expansion, and in the within-firm spillover patterns from Section 4. I do not target these moments explicitly due to the associated computational burden,[40] but I later verify that the model indeed yields a similar patterns to the estimates in Section 4 as an over-identification check. To match the between-firm component of inequality, I target the shape parameter of the productivity distribution, $\theta$, which determines the magnitude of differences between firms. Note that both a higher $\gamma$ and a lower $\theta$ raise inequality between firms in the economy, but they have opposing effects on inequality within firms.

The other parameters are those in the cost of expansion $C$ and the entry cost. For the distance elasticity in the cost of expansion, $\rho$, I target the distance elasticity of headquarters-branch linkages from Table 16, by running the gravity regression in the simulated model. For the state-border effect, $\kappa$, I target the state-border effect from the same gravity regression. For the scale of the cost of new establishments, $\tau_0$, I target the average number of establishments per firm. For the span of control component of the cost function, $\delta$, I target the employment-weighted number of establishments per firm. Finally, for the entry cost $f_e$, I target average firm size, i.e. the ratio of total employment (normalized to 1 in the model) to the total mass of firms.

To estimate $\Theta$, I minimize the loss function $\mathcal{L}(\Theta) \equiv (m(\Theta) - \tilde{m})' \boldsymbol{W} (m(\Theta) - \tilde{m})$ where $m(\Theta)$ is the vector of simulated moments from the model; $\tilde{m}$ are the equivalent moments for 1980 in the data; and $\boldsymbol{W}$ is a weighting-matrix, which I set to be diagonal and inversely proportional to the squared values of $\tilde{m}$, expressing the moments in percentage-deviation terms.[41] Results from this estimation procedure are provided in Table 5.

**Table 5:** Estimated parameters for the production block of the model

| Parameter | Interpretation | Estimate | Main targeted moment | Value (model/data) |
|---|---|---|---|---|
| Production function and productivity distribution | | | | |
| $\eta$ | EoS between headquarters and branches | 1.28 | Co-movement of relative expenditure on HQ and firm scope (Table 13) | 0.15 / 0.15 |
| $\gamma$ | Productivity shifter for HQ workers | 0.19 | Within-firm component of (within-sector) variance of log wages across establishments (Census LBD) | 0.11 / 0.12 |
| $\theta$ | Dispersion of firm productivity | 7.89 | Between-firm component of (within-sector) variance of log wages across establishments (Census LBD) | 0.12 / 0.12 |
| Cost of expansion ($C$) and entry cost | | | | |
| $\rho$ | Distance elasticity in cost of new establishments | 0.76 | Distance elasticity of HQ-branch linkages (Table 16) | -0.95 / -0.98 |
| $\kappa$ | State border effect in cost of new establishments | 2.50 | State border effect in HQ-branch linkages (Table 16) | -1.30 / -1.28 |
| $\tau_0$ | Shifter of cost of new establishments | 2.15 | Average number of establishments per firm (Census BDS) | 1.19 / 1.22 |
| $\delta$ | Shifter of span of control in the cost of expansion | 5.12 | Average number of establishments per firm (weighted, Census BDS) | 15.4 / 15.2 |
| $f_e$ | Entry cost | $10^6$ | Firm size (employment per firm, Census BDS) | 26.5 / 24.5 |

## 5.6 Over-identification

As one assessment of the model's performance, I now briefly report how the quantified model performs for moments that were not explicitly targeted in the estimation procedure.

**Passthrough of firm growth to wages.** As discussed in Section 4, one key implication of the non-rivalry assumption is differential pass-through of firm expansion

[40] It requires an additional computation of the equilibrium for every guess of $\Theta$, in addition to the model inversion that I perform for each guess.

[41] In practice, I employ the TikTak algorithm for global optimization from Arnoud et al. (2019) with 500 starting points, setting the Nelder–Mead method as the local minimizer.

to wages. In particular, I revisit in the model how an increase in firm size as a result of simulated demand shocks (a change in $\Upsilon_j$) affects firms' relative headquarters wages in the model. Replicating the fixed-effect regression from Table 2, I obtain an estimate of approximately 0.175, slightly above the 2SLS point estimate and well within its 95% confidence interval. As one benchmark to assess this magnitude, recall that in the case of symmetric space and $\eta = 1$, this elasticity is given by $1/(1+\epsilon) \approx 0.166$ in the model. With asymmetric space and $\eta \neq 1$, we obtain a slightly higher value. The model thus replicates well the differential pass-through of firm expansion to wages that is implied by the non-rivalry assumption.

**Within-firm spillovers.** Another implication of the non-rivalry assumption is the existence of cross-region spillovers within firms. I measured such spillovers in Table 3. Replicating the regression that captures how a firm's expansion in a given market responds to exogenous income growth in the firm's other markets within the model, I find a coefficient of around 0.8. This is slightly lower than the 2SLS point-estimate from Table 3, but within its 95% confidence interval. The model thus replicates well the cross-region spillovers within firms that are implied by the non-rivalry assumption.

**Headquarters expenditure share.** Due to the challenges of measuring firms' headquarters expenditure share in the data (see discussion in Section 4), I do not target its level explicitly, and prefer to target the (more accurate) variance of log wages within firms. Nevertheless, we can compare the output of the model to different approaches in the data. Consider the approach of measuring headquarters payroll share using establishments' industry classifications in the LBD, as described in Section 2.1. According to this approach, the expenditure share on headquarters workers out of total payroll to headquarters and branches is around 16%. A potential upper bound is to treat only establishments in the firm's main NAICS-4 industry as "branches", which yields an headquarters income share of 23%; a potential lower bound is to treat only establishments in the NAICS 55 sector (which are explicitly reported as headquarters activity) as "headquarters", and all other establishments (including other within-firm business services) as "branches", which yields 8%. In the quantified model, the payroll-weighted average of this statistic across firms is 14%, in-between the upper and lower bounds, and very close to my preferred estimate.[42]

**Spatial distribution of headquarters.** While the allocation of population, wages, and skills was explicitly targeted as part of the model's inversion, I did not use data on the spatial distribution of firms and headquarters. The correlation of the log measure of headquarters between the model and the data is above 0.9. This high correlation results from the fact that this measure scales very strongly with total regional employment. In

[42] An alternative approach to measure this moment is the data is to use the notion of SG&A costs (Selling, General, and Administrative Expenses) for publicly-listed firms. This measure is often used to proxy for headquarters and other overhead costs. Depending on the specific data restrictions, the share of these costs in total costs is around 15%-20%, again very close to the share in the model.

the model, the tight link to regional employment results from the combination of free entry and high spatial frictions for headquarters-branch linkages.

# 6 Quantitative experiments

In this section, I use the quantitative model to evaluate the aggregate and distributional implications of firms' spatial expansion. I focus on a particular set of shocks to the baseline equilibrium: changes in spatial frictions to firm expansion, i.e. on the geographical aspects of the cost of expansion $C(\cdot)$ – the distance elasticity $\rho$ and the state-border effect $\kappa$. There are a few key advantages in focusing on these shocks when studying the importance of firm expansion. First, since they are part of the cost of expansion $C(\cdot)$, all their impacts on equilibrium, labor markets, and welfare are materialized through firms' market penetration decisions $x_{ij}(z)$, and thus through the novel mechanisms outlined in this paper. These shocks have no effect when holding constant expansion decisions, in contrast to other primitives such as shocks to demand, firms' production function, or regional fundamentals. Second, while many primitives in the model can induce an increase in firms' geographical scope, we can easily discipline changes in these geographical frictions from data on bilateral headquarters-branch linkages. Moreover, as detailed below, I use only information on changes in the number branches in distant markets *relative* to branches in firms' headquarters market, thus abstracting from any expansion patterns that are unrelated to spatial frictions. Lastly, these shocks have a clear interpretation of lower frictions to operate far-away establishments. Thus, they capture specific technological and policy changes that shape firms' communication frictions across space, such as improvements in information technologies, the expansion of air-travel, and regulatory changes.

## 6.1 Lower frictions to operate far-away branches

The main shock that I explore is a realistic decline in spatial frictions to firm expansion, driven by an economy-wide decline in the distance elasticity $\rho$. This shock is motivated by a sharp decline in the distance-elasticity for headquarters-branch linkages in the data, as can be seen in Appendix Table 16. Between 1980 and 2017, this distance elasticity has declined (in absolute value) by around 0.2-0.4. I thus change $\rho$ in the model to achieve a decline of around 0.3 in this distance elasticity, holding other primitives constant. Importantly, I use only data on relative distances between headquarters and branches – and no labor market data – to discipline this shock.[43]

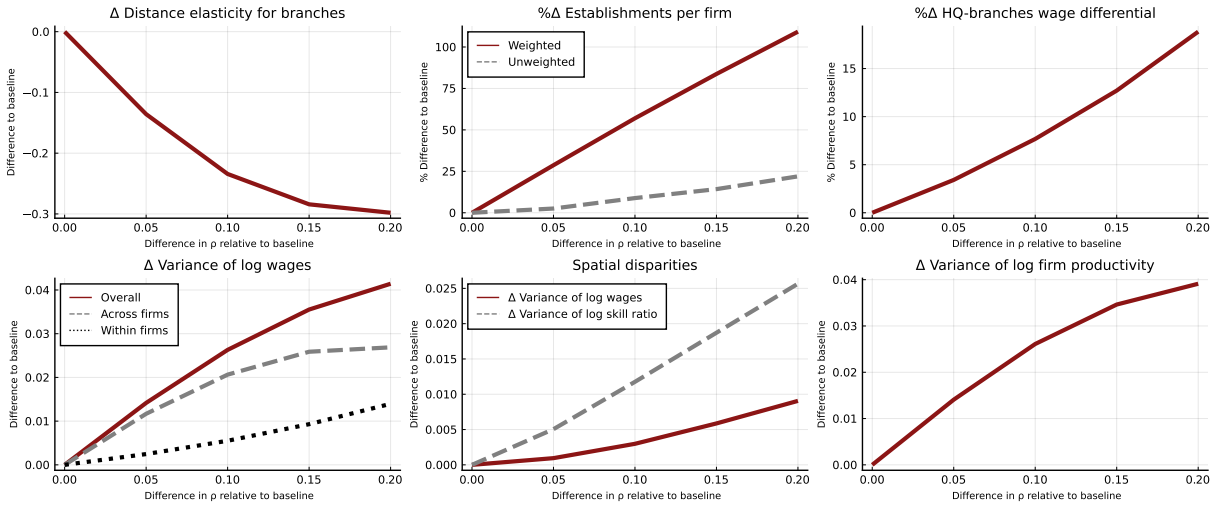
[43]Note that while the model does not admit an exact log-linear gravity equation, I verify that I replicate the change in the empirical distance elasticity by running the exact same gravity regression in the simulated model. The $R^2$ coefficient from this gravity regression in the model is over 0.9, suggesting that it approximates very closely the log-linear specification that I consider in the data.

Figure 7 displays changes in key aggregate moments of interest in response to this decline in the distance elasticity $\rho$, where I vary the values of $\rho$ relative to the baseline equilibrium on the x-axis, ranging from zero (the baseline equilibrium) to a difference of 0.2 relative to the baseline 1980 equilibrium. I then re-compute the equilibrium for every new value of $\rho$, and plot changes in selected outcomes relative to the baseline equilibrium on the y-axis. By construction, the gravity coefficient of distance (top-left subplot) declines by a similar magnitude to its decline in the data. Since this decline reflects a lower cost of expansion $C(\cdot)$, firms increase their average number of establishments (top-middle subplot). In line with the empirical evidence from Figure 1, this increase is driven by the large firms, such that the employment-weighted expansion is far greater than the unweighted expansion. Overall, the weighted expansion in the model is above 100%, over half of the equivalent increase in the data between 1980-2017.

Turning to labor market outcomes, the above spatial expansion leads to higher economy-wide headquarters-branch wage dispersion (top-right subplot), as in Proposition 1. In the model, the decline in distance elasticity raises this difference by close to 20% . In the data, this moment varies with the exact definition of headquarters and branches, but employing the definitions in Section 2 reveals an increase of about 30%, such that the distance-elasticity shock in the model accounts for around two-thirds of the trend in the data.

When investigating broader measures of inequality, the overall variance in log wages (bottom-left panel) rises by around 0.04 points relative to the baseline equilibrium. Compared to an increase of around 0.09 in the (within-industry) variance of log wages in the data, the decline in spatial frictions to expansion can rationalize above 40% of the observed trend. When considering the overall increase in the variance of log wages (and not just within-industry), the above shock can rationalize close to a quarter of it, though, recall that the model cannot speak to the cross-industry component which is not modeled. Moreover, the model gives rise to higher inequality both across firms and within them, in line with the patterns described in Section 2, and in contrast to theories that highlight only greater differentiation between firms (such as rising outsourcing). To see why inequality across firms rises in the model, recall that the model allows for endogenous changes in the dispersion of firm productivity – defined as sales-per-worker – due to the existence of within-firm non-rivalries. Indeed, in the quantitative model, the dispersion in firm productivity rises following a reduction in the cost of expansion, as depicted in the bottom-right panel. In principle, greater firm scope might also lead to lower dispersion in productivity and wages across firms, if expansion is concentrated in smaller firms. This is not the case in the quantified model or in the data, but it introduces concavity in the relationship between the cost of expansion and dispersion between firms.

**Figure 7: Labor market implications of lower spatial frictions to firm expansion**



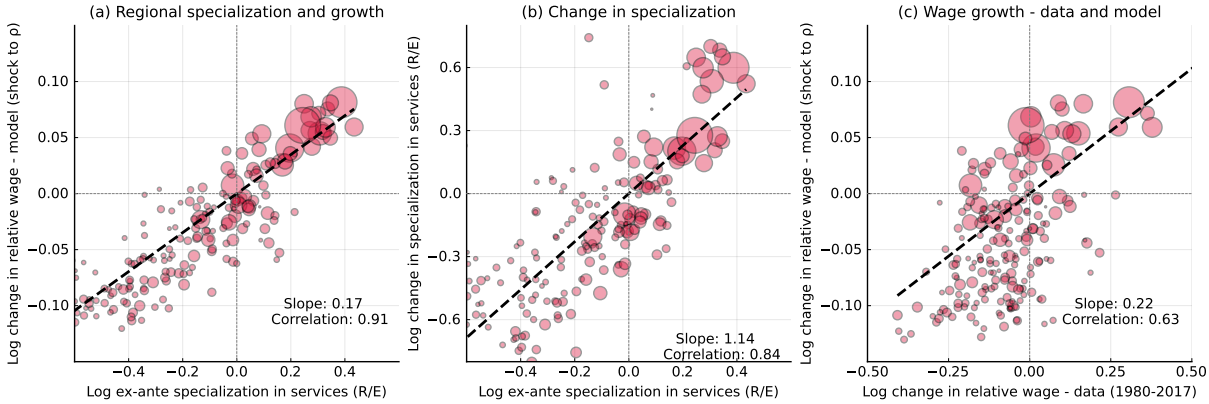

Note: Distributional implications of lower spatial frictions to firm expansion. The x-axis depicts changes in the distance elasticity in the cost of expansion, $\rho$, expressed in terms of the difference from the baseline equilibrium (with zero representing the baseline equilibrium). Positive values indicate lower distance elasticity in absolute value. The y-axis in each subplot captures changes in selected moments of interest relative to the baseline equilibrium, after re-computing the equilibrium for every new value of $\rho$: change in the gravity coefficient of headquarters-branch linkages (top-left panel); % change in average establishments per firm (employment-weighted and unweighted, top-middle); % change in mean wage dispersion between headquarters and branches (top-right); change in variance of log wages (overall, across firms, and within firms; bottom-left); change in the variance of log wages and the variance of high-skilled to low-skilled employment (bottom-middle); change in the variance of log firm productivity (bottom-right).

Finally, spatial disparities also increase, in line with Proposition 4. The bottom-middle panel of Figure 7 depicts changes in spatial inequality, as measured by the variance of log wages across regions (solid red line); and changes in spatial segregation, as measured by the variance of log skill-to-unskilled employment ratio across space (dashed gray line). Both rise when the distance elasticity of expansion falls, by a magnitude that equals around a third of the rise in spatial inequality in the data.

To understand which regions gain the most from lower barriers to firm expansion, panel (a) in Figure 8 depicts regional wage growth following the above shock to $\rho$ against a measure of ex-ante regional specialization in providing headquarters services. As in Proposition 4, I measure such specialization using the ratio of total national sales of locally-headquartered firms to regional expenditure on locally-active branches, $R_i/E_i$. A few patterns emerge. First, specialization in headquarters services (relative to production of tradables) is a strong predictor of regional wage growth following a decline in the cost of expansion, since the markets that benefit the most are those with comparative advantage in providing headquarters services to other regions. Note that this comparative advantage can arise in the model either directly through high regional productivity for local headquarters (higher $A_i$); or indirectly through greater market access (low outgoing bilateral frictions $\tau_{ij}$) and/or high abundance of relatively skilled labor (e.g. due to high skill-specific amenities $B_{is}$). Second, it is the relatively larger markets that specialize

in such activity, and consequently that gain the most from a national increase in firms' scope. This aligns with the "urban-bias" of regional growth in recent decades.

**Figure 8: Regional implications of lower spatial frictions to firm expansion**




Note: Panel (a) shows regional wage growth in the model following a reduction of the distance elasticity $\rho$ against a measure of regional specialization in headquarters services in the baseline equilibrium. See Section 6 for additional details. Panel (b) shows changes in this measure of specialization following a shock to $\rho$ against its value in the baseline equilibrium. Panel (c) shows regional wage growth in the model following a reduction of the distance elasticity $\rho$ against regional wage growth in the data between 1980 and 2017. In all figures, each circle is one of 200 U.S. local labor markets as defined in Section 5, weighted by total employment in 1980. Dashed black lines capture employment-weighted linear fits. All values are demeaned by the employment-weighted average of each measure.

One key channel through which lower barriers to firm expansion lead to higher spatial disparities is the magnification of the above regional specialization in headquarters services, as summarized in Proposition 4. Indeed, in the quantified model, a reduction in the cost of expansion leads to greater variation in the specialization measure $R_i/E_i$. This can be seen in Panel (b) of Figure 8, which plots log changes in $R_i/E_i$ across local markets following a decline in $\rho$ against the ex-ante values of this measure in the baseline equilibrium. Large markets with high ex-ante specialization in headquarters services such as New-York City become even more specialized in such activities when firms can open branches in more markets. In return, these regions increasingly import tradable goods from smaller markets, which become increasingly specialized in the production of goods. This aligns with the empirical pattern that regions with greater expansion of locally-headquartered firms are also the ones that see a greater decline in their manufacturing activity, as discussed in Section 4.

Finally, a natural question is how do the above regional manifestations compare to actual patterns of regional growth in the data. Panel (c) of Figure 8 plots log changes in regional average wages in the model, following a shock to $\rho$, against these changes in the data between 1980-2017 (both demeaned to capture relative changes in wages).[44]

[44]To improve visual clarity, the figure omits one notable outlier: the local labor market of San Jose, CA, which includes Silicon Valley. Wage growth in this market has been exceptionally high in the data since the 1980s. While it is also positive and high in the model, it does not constitute a particular outlier. This market is included in the computation of the regression slope and the correlation.

Evidently, there is a strong positive relationship between the model and the data. The scale of changes in the model is smaller than the data, suggesting that a lower cost of firm expansion cannot be the only explanation for the observed patterns. Indeed, as noted above, this shock can explain around a third of the rise in variance of log wages across space that was observed in the data. At the same time, the prediction that larger markets that specialize in headquarters activities should see the greater increase in wages aligns with the experience of the U.S. economy in recent decades. It is also worth recalling that the above experiment considers just a homogeneous increase in the ability to operate far-away branches across all markets. Thus, the model rationalizes much of the regional experiences in the U.S. economy without delving into region-specific shocks and histories.

## 6.2 Deregulation of cross-state firm activity

Finally, I consider a counterfactual that lowers the cross-state frictions in firms' spatial expansion ($\kappa$), holding the importance of distance ($\rho$) and other primitives constant. This shock is motivated by the decline in the importance of state-border effects for cross-region headquarters-branch linkages, as can be seen in Table 16. Between 1980 and 2017, this state-border effect has declined (in absolute value) by approximately 0.1-0.4, depending on the chosen estimator. One natural interpretation for the decline in state border effects are changes in regulations and compliance requirements across U.S. states. Over the 1980s and 1990s multiple service sectors have undergone substantial deregulation that increased the ability of firms to operate cross-state operations, including in Financial Services (e.g. the Riegle–Neal Interstate Banking and Branching Efficiency Act of 1994, alongside many state-level changes), Telecommunications (The Telecommunications Act of 1996), Air Travel (The Airline Deregulation Act of 1978), Trucking (Motor Carrier Act of 1980), and Electricity (Energy Policy Act of 1992).

I study such changes in the model by shocking $\kappa$ until the state border effect in the model declines by 0.2. A replication of Figure 7 for this experiment is provided in Appendix E.5. The effect on firm expansion is much smaller relative to the decline in the distance elasticity $\rho$, raising the (employment-weighted) average number of establishments per firm by less than 10%. This suggests that it is hard to rationalize the vast expansion in the data using the wave of deregulation in the 1980s and the 1990s. Accordingly, the effect on inequality is also minor, with the variance of log wages rising by approximately 0.005. Average welfare increases by 0.8%, capturing gains from more variety in the economy (as each firm can provide its product in more markets); increased productivity, due to the non-rival nature of headquarters inputs (rising by 0.2% for the average firm); and increased regional specialization. Of course, such an experiment cannot capture the full welfare implications of the above deregulations, which are often associated with specific industry-level distortions. Nevertheless, it provides a sense of the equity-efficiency tradeoff

that arises from policies that shape firms' expansion opportunities for the typical sector, on top of the role of sector-specific distortions. Note also that the above manifestations are consistent with the empirical findings in Philippon and Reshef (2012), in which the deregulation of the U.S. financial industry is associated with an increase in the skill-premium and top executive compensation in that sector.

# 7 Conclusion

A key trend in the U.S. economy in recent decades is the increase in firms' spatial scope, particularly in services-producing sectors. This paper offers new theory and evidence that link this trend to changes in the distribution of wages, motivated by the central role of multi-location service firms in the rise of wage inequality. The key idea is that when the output of some workers is non-rival across the firm's locations, changes in firms' scope can generate rich distributional implications. I develop a model that formalizes this idea, yielding new micro-foundations for skill-biased technical-change, rising spatial disparities, and a simultaneous increase in wage dispersion across and within firms – all of which are observed in the data in recent decades. I provide reduced-form evidence for these predictions and for the assumption of within-firm non-rivalries, and quantify their aggregate relevance by estimating the model for the U.S. economy. I show that a reduction in spatial frictions to firm expansion can rationalize much of the observed labor market trends in the U.S. economy since the 1980s. I also demonstrate how this framework can be used to evaluate the implications of policies that shape firms' ability to span multiple markets, such as a deregulation of cross-state firm activity.

I highlight four main benefits of the framework for the broader question of wage inequality. First, compared to other theories of rising inequality such as foreign trade or automation, it is especially well suited to the services sector, which accounts for most of the economy and much of the increase in inequality. Second, it can simultaneously rationalize multiple labor market trends that have largely been studied in isolation. Third, it aligns with observed changes in the organization of production and the expansion of multi-establishment firms. Finally, it generates distinctive predictions absent from existing theories of wage inequality – such as the importance of non-rivalry in explaining occupational inequality and the rising dispersion across establishments within firms. More broadly, the framework can also be applied beyond wage inequality, for example to analyze the shift toward services, the globalization of services firms, the rise of work-from-home, and the design of place-based policies and local competition for firms.

# References


**Aghion, Philippe, Antonin Bergeaud, Timo Boppart, Peter J Klenow, and Huiyu Li**, "A theory of falling growth and rising rents," *Review of Economic Studies*, 2023, *90* (6), 2675–2702.

**Alviarez, Vanessa, Javier Cravino, and Natalia Ramondo**, "Firm-embedded productivity and cross-country income differences," *Journal of Political Economy*, 2023, *131* (9), 2289–2327.

**Amiti, Mary and Donald R Davis**, "Trade, firms, and wages: Theory and evidence," *The Review of economic studies*, 2012, *79* (1), 1–36.

**Argente, David, Sara Moreira, Ezra Oberfield, and Venky Venkateswaran**, "Scalable Expertise: How Standardization Drives Scale and Scope," Technical Report, National Bureau of Economic Research 2025.

**Arkolakis, Costas**, "Market penetration costs and the new consumers margin in international trade," *Journal of political economy*, 2010, *118* (6), 1151–1199.

**_ , Natalia Ramondo, Andrés Rodríguez-Clare, and Stephen Yeaple**, "Innovation and production in the global economy," *American Economic Review*, 2018, *108* (8), 2128–73.

**Arnoud, Antoine, Fatih Guvenen, and Tatjana Kleineberg**, "Benchmarking global optimizers," Technical Report, National Bureau of Economic Research 2019.

**Atalay, Enghin, Ali Hortaçsu, and Chad Syverson**, "Vertical integration and input flows," *American Economic Review*, 2014, *104* (4), 1120–48.

**Autor, David, David Dorn, and Gordon Hanson**, "When work disappears: Manufacturing decline and the falling marriage market value of young men," *American Economic Review: Insights*, 2019, *1* (2), 161–178.

**Autor, David H and David Dorn**, "The growth of low-skill service jobs and the polarization of the US labor market," *American economic review*, 2013, *103* (5), 1553–1597.

**Azar, José A, Steven T Berry, and Ioana Marinescu**, "Estimating labor market power," Technical Report, National Bureau of Economic Research 2022.

**Barnatchez, Keith, Leland Dod Crane, and Ryan Decker**, "An assessment of the national establishment time series (nets) database," *Finance and Economics Discussion Series (FEDS)*, 2017.

**Barth, Erling, Alex Bryson, James C Davis, and Richard Freeman**, "It's where you work: Increases in the dispersion of earnings across establishments and individuals in the United States," *Journal of Labor Economics*, 2016, *34* (S2), S67–S97.

**Bartik, Timothy J**, "Who benefits from state and local economic development policies?," 1991.

**Berger, David, Kyle Herkenhoff, and Simon Mongey**, "Labor market power," *American Economic Review*, 2022, *112* (4), 1147–93.

**Berry, Christopher R and Edward L Glaeser**, "The divergence of human capital levels across cities," *Papers in regional science*, 2005, *84* (3), 407–444.

**Cao, Dan, Henry R Hyatt, Toshihiko Mukoyama, and Erick Sager**, "Firm growth through new establishments," *Available at SSRN 3361451*, 2017.

**Card, David, Ana Rute Cardoso, Joerg Heining, and Patrick Kline**, "Firms and labor market inequality: Evidence and some theory," *Journal of Labor Economics*, 2018, *36* (S1), S13–S70.

**_ , Jesse Rothstein, and Moises Yi**, "Location, location, location," *American Economic Journal: Applied Economics*, 2025, *17* (1), 297–336.

**Costinot, Arnaud and Andrés Rodríguez-Clare**, "Trade theory with numbers: Quantifying the consequences of globalization," in "Handbook of international economics," Vol. 4, Elsevier, 2014, pp. 197–261.

**Davis, Donald R and James Harrigan**, "Good jobs, bad jobs, and trade liberalization," *Journal of international Economics*, 2011, *84* (1), 26–36.

**Davis, Morris A and François Ortalo-Magné**, "Household expenditures, wages, rents," *Review of Economic Dynamics*, 2011, *14* (2), 248–261.

**Diamond, Rebecca**, "The determinants and welfare implications of US workers' diverging location choices by skill: 1980-2000," *American Economic Review*, 2016, *106* (3), 479–524.

**Ding, Xiang**, *Industry linkages from joint production*, US Census Bureau, Center for Economic Studies, 2023.

**_ , Teresa C Fort, Stephen J Redding, and Peter K Schott**, "Structural change within versus across firms: Evidence from the United States," Technical Report, National Bureau of Economic Research 2022.

**Eckert, Fabian**, "Growing apart: Tradable services and the fragmentation of the us economy," *mimeograph, Yale University*, 2019.

_ **, Sharat Ganapati, and Conor Walsh**, "Urban-Biased Growth: A Macroeconomic Analysis," Technical Report, National Bureau of Economic Research 2022.

**Fort, Teresa C**, "Technology and production fragmentation: Domestic versus foreign sourcing," *The Review of Economic Studies*, 2017, *84* (2), 650–687.

_ **and Shawn Klimek**, "The effects of industry classification changes on US employment composition," *CES Working Paper 18-28, US Census Bureau*, 2018.

**Galle, Simon, Andrés Rodríguez-Clare, and Moises Yi**, "Slicing the pie: Quantifying the aggregate and distributional effects of trade," *The Review of Economic Studies*, 2023, *90* (1), 331–375.

**Garicano, Luis and Esteban Rossi-Hansberg**, "Inequality and the Organization of Knowledge," *American Economic Review*, 2004, *94* (2), 197–202.

_ **and** _ , "Organization and inequality in a knowledge economy," *The Quarterly journal of economics*, 2006, *121* (4), 1383–1435.

**Giannone, Elisa**, "Skilled-biased technical change and regional convergence," *University of Chicago. Unpublished manuscript*, 2017.

**Giroud, Xavier and Joshua Rauh**, "State taxation and the reallocation of business activity: Evidence from establishment-level data," *Journal of Political Economy*, 2019, *127* (3), 1262–1316.

_ **, Simone Lenzu, Quinn Maingi, and Holger Mueller**, "Propagation and amplification of local productivity spillovers," *Econometrica*, 2024, *92* (5), 1589–1619.

**Haltiwanger, John C, Henry R Hyatt, and James Spletzer**, "Industries, mega firms, and increasing inequality," Technical Report, National Bureau of Economic Research 2022.

**Helpman, Elhanan**, "A simple theory of international trade with multinational corporations," *Journal of political economy*, 1984, *92* (3), 451–471.

_ **, Oleg Itskhoki, and Stephen Redding**, "Inequality and unemployment in a global economy," *Econometrica*, 2010, *78* (4), 1239–1283.

**Holmes, Thomas J**, "The diffusion of Wal-Mart and economies of density," *Econometrica*, 2011, *79* (1), 253–302.

_ **and John J Stevens**, "An alternative theory of the plant size distribution, with geography and intra-and international trade," *Journal of Political Economy*, 2014, *122* (2), 369–421.

**Hsieh, Chang-Tai and Esteban Rossi-Hansberg**, "The industrial revolution in services," *Journal of Political Economy Macroeconomics*, 2023, *1* (1), 3–42.

**Jia, Panle**, "What happens when Wal-Mart comes to town: An empirical analysis of the discount retailing industry," *Econometrica*, 2008, *76* (6), 1263–1316.

**Jiang, Xian**, "Information and Communication Technology and Firm Geographic Expansion," Technical Report, working paper 2021.

**Katz, Lawrence F and Kevin M Murphy**, "Changes in relative wages, 1963–1987: supply and demand factors," *The quarterly journal of economics*, 1992, *107* (1), 35–78.

**Kline, Patrick, Neviana Petkova, Heidi Williams, and Owen Zidar**, "Who profits from patents? rent-sharing at innovative firms," *The quarterly journal of economics*, 2019, *134* (3), 1343–1404.

**Krusell, Per, Lee E Ohanian, José-Víctor Ríos-Rull, and Giovanni L Violante**, "Capital-skill complementarity and inequality: A macroeconomic analysis," *Econometrica*, 2000, *68* (5), 1029–1053.

**Lamadon, Thibaut, Magne Mogstad, and Bradley Setzler**, "Imperfect competition, compensating differentials, and rent sharing in the US labor market," *American Economic Review*, 2022, *112* (1), 169–212.

**Moretti, Enrico**, *The new geography of jobs*, Houghton Mifflin Harcourt, 2012.

**Oberfield, Ezra, Esteban Rossi-Hansberg, Pierre-Daniel Sarte, and Nicholas Trachter**, "Plants in space," *Journal of Political Economy*, 2024, *132* (3), 000–000.

**Philippon, Thomas and Ariell Reshef**, "Wages and human capital in the US finance industry: 1909–2006," *The Quarterly Journal of Economics*, 2012, *127* (4), 1551–1609.

**Ramondo, Natalia and Andrés Rodríguez-Clare**, "Trade, multinational production, and the gains from openness," *Journal of Political Economy*, 2013, *121* (2), 273–322.

**Rossi-Hansberg, Esteban, Pierre-Daniel Sarte, and Nicholas Trachter**, "Diverging trends in national and local concentration," *NBER Macroeconomics Annual*, 2021, *35* (1), 115–150.

**Rubinton, Hannah**, "The geography of business dynamism and skill biased technical change," *FRB St. Louis Working Paper*, 2020.

**Saiz, Albert**, "The geographic determinants of housing supply," *The Quarterly Journal of Economics*, 2010, *125* (3), 1253–1296.

**Setzler, Bradley and Felix Tintelnot**, "The effects of foreign multinationals on workers and firms in the United States," *The Quarterly Journal of Economics*, 2021, *136* (3), 1943–1991.

**Silva, JMC Santos and Silvana Tenreyro**, "The log of gravity," *The Review of Economics and statistics*, 2006, *88* (4), 641–658.

**Song, Jae, David J Price, Fatih Guvenen, Nicholas Bloom, and Till Von Wachter**, "Firming up inequality," *The Quarterly journal of economics*, 2019, *134* (1), 1–50.

**Tintelnot, Felix**, "Global production with export platforms," *The Quarterly Journal of Economics*, 2017, *132* (1), 157–209.

**Tolbert, Charles M and Molly Sizer**, "US commuting zones and labor market areas: A 1990 update," Technical Report 1996.

**Wenning, Derek**, "Equal prices, unequal access: The effects of national pricing in the us life insurance industry," *Unequal Access: The Effects of National Pricing in the US Life Insurance Industry (February 29, 2024)*, 2024.

# Appendices

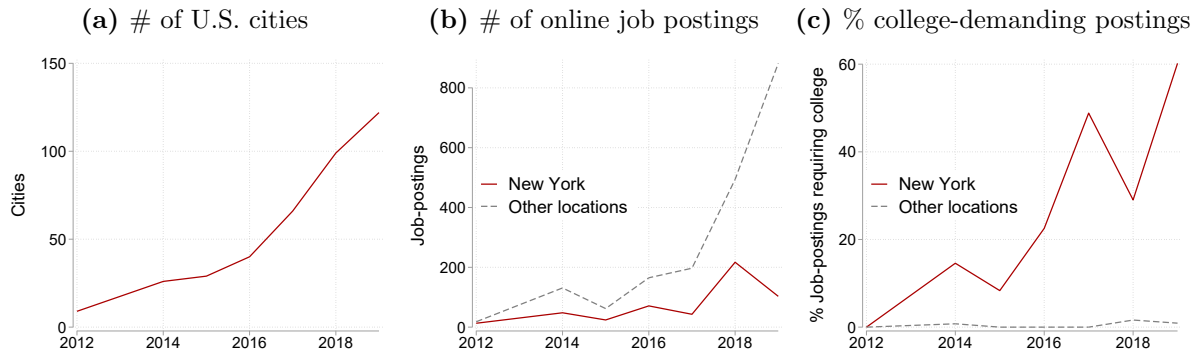
## A Case studies for firm spatial expansion

In this section, I provide two case studies that demonstrate the heterogeneous labor market implications of the spatial expansion of firms. Note that in both examples, I make no use of confidential Census Bureau data.

### A.1 The expansion of Shake Shack

The first case study is the expansion of Shake Shack, an American fast casual restaurant chain based in New York City, which opened its first restaurant outside of New York in 2010. I explore its demand for labor through the lens of its online job postings, as collected by Lightcast. Panel (a) in Figure 9 below shows the number of cities with postings by Shake Shack in the Lightcast data over time, which has risen in parallel to its expansion into more cities across the U.S. Panel (b) shows that most of these postings were in Shake Shack's new locations, with a significant share of them still posted in the original market of New York City. Panel (c) shows the skill intensity of these jobs, as measured by the share of total postings that explicitly require a college degree. While most of the posted jobs are low-skilled according to this measure, the jobs that Shake Shack opens in New York City are increasingly high-skilled, in line with the expansion of its headquarters in that location. By 2019, 60% of the New York City job postings required a college degree. A closer look reveals that these jobs include traditional headquarters-level occupations, such as management, design, marketing, and information-technology specialists.

**Figure 9: Example: the expansion of Shake Shack (online job-postings data)**




This example demonstrates the heterogeneous effects of firm expansion across two dimensions that are considered in the model: first, in the cross-section of establishments, between firm headquarters and its branches; second, across space, between New York City and other locations in the U.S. This example suggests potentially meaningful distribu-

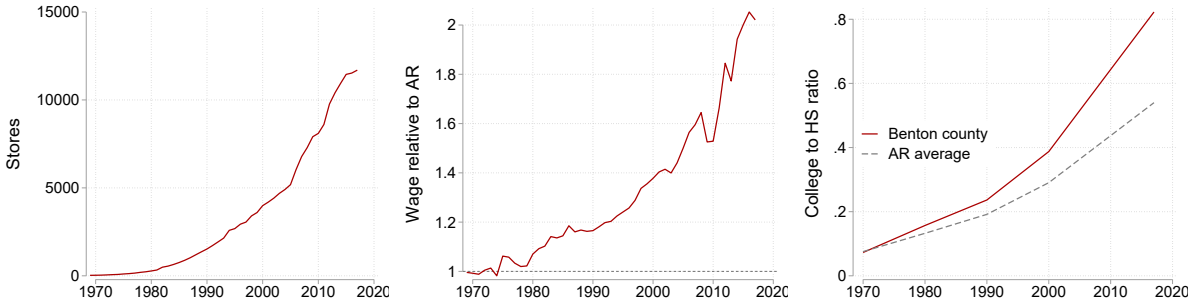
tional implications in the economy as a whole, given that the *average firm* in the economy has exhibited this kind of spatial expansion in recent decades.

## A.2 The expansion of Walmart

The second example is the expansion of the retail corporation Walmart, headquartered in Benton County (Bentonville City), Arkansas. Walmart is an outlier in the sense that it is a particularly large firm with headquarters in (what used to be) a relatively remote location. However, precisely because of this property, it provides a good example of the regional effects of firm spatial expansion at its headquarters market. Panel (a) of Figure 10 below shows the spatial expansion of Walmart as measured by its total number of stores (data from Walmart Inc.). The chain's expansion outside of Arkansas started in the late 1960s and has been growing exponentially since then. Panels (b) and (c) show regional outcomes for Benton County, Arkansas, relative to the Arkansas average. In the early 1970s, Benton County resembled an average Arkansas county, with a similar wage to the rest of the state and with similar skill intensity, as measured by the ratio of workers with a college degree to workers with only a high school diploma. Since then, in parallel with the expansion of Walmart's headquarters in that county, the average wage has diverged from the rest of the state, such that by now it is more than double the average Arkansas wage. At the same time, it has experienced an influx of college graduates, leading to stronger skill-deepening than the rest of the state. These outcomes are in line with the central mechanism of the model, in which the aggregate spatial expansion of firms leads to greater growth of income and demand for skilled labor in locations that specialize in providing headquarters services.

**Figure 10: Example: the expansion of Walmart (Walmart Inc. data)**

**(a)** # of Walmart stores overall **(b)** Benton to AR avg. wage ratio **(c)** Benton & AR college/HS ratio




This example demonstrates the heterogeneous effects of firm spatial expansion once aggregated to the regional level, with divergence in income and demand for skill across different areas within a single state. As in the case of the previous example of Shake Shack's expansion, it is suggestive of potentially important heterogeneous effects for the economy as a whole, given the observed spatial expansion of the average firm in the data.

# B Model derivations

## B.1 Proof of proposition 1

Recall that $x_{ij}(z) W_{\ell j} \ell_{ij}(z)^{\frac{\epsilon+1}{\epsilon}}$ is the firm's payroll in region $j$ and denote by $x_{ij}(z) r_{ij}(z)$ the firm's sales in region $j$. The share of $x_{ij}(z) r_{ij}(z)$ that is allocated toward labor compensation (in both the branch and the headquarters) is given by $\psi \equiv \frac{\epsilon}{\epsilon+1}\frac{\sigma-1}{\sigma}$. Define the share of payroll originating from market $j$ ($\psi\, x_{ij}(z) r_{ij}(z)$) that is allocated toward headquarters labor as $\Gamma_{ij}(z)$, given by

$$\Gamma_{ij}(z) \equiv 1 - \frac{W_{\ell j} \ell_{ij}(z)^{\frac{\epsilon+1}{\epsilon}}}{\psi\, r_{ij}(z)}.$$

It is also helpful to define $h_{ij}(z) \equiv \left(\frac{\Gamma_{ij}(z)}{1-\Gamma_{ij}(z)}\frac{W_{\ell j}}{W_{hi}}\right)^{\frac{\epsilon}{\epsilon+1}} \ell_{ij}(z)$ as the hypothetical market-specific allocation of headquarters labor. Note that $W_{hi} h_i(z)^{\frac{\epsilon+1}{\epsilon}} = \sum_{j=1}^{N} x_{ij}(z) W_{hi} h_{ij}(z)^{\frac{\epsilon+1}{\epsilon}}$, or $h_i(z) = \left[\sum_{j=1}^{N} x_{ij}(z) h_{ij}(z)^{\frac{\epsilon+1}{\epsilon}}\right]^{\frac{\epsilon}{\epsilon+1}}$.

By the firm's first-order condition with respect to $\ell_{ij}(z)$, the ratio of the $j$-originated headquarters payroll to the branch payroll in $j$ is

$$\begin{aligned}
\frac{x_{ij}(z) W_{hi} h_{ij}(z)^{\frac{\epsilon+1}{\epsilon}}}{x_{ij}(z) W_{\ell j} \ell_{ij}(z)^{\frac{\epsilon+1}{\epsilon}}} &= \frac{x_{ij}(z)\psi r_{ij}(z) - x_{ij}(z) W_{\ell j} \ell_{ij}(z)^{\frac{\epsilon+1}{\epsilon}}}{x_{ij}(z) W_{\ell j} \ell_{ij}(z)^{\frac{\epsilon+1}{\epsilon}}} \\
&= \frac{\Gamma_{ij}(z)}{1-\Gamma_{ij}(z)} \\
&= \gamma \left(\frac{h_i(z)}{\ell_{ij}(z)}\right)^{\frac{\eta-1}{\eta}}.
\end{aligned}$$

Rearranging both sides of this expression we obtain

$$\frac{W_{hi}}{W_{\ell j}} \left(\frac{h_{ij}(z)}{\ell_{ij}(z)}\right)^{\frac{\epsilon+1}{\epsilon}} = \left(\frac{W_{hi}}{W_{\ell j}}\right)^{-\frac{\frac{\eta-1}{\eta}}{\frac{\epsilon+1}{\epsilon}-\frac{\eta-1}{\eta}}} \left[\gamma \left(\frac{h_i(z)}{h_{ij}(z)}\right)^{\frac{\eta-1}{\eta}}\right]^{\frac{\frac{\epsilon+1}{\epsilon}}{\frac{\epsilon+1}{\epsilon}-\frac{\eta-1}{\eta}}}.$$

The headquarters payroll share at the firm level can then be obtained as

$$\begin{aligned}
\frac{\Gamma_i(z)}{1-\Gamma_i(z)} &= \frac{\sum_{j=1}^{N} x_{ij}(z) W_{hi} h_{ij}(z)^{\frac{\epsilon+1}{\epsilon}}}{\sum_{k=1}^{N} x_{ik}(z) W_{\ell k} \ell_{ik}(z)^{\frac{\epsilon+1}{\epsilon}}} \\
&= \sum_{j=1}^{N} \frac{x_{ij}(z) W_{\ell j} \ell_{ij}(z)^{\frac{\epsilon+1}{\epsilon}}}{\sum_{k=1}^{N} x_{ik}(z) W_{\ell k} \ell_{ik}(z)^{\frac{\epsilon+1}{\epsilon}}} \frac{W_{hi} h_{ij}(z)^{\frac{\epsilon+1}{\epsilon}}}{W_{\ell j} \ell_{ij}(z)^{\frac{\epsilon+1}{\epsilon}}} \\
&= \gamma^{\frac{1}{1-\frac{\epsilon}{\epsilon+1}\frac{\eta-1}{\eta}}} \sum_{j=1}^{N} \frac{x_{ij}(z) W_{\ell j} \ell_{ij}(z)^{\frac{\epsilon+1}{\epsilon}}}{\sum_{k=1}^{N} x_{ik}(z) W_{\ell k} \ell_{ik}(z)^{\frac{\epsilon+1}{\epsilon}}} \left( \left( \frac{W_{\ell j}}{W_{hi}} \right) \left( \frac{h_i(z)}{h_{ij}(z)} \right)^{\frac{\epsilon+1}{\epsilon}} \right)^{\frac{\frac{\epsilon}{\epsilon+1}\frac{\eta-1}{\eta}}{1-\frac{\epsilon}{\epsilon+1}\frac{\eta-1}{\eta}}} \\
&= \gamma^{\frac{1}{1-\frac{\epsilon}{\epsilon+1}\frac{\eta-1}{\eta}}} \sum_{j=1}^{N} \frac{x_{ij}(z) W_{\ell j} \ell_{ij}(z)^{\frac{\epsilon+1}{\epsilon}}}{\sum_{k=1}^{N} x_{ik}(z) W_{\ell k} \ell_{ik}(z)^{\frac{\epsilon+1}{\epsilon}}} \\
&\quad \times \left( \frac{W_{\ell j}}{W_{hi}} \frac{\Gamma_i(z)}{\Gamma_{ij}(z)} \sum_{m=1}^{N} \frac{x_{im}(z)}{\sum_{k=1}^{N} x_{ik}(z)} \frac{r_{im}(z)}{r_{ij}(z)} \right)^{\frac{\frac{\epsilon}{\epsilon+1}\frac{\eta-1}{\eta}}{1-\frac{\epsilon}{\epsilon+1}\frac{\eta-1}{\eta}}} \\
&= \gamma^{\bar{\eta}} \left( \frac{\bar{W}_{\ell i}(z)}{W_{hi}} \left( \sum_{j=1}^{N} x_{ij}(z) \right) \right)^{\bar{\eta}-1}.
\end{aligned}$$

where

$$\bar{W}_{\ell i}(z) \equiv \left( \sum_{j=1}^{N} \Omega_{ij}(z) W_{\ell j}^{\bar{\eta}-1} \right)^{\frac{1}{\bar{\eta}-1}},$$

$$\bar{\eta} \equiv \frac{\eta}{\eta - \frac{\epsilon}{\epsilon+1}(\eta-1)},$$

$$\Omega_{ij}(z) \equiv \frac{x_{ij}(z) W_{\ell j} \ell_{ij}(z)^{\frac{\epsilon+1}{\epsilon}}}{\sum_{k=1}^{N} x_{ik}(z) W_{\ell k} \ell_{ik}(z)^{\frac{\epsilon+1}{\epsilon}}} \times \left( \frac{\Gamma_i(z)}{\Gamma_{ij}(z)} \times \sum_{m=1}^{N} \frac{x_{im}(z)}{\sum_{k=1}^{N} x_{ik}(z)} \frac{r_{im}(z)}{r_{ij}(z)} \right)^{\bar{\eta}-1}.$$

In the case of symmetric space, the above expressions reduce to:

$$\frac{\Gamma(z)}{1-\Gamma(z)} = \gamma^{\bar{\eta}} \left( \frac{W_\ell}{W_h} N x(z) \right)^{\bar{\eta}-1},$$

where $x(z)$ is the share of locations served by a firm of type $z$ in each of the $N$ regions in the economy. When, in addition, $\epsilon \to \infty$, we get that $W_h$ and $W_\ell$ are the economy-wide prices of headquarters labor and branch-level labor, respectively, and that the headquarters payroll share for firm $z$ is

$$\Gamma(z) = \frac{\gamma^{\eta} \left( \frac{W_\ell}{W_h} N x(z) \right)^{\eta-1}}{1 + \gamma^{\eta} \left( \frac{W_\ell}{W_h} N x(z) \right)^{\eta-1}}.$$

Let $\bar{H}$ and $\bar{L}$ be the endowments of headquarters labor and branch-level labor in the

economy, and let $F(z)$ be the cumulative distribution function for firms of type $z$. In this case:

$$\frac{\Gamma(z)}{1-\Gamma(z)} = \frac{W_h}{W_\ell}\frac{h(z)}{N\,x(z)\,\ell(z)} = \gamma^\eta \left(\frac{W_\ell}{W_h} N\,x(z)\right)^{\eta-1},$$

and

$$h(z) = \gamma^\eta \left(\frac{W_\ell}{W_h}\right)^\eta (N\,x(z))^{\eta-1} N\,x(z)\,\ell(z).$$

Labor market clearing for headquarters workers implies

$$\begin{aligned}\bar{H} &= \gamma^\eta \left(\frac{W_\ell}{W_h}\right)^\eta \int_z (N\,x(z))^{\eta-1} N\,x(z)\,\ell(z)\; dF(z) \\ &= \gamma^\eta \left(\frac{W_\ell}{W_h}\right)^\eta \int_z \frac{N\,x(z)\,\ell(z)}{\bar{L}} (N\,x(z))^{\eta-1}\; dF(z)\; \bar{L}.\end{aligned}$$

Therefore the ratio of headquarters wages to branch-level wages is given by

$$\frac{W_h}{W_\ell} = \gamma\, \bar{x}^{\frac{\eta-1}{\eta}} \left(\frac{\bar{H}}{\bar{L}}\right)^{-\frac{1}{\eta}},$$

where $\bar{x}$ is a weighted power mean of firms' geographical scope in the economy,

$$\bar{x} \equiv \left(\int_z s_\ell(z)\,(N\,x(z))^{\eta-1}\,dF(z)\right)^{\frac{1}{\eta-1}}, \quad s_\ell(z) \equiv \frac{N\,x(z)\,\ell(z)}{\bar{L}},$$

with branch-level employment of each firm ($s_\ell(z)$) as the weights.

## B.2 Proof of propositions 2 and 3

Recall the firm's problem

$$\begin{aligned}\max_{h_i(z),\{\ell_{ij}(z)\},\{x_{ij}(z)\}} &\sum_{j=1}^N x_{ij}(z)\,\Upsilon_j \left(A_i z \left(\gamma h_i(z)^{\frac{\eta-1}{\eta}} + \ell_{ij}(z)^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta}{\eta-1}}\right)^{\frac{\sigma-1}{\sigma}} \\ &- \sum_{j=1}^N x_{ij}(z)\, W_{\ell j}\ell_{ij}(z)^{\frac{\epsilon+1}{\epsilon}} - W_{hi}h_i(z)^{\frac{\epsilon+1}{\epsilon}} - C\left(\{x_{ij}(z)\}_{j=1}^N\right).\end{aligned}$$

The first-order condition with respect to $\ell_{ij}(z)$ yields

$$\ell_{ij}(z) = \left(\psi\,(1-\Gamma_{ij}(z))\,\frac{r_{ij}(z)}{W_{\ell j}}\right)^{\frac{\epsilon}{\epsilon+1}},$$

where $\psi \equiv \frac{\epsilon}{\epsilon+1}\frac{\sigma-1}{\sigma}$, $r_{ij}(z)$ is the firm's local sales, and we define the intensity of headquarters labor in local production as $\Gamma_{ij}(z) \equiv \frac{\gamma h_i(z)^{\frac{\eta-1}{\eta}}}{\gamma h_i(z)^{\frac{\eta-1}{\eta}} + \ell_{ij}(z)^{\frac{\eta-1}{\eta}}}$. Noting that $w_{\ell,ij}(z) =$

$W_{\ell j}\ell_{ij}(z)^{\frac{1}{\epsilon}}$ and rearranging, we obtain

$$\log w_{\ell,ij}(z) = \log\psi + \frac{\epsilon}{\epsilon+1}\log W_{\ell j} + \frac{1}{\epsilon+1}\log\left(1-\Gamma_{ij}(z)\right) + \frac{1}{\epsilon+1}\log r_{ij}(z).$$

The first-order condition with respect to $h_i(z)$ yields

$$W_{hi}h_i(z)^{\frac{\epsilon+1}{\epsilon}} = \psi\sum_{j=1}^{N}\Gamma_{ij}(z)\,x_{ij}(z)\,r_{ij}(z).$$

Noting that $w_{h,i}(z) = W_{hi}h_i(z)^{\frac{1}{\epsilon}}$ and rearranging, we obtain

$$w_{h,i}(z)^{\epsilon+1}W_{hi}^{-\epsilon} = \psi\underbrace{\left(\sum_{j=1}^{N}\frac{x_{ij}(z)\,r_{ij}(z)}{\sum_{k=1}^{N}x_{ik}(z)\,r_{ik}(z)}\Gamma_{ij}(z)\right)}_{=\Gamma_i(z)}\left(\sum_{k=1}^{N}x_{ik}(z)\,r_{ik}(z)\right),$$

which yields

$$\log w_{h,i}(z) = \log\psi + \frac{\epsilon}{\epsilon+1}\log W_{hi} + \frac{1}{\epsilon+1}\log\Gamma_i(z) + \frac{1}{\epsilon+1}\log\left(\sum_{k=1}^{N}x_{ik}(z)\,r_{ik}(z)\right).$$

We now consider a small homogeneous reduction in the cost function $C(\cdot)$ that results in positive expansion in all markets, $d\log x_{ij}(z) > 0$. Log-linearizing the first-order condition for headquarters labor, we obtain

$$\frac{\epsilon+1}{\epsilon}\,d\log h_i(z) = \sum_{j=1}^{N}\phi_{ij}(z)\left[d\log\Gamma_{ij}(z) + d\log x_{ij}(z) + d\log r_{ij}(z)\right],$$

where we define

$$\phi_{ij}(z) \equiv \frac{\Gamma_{ij}(z)\,x_{ij}(z)\,r_{ij}(z)}{\sum_{k=1}^{N}\Gamma_{ik}(z)\,x_{ik}(z)\,r_{ik}(z)}.$$

Log-linearizing the first-order condition for branch-level labor, we obtain

$$d\log\ell_{ij}(z) = \frac{\epsilon}{\epsilon+1}d\log r_{ij}(z) - \frac{\epsilon}{\epsilon+1}\frac{\Gamma_{ij}(z)}{1-\Gamma_{ij}(z)}d\log\Gamma_{ij}(z),$$

Log-linearizing the local labor intensity, we obtain

$$d\log\Gamma_{ij}(z) = \left(1-\Gamma_{ij}(z)\right)\frac{\eta-1}{\eta}\left(d\log h_i(z) - d\log\ell_{ij}(z)\right).$$

Log-linearizing local sales $r_{ij}(z)$, we obtain

$$d\log r_{ij}(z) = \frac{\sigma-1}{\sigma}\left(\Gamma_{ij}(z)\,d\log h_i(z) + \left(1-\Gamma_{ij}(z)\right)d\log\ell_{ij}(z)\right).$$

Combining the above expressions, we obtain two relationships between $d\log \ell_{ij}(z)$ and $d\log h_i(z)$:

$$\begin{aligned} d\log h_i(z) = &\frac{\epsilon}{\epsilon+1}\sum_{j=1}^{N}\phi_{ij}(z)\, d\log x_{ij}(z) \\ &+\sum_{j=1}^{N}\phi_{ij}(z)\left((1-\Gamma_{ij}(z))\frac{\eta-1}{\eta}+\Gamma_{ij}(z)\frac{\sigma-1}{\sigma}\right) d\log h_i(z) \\ &+\sum_{j=1}^{N}\phi_{ij}(z)(1-\Gamma_{ij}(z))\left(\frac{\sigma-1}{\sigma}-\frac{\eta-1}{\eta}\right) d\log \ell_{ij}(z), \end{aligned} \tag{A.1}$$

$$d\log \ell_{ij}(z) = \tilde{\Gamma}_{ij}(z)\, d\log h_i(z), \tag{A.2}$$

where we denote

$$\tilde{\Gamma}_{ij}(z) \equiv \frac{\frac{\epsilon}{\epsilon+1}\left(\frac{\sigma-1}{\sigma}-\frac{\eta-1}{\eta}\right)\Gamma_{ij}(z)}{1-\frac{\epsilon}{\epsilon+1}\left(\frac{\sigma-1}{\sigma}(1-\Gamma_{ij}(z))+\frac{\eta-1}{\eta}\Gamma_{ij}(z)\right)} < 1.$$

First, we show that $d\log h_i(z) > 0$ following a small reduction in the cost of expansion. Combining (A.1) and (A.2) and rearranging, we obtain

$$d\log h_i(z) = \frac{\epsilon}{\epsilon+1}\left[1-\left(\left(1-\tilde{\phi}_i(z)\right)\frac{\sigma-1}{\sigma}+\tilde{\phi}_i(z)\frac{\eta-1}{\eta}\right)\right]^{-1}\sum_{j=1}^{N}\phi_{ij}(z)\, d\log x_{ij}(z),$$

where we define

$$\tilde{\phi}_i(z) \equiv \sum_{j=1}^{N}\phi_{ij}(z)(1-\Gamma_{ij}(z))\left(1-\tilde{\Gamma}_{ij}(z)\right).$$

Since $\tilde{\phi}_i(z) < 1$, we get that $\left(1-\tilde{\phi}_i(z)\right)\frac{\sigma-1}{\sigma}+\tilde{\phi}_i(z)\frac{\eta-1}{\eta} < 1$ and $d\log h_i(z) > 0$ when $\sum_{j=1}^{N}\phi_{ij}(z)\, d\log x_{ij}(z) > 0$. Consequently, also $d\log w_{h,i}(z) > 0$.

To see that headquarters wages increase more than branch-level wages, note that, holding constant the labor-supply shifters (or allowing them to change but comparing headquarters and branch labor in the same market such that $i=j$), we get $d\log w_{h,i}(z) - d\log w_{\ell,ij}(z) = \frac{1}{\epsilon}\left(\frac{1}{\epsilon}d\log h_i(z) - d\log \ell_{ij}(z)\right)$. Since $\tilde{\Gamma}_{ij}(z)$ is always smaller than 1, a decline in the cost of expansion that keeps market-specific demand shifters $\Upsilon_j$ constant results in $d\log \ell_{ij}(z) < d\log h_i(z)$. Consequently, $d\log w_{\ell,ij}(z) < d\log w_{h,i}(z)$.

Finally, whenever $d\log h_i(z) > 0$, then $d\log w_{\ell,ij}(z) > 0$ if and only if $\tilde{\Gamma}_{ij}(z) > 0$. This is the case if $\frac{\sigma-1}{\sigma}-\frac{\eta-1}{\eta} > 0$, or $\sigma > \eta$. This completes the proof.

## B.3 Proof of proposition 4

Denote by $R_i$ and $E_i$ the total sales of locally headquartered firms in region $i$ and total expenditure on locally active branches in region $i$, respectively. $R_i$ and $E_i$ are given by

$$R_i = \sum_{j=1}^{N} M_i \int_z x_{ij}(z)\, r_{ij}(z)\, dG_i(z),$$

$$E_i = \sum_{n=1}^{N} M_n \int_z x_{ni}(z)\, r_{ni}(z)\, dG_n(z),$$

where $M_i$ is the mass of firms headquartered in region $i$ and $G_n(z)$ is the productivity distribution of these firms. In addition, define the headquarters intensity of locally headquartered firms, $\bar{\Gamma}_{out,i}$, and of locally active branches, $\bar{\Gamma}_{in,i}$, as

$$\bar{\Gamma}_{out,i} = \sum_{j=1}^{N} \int_z \left( \frac{x_{ij}(z)\, r_{ij}(z)}{\int_{z'} x_{ij}(z')\, r_{ij}(z')\, dG_i(z')} \right) \Gamma_{ij}(z)\, dG_i(z),$$

$$\bar{\Gamma}_{in,i} = \sum_{n=1}^{N} \int_z \left( \frac{x_{ni}(z)\, r_{ni}(z)}{\int_{z'} x_{ni}(z')\, r_{ni}(z')\, dG_n(z')} \right) \Gamma_{ni}(z)\, dG_n(z).$$

We can now write the total headquarters payroll in region $i$ as

$$\psi\, \bar{\Gamma}_{out,i} R_i,$$

total payroll to branch workers in region $i$ as

$$\psi\, (1 - \bar{\Gamma}_{in,i}) E_i,$$

and total payroll in the goods-producing sector as $\psi\, R_{g,i}$, where $R_{g,i}$ stands for total sales from tradable goods in region $i$.

We now assume that the bundles of headquarters labor, branch-level labor, and tradable-goods labor are all Cobb–Douglas aggregators of labor from different skill groups $s \in \{1, ..., S\}$, with skill intensities $\alpha_{h,s}, \alpha_{\ell,s}, \alpha_{g,s}$, respectively. Define $\mathcal{S}_{is}$ as the share of regional income in $i$ that is paid to skill group $s$. From the above derivations, it can be written as

$$\mathcal{S}_{is} = \frac{\alpha_{g,s}\psi R_{g,i} + \alpha_{h,s}\bar{\Gamma}_{out,i}\psi R_i + \alpha_{b,s}\left(1 - \bar{\Gamma}_{in,i}\right)\psi E_i}{Y_i},$$

where $Y_i$ is total regional income.

Since income must equal expenditure, $E_i = \beta Y_i$, the above can be rewritten as

$$\mathcal{S}_{is} = \beta\psi\alpha_{g,s}\frac{R_{g,i}}{E_i} + \beta\psi\alpha_{h,s}\bar{\Gamma}_{out,i}\frac{R_i}{E_i} + \beta\psi\alpha_{b,s}\left(1 - \bar{\Gamma}_{in,i}\right).$$

Since overall trade in goods and services is balanced in the model, total outflow of headquarters services must equal the total inflow of tradable goods:

$$R_{g,i} = \left(\frac{1}{\beta\psi} - \left(1 - \bar{\Gamma}_{in,i}\right)\right) E_i - \bar{\Gamma}_{out,i} R_i.$$

Combining the above expressions, we obtain

$$\mathcal{S}_{is} = \left(1 - \tilde{\beta}\right) \alpha_{g,s} + \tilde{\beta} \left[\alpha_{g,s} + \left(\alpha_{h,s} - \alpha_{g,s}\right) \bar{\Gamma}_{out,i} \frac{R_i}{E_i} + \left(\alpha_{b,s} - \alpha_{g,s}\right) \left(1 - \bar{\Gamma}_{in,i}\right)\right], \quad \text{(A.3)}$$

where we denote $\tilde{\beta} = \beta\psi$.

Consider now a few special cases of Equation A.3. Suppose that $s$ is the group of skilled labor. When headquarters production uses only skilled labor ($\alpha_{h,s} = 1$), branch production uses only low-skilled labor ($\alpha_{b,s} = 1$), and the skilled intensity of tradable goods is given by $\alpha = \alpha_{g,s}$, then the skilled income share in region $i$ is given by

$$\mathcal{S}_{is} = \left(1 - \tilde{\beta}\right) \alpha + \tilde{\beta} \left[\left(1 - \alpha\right) \bar{\Gamma}_{out,i} \frac{R_i}{E_i} + \alpha \bar{\Gamma}_{in,i}\right],$$

as in Proposition 4.

For the economy as a whole, the income share of skill group $s$ is given by

$$\mathcal{S}_s = \sum_{i=1}^{N} \frac{Y_i}{\sum_{j=1}^{N} Y_j} \mathcal{S}_{is} = \left(1 - \tilde{\beta}\right) \alpha + \tilde{\beta} \bar{\Gamma},$$

where $\bar{\Gamma}$ is the aggregate intensity of the factors producing the non-rival headquarters input in the production of services.

When $\eta = 1$, $\bar{\Gamma} = \bar{\Gamma}_{out,i} = \bar{\Gamma}_{in,i} = \frac{\gamma}{\gamma+1}$. In this case, the aggregate skill intensity is invariant to the spatial distribution of economic activity and given by

$$\mathcal{S}_s = \left(1 - \tilde{\beta}\right) \alpha + \tilde{\beta} \frac{\gamma}{\gamma + 1}.$$

In this case, regional skill intensity is summarized by regional specialization in providing headquarters services, captured by the ratio $\frac{R_i}{E_i}$:

$$\mathcal{S}_{is} = \left(1 - \tilde{\beta}\right) \alpha + \tilde{\beta} \left[\left(1 - \alpha\right) \frac{R_i}{E_i} + \alpha\right] \frac{\gamma}{\gamma + 1}.$$

In particular, when there is no cross-region firm activity (e.g., when the cost of expansion is infinite and $x_{ij}(z) > 0$ only for $i = j$), then all spatial differences in skill intensity are eliminated, and skill intensity in all regions is given by $\left(1 - \tilde{\beta}\right) \alpha + \tilde{\beta} \frac{\gamma}{\gamma+1}$.

# C Model inversion

The set of location fundamentals $\mathcal{A} \equiv \{A_i, B_{is}, A_{g,i}\}_{i,s}$ includes the productivity shifters in the services sector, regional amenities, and regional productivity in tradable-goods. I obtain $\mathcal{A}$ by inverting the model's equilibrium conditions, conditional on all other parameters.

To this end, I match three sets of regional outcomes in the data: total regional labor income, $Y_i$, regional employment in each skill group, $L_{is}$, and relative payroll for skilled workers, which I label $\omega_i$. I define $\omega_i$ as the ratio between total payroll of workers in the upper half of the skill distribution (i.e., the top 2 skill groups that I defined in Section 5) to the income of workers in the bottom half of the skill distribution (the bottom 2 skill groups from Section 5).

## C.1 Preliminaries: obtaining regional sales and expenditure

It is first useful to recall the definition of key regional outcomes in the model, which I defined in Sections 3 and B.3. I denote by $R_i$ and $E_i$ total sales of the firms that are headquartered in region $i$ and expenditure on branches that operate in region $i$, respectively. I denote by $R_{g,i}$ total regional sales of tradable goods in region $i$. I define by $\bar{\Gamma}_{out,i}$ and $\bar{\Gamma}_{in,i}$ the headquarters-intensity of locally-headquartered firms and of locally active branches, respectively.[45] Finally recall that $\alpha_{hs}, \alpha_{\ell s}, \alpha_{gs}$ denote the intensity of group $s$ in the headquarters labor bundle, branch-level labor bundle, and tradable-goods labor bundle, respectively. With the above definition, total income of group $s$ in region $i$ is given by

$$Y_{is} = \psi \left( \alpha_{hs} \bar{\Gamma}_{out,i} R_i + \alpha_{bs} \left( 1 - \bar{\Gamma}_{in,i} \right) E_i + \alpha_{gs} R_{g,i} \right),$$

where recall that $\psi \equiv \frac{\sigma-1}{\sigma} \frac{\epsilon}{\epsilon+1}$ is the payroll share of sales. We then have that $\omega_i$ in the model is defined as

$$\omega_i \equiv \frac{Y_{i1} + Y_{i2}}{Y_{i3} + Y_{i4}},$$

where groups 1 and 2 are the top two skill groups and groups 3 and 4 are the bottom two skill groups.

The first step of the inversion is to obtain $E_i$, $R_i$, and $R_{g,i}$ that are consistent with the model's structure and with the above observables $Y_i$ and $\omega_i$. Since $E_i$ is total local

[45] I elaborate on these definitions in Appendix B.3:

$$\bar{\Gamma}_{out,i} = \sum_{j=1}^{N} \int_z \left( \frac{x_{ij}(z) r_{ij}(z)}{\int_{z'} x_{ij}(z') r_{ij}(z') \, dG_i(z')} \right) \Gamma_{ij}(z) \, dG_i(z),$$

$$\bar{\Gamma}_{in,i} = \sum_{n=1}^{N} \int_z \left( \frac{x_{ni}(z) r_{ni}(z)}{\int_{z'} x_{ni}(z') r_{ni}(z') \, dG_n(z')} \right) \Gamma_{ni}(z) \, dG_n(z).$$

expenditure on local branches, it is simply given by $E_i = \beta Y_i$, where recall that $\beta$ is exogenous expenditure share on services. Now, let $s$ denote the upper half of the skill distribution and $k$ denote the lower half. From the above expression for total income of group $s$ in region $i$, we have that

$$\left(\omega_i \alpha_{gk} - \alpha_{gs}\right) R_{g,i} = \left(\alpha_{hs} - \omega_i \alpha_{hk}\right) \bar{\Gamma}_{out,i} R_i + \left(\alpha_{\ell s} - \omega_i \alpha_{\ell k}\right) \left(1 - \bar{\Gamma}_{in,i}\right) E_i.$$

In addition, regional balance of payments in the model implies that

$$R_{g,i} = \left(\frac{1}{\beta\psi} - \left(1 - \bar{\Gamma}_{in,i}\right)\right) E_i - \bar{\Gamma}_{out,i} R_i.$$

Combining these two equations, we obtain two equations for $R_i$ and $R_{g,i}$, conditional on parameters, observables, and guesses for $\bar{\Gamma}_{in,i}$ and $\bar{\Gamma}_{out,i}$:

$$R_{g,i} = \frac{\left(\beta\psi\right)^{-1} \left(\alpha_{hs} - \omega_i \alpha_{hk}\right) + \left(\left(\alpha_{\ell s} - \omega_i \alpha_{\ell k}\right) - \left(\alpha_{hs} - \omega_i \alpha_{hk}\right)\right) \bar{\nu}_{in,i}}{\left(\alpha_{hs} - \omega_i \alpha_{hk}\right) - \left(\alpha_{gs} - \omega_i \alpha_{gk}\right)} E_i$$

$$R_i = \frac{1}{\psi \bar{\Gamma}_{out,i}} \left[E_i \left(1/\beta - \psi \left(1 - \bar{\Gamma}_{in,i}\right)\right) - \psi R_{g,i}\right]$$

## C.2 Inverting $A_i$

We now invert $A_i$ using the recovered value for $R_i$. Recall that sales of a firm with productivity $z$ from region $i$ in region $j$ are given by

$$x_{ij}\left(z\right) r_{ij}\left(z\right) = x_{ij}\left(z\right) E_j Q_j^{-\frac{\sigma-1}{\sigma}} A_i^{\frac{\sigma-1}{\sigma}} z^{\frac{\sigma-1}{\sigma}} \gamma^{\frac{\sigma-1}{\sigma}\frac{\eta}{\eta-1}} \Gamma_{ij}\left(z\right)^{-\frac{\sigma-1}{\sigma}\frac{\eta}{\eta-1}} h_i\left(z\right)^{\frac{\sigma-1}{\sigma}}$$

Aggregating across all $i$-based firms, we obtain:

$$R_{ij} = \int_z x_{ij}\left(z\right) r_{ij}\left(z\right) dG_i\left(z\right) = E_j Q_j^{-\frac{\sigma-1}{\sigma}} A_i^{\frac{\sigma-1}{\sigma}} \gamma^{\frac{\sigma-1}{\sigma}\frac{\eta}{\eta-1}} \Xi_{ij},$$

where we define

$$\Xi_{ij} \equiv \int_z x_{ij}\left(z\right) z^{\frac{\sigma-1}{\sigma}} \Gamma_{ij}\left(z\right)^{-\frac{\sigma-1}{\sigma}\frac{\eta}{\eta-1}} h_i\left(z\right)^{\frac{\sigma-1}{\sigma}} dG_i\left(z\right).$$

Total sales by $i$-headquartered firms are given by

$$R_i = A_i^{\frac{\sigma-1}{\sigma}} \gamma^{\frac{\sigma-1}{\sigma}\frac{\eta}{\eta-1}} \sum_{j=1}^{N} E_j Q_j^{-\frac{\sigma-1}{\sigma}} \Xi_{ij}.$$

Thus, we can update our guess for $A_i$ by iterating over the following equation as part of the solution of equilibrium:

$$A_i = \gamma^{-\frac{\eta}{\eta-1}} \left( \frac{R_i}{\sum_{j=1}^{N} E_j Q_j^{-\frac{\sigma-1}{\sigma}} \Xi_{ij}} \right)^{\frac{\sigma}{\sigma-1}}.$$

Note that we have replaced the equilibrium condition that derives $R_i$ based on knowledge of $A_i$ with a condition that recovers $A_i$ based on knowledge of $R_i$.

## C.3 Inverting $A_{g,i}$

Using the information on $R_{g,i}$, we can similarly obtain a new guess for $A_{g,i}$ from the first order condition of firms in $i$'s tradable sector, yielding:

$$A_{g,i} = \left( \frac{R_{g,i}}{M_{g,i} E_g Q_g^{-\frac{\sigma-1}{\sigma}} \left( \prod_{s=1}^{S} \ell_{g,is}^{\alpha_{gs}} \right)^{\frac{\sigma-1}{\sigma}}} \right)^{\frac{\sigma}{\sigma-1}},$$

where $M_{g,i}$ is the guess for the mass of firms in the tradable sector; $E_g$ and $Q_g$ are the economy-wide expenditure and quantity index in the tradable sector; and $\prod_{s=1}^{S} \ell_{g,is}^{\alpha_{gs}}$ is the Cobb-Douglas labor aggregator in the tradable production function.

## C.4 Inverting $B_{is}$

Inverting the location choice equation 10, we obtain a new guess for $B_{is}$:

$$B_{is} = \frac{\bar{P}_i}{\bar{W}_{is}} \left( \sum_{j=1}^{N} \left( B_{js} \bar{W}_{js} / \bar{P}_j \right)^{\xi} \frac{L_{is}}{\sum_{j=1}^{N} L_{js}} \right)^{\frac{1}{\xi}}$$

where $\bar{P}_i$ is the guess for the regional price index (including housing) and $\bar{W}_{is} \equiv L_{is}^{\frac{1}{\epsilon}} W_{is}$ is the regional ideal price index for workers of type $s$.

# D Additional details on the data

## D.1 U.S. Census Longitudinal Business Database (LBD)

**Sample selection.** I follow a similar sample selection procedure to Barth et al. (2016). I drop observations with non-positive employment or payroll, as well as establishments with over 100,000 employees which are likely to capture miscoded records. I compute average wages as the ratio of total annual payroll to total establishment employment. I convert wages to 1982 dollars using the Consumer Price Index and exclude establishments that have an average wage less than half the yearly equivalent of the 1982 minimum wage of $3.35 an hour for a 40-hour week. I also omit firms in the utilities sector.

**Firms and firm-level industry code.** My definition of a firm follows the standard Census Bureau firm identifiers that link different establishments together based on IRS employer identification numbers (EINs) and ownership data from enterprise-level surveys. A multi-establishment firm is defined as having at least two establishments.

In the LBD, establishments are classified into industries, but multi-establishment firms do not have a unique industry identifier. I define the firm's industry according to the 4-digit NAICS code that accounts for the largest share of the firm's payroll.[46] For example, a firm's industry is classified as "Restaurants and Other Eating Places" if establishments with a NAICS code of 7225 constitute most of the firm's total wage bill.[47] Similarly, I define firm-level sector as the 2-digit NAICS sector that accounts for the largest part of its payroll. In the example above, the firm's sector would be "Accommodation and Food Services" (NAICS 72). To classify firms into goods and services, I follow the standard Bureau of Economic Analysis (BEA) definitions for "goods-producing industries" and "services-producing industries".[48] I define a firm as a "service firm" if establishments in services-producing sectors account for at least half of its total payroll.

**Geography.** In all data sources, I define a region or a local labor market according to 1990 commuting zones (CZs) as in Tolbert and Sizer (1996). In the model, I consider an aggregation of these commuting zones that groups together small neighboring commuting zones. I focus on the contiguous U.S., excluding Alaska, Hawaii and the American territories, yielding a total of 722 commuting zones that aggregate into 200 local markets.

[46] Throughout the paper, I employ the longitudinally consistent industry codes from Fort and Klimek (2018) that address changes in U.S. industry classification schemes over time.

[47] I construct a separate category for firms without a clear industry classification when no single industry covers at least 40% of the firm's payroll. This group of firms accounts for only a small share of total employment and wage bill, so most firms in the data have a clear firm-level industry identifier.

[48] "Goods-producing industries" include agriculture, mining, construction and manufacturing. "Services-producing industries" include all other sectors with the exception of utilities which are excluded from the analysis.

## D.2 Dun & Bradstreet

I use the Dun & Bradstreet Historical Records which provide data on U.S. private and public companies going back to 1969, with the exception of 1981 and 1984. Each establishment in a multi-establishment firm is linked to its headquarters establishment, and therefore it is possible to infer the headquarters location for all establishments in the data. I omit observations with missing firm linkages.

This data source has pros and cons relative to the Census LBD. For data on employment, it is less accurate than the LBD – especially for data on establishment-level employment – and it lacks data on wages. However, a key advantage relative to the LBD is that it clearly distinguishes between firm headquarters and branches, and it has a headquarters identifier for each firm. Therefore, I use this data to compute cross-region headquarters-branch linkages. See Barnatchez et al. (2017) for additional discussion of this dataset, including evidence on its good coverage of the spatial allocation of firms.

**Sample selection.** I drop observations with missing geographical data; missing industry; and missing data on establishment type and firm linkages. I omit establishments in Public Administration and other selected industries that are beyond the scope of the LBD dataset such as Membership Organizations, U.S. Postal Service, Federal and Federally-Sponsored Credit Agencies, Private Households. I also drop Educational Services, since it is hard to distinguish between privately-owned and government-owned establishments.

## D.3 Lightcast

Data on online job postings is obtained from Lightcast, a business analytics company, which extracts information from the near-universe of online job postings from a variety of online sources such as job boards and company websites. Lightcast employs a designated algorithm to avoid double counting of postings across multiple sources. The data covers 2010-2024, although I focus mostly on pre-Covid years 2010-2019. It includes extracted information on employer, job location, occupation, education-requirement, and for approximately a fifth of total observations also posted salaries.

In these data, a firm is defined based on the set of job postings that share the same codified employer name.[49] The Lightcast data lacks establishment identifiers, but the geographic location of the posting is known, allowing me to compare firms' postings across different commuting zones. These data are only available since 2010, but it is nevertheless useful for understanding how multi-location service firms are structured.

To distinguish firm headquarters from branches, I merge the Lightcast data with Dun & Bradstreet firm geography using name and location matches. The match is performed year by year, with the goal of retrieving headquarters FIPS codes from the Dun

[49]Lightcast claim to invest much effort into name codification to ensure that they capture the same entities.

& Bradstreet data. To standardize names across datasets, I apply the `stnd_compname` command in Stata, which removes nonessential words while retaining the core company name. Matching proceeds in two stages: first, through exact name matches; and second, through a fuzzy match that requires the first seven letters to coincide and allows a maximum string distance of two, implemented using the `stringdist` command in R. Table 7 shows the number of matches by year, with the number of matched firms ranges between 22,000 and 60,000 per year for the period 2010-2019.

**Table 7:** Match counts by year

| Year | Exact match | Fuzzy match |
|---|---|---|
| 2010 | 22,240 | 3,872 |
| 2011 | 27,201 | 4,835 |
| 2012 | 23,879 | 3,984 |
| 2013 | 30,451 | 5,416 |
| 2014 | 30,650 | 5,472 |
| 2015 | 38,992 | 7,514 |
| 2016 | 35,801 | 6,730 |
| 2017 | 34,784 | 6,181 |
| 2018 | 51,536 | 12,176 |
| 2019 | 58,999 | 14,359 |

# E Additional empirical and quantitative results

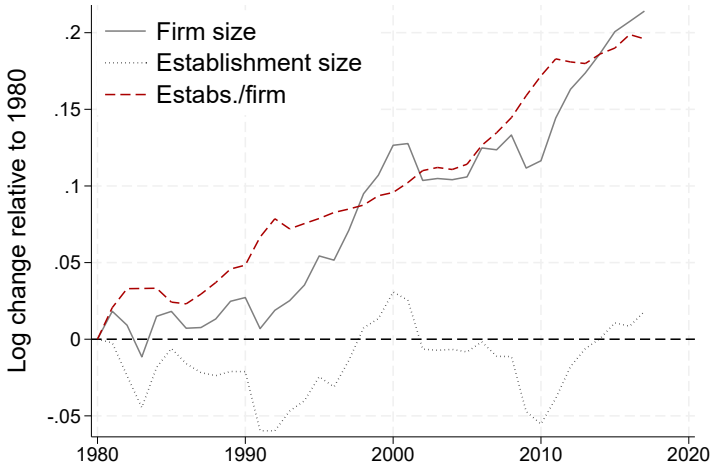
## E.1 Additional information on firm expansion (Section 2.2)

**Table 8: Statistics on firm spatial expansion**

| Type of firm | Year | Firms | % of workforce | Emp. per firm | Estabs. per firm | CZs per firm |
|---|---|---|---|---|---|---|
| Goods-producing firms | 1980 | 639800 | 31% | 38.0 | 1.2 | 1.1 |
| Goods-producing firms | 2017 | 778400 | 14% | 23.3 | 1.1 | 1.1 |
| Services, single estab. | 1980 | 2407000 | 40% | 12.7 | 1.0 | 1.0 |
| Services, single estab. | 2017 | 3769000 | 38% | 13.1 | 1.0 | 1.0 |
| Services, multi-estab. and single-CZ | 1980 | 69000 | 5% | 53.2 | 2.9 | 1.0 |
| Services, multi-estab. and single-CZ | 2017 | 72500 | 7% | 125.3 | 3.5 | 1.0 |
| Services, multi-estab. and multi-CZ | 1980 | 43500 | 24% | 432.6 | 12.2 | 5.0 |
| Services, multi-estab. and multi-CZ | 2017 | 72500 | 41% | 748.2 | 20.4 | 6.4 |

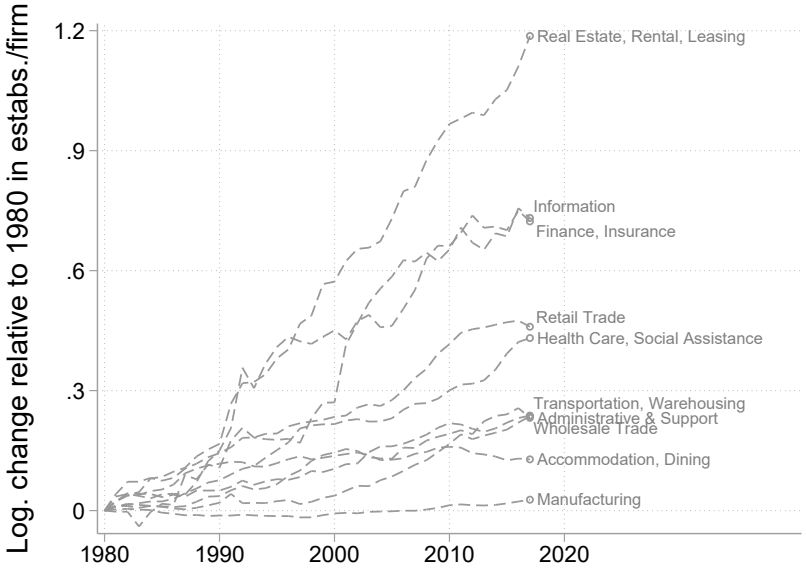
Note: Data from the Census Bureau Longitudinal Business Database. Goods- and services-producing firms stand for firms with at least half of their wage bill across all establishments in goods- and services-producing sectors, respectively. Multi-estab. stands for a firm with at least two establishments, and multi-CZ stands for a firm with establishments in at least two 1990 commuting zones

**Figure 11:** Firm expansion through the extensive margin




Note: This figure shows that changes in the average number of establishments per firm account for all the increase in average firm size (defined as employment per firm) since 1980, while average establishment size (employment per establishment) remained constant. Data from the U.S. Census Business Dynamics Dataset for firms with at least five employees.

**Figure 12:** Firm expansion by sector




Note: This figure shows the log change in the average number of establishments per firm relative to 1980 by economic sector, for firms with at least 5 employees. Data from the Business Dynamics Statistics dataset.

**Table 9: Additional statistics on firm structure in goods-producing sectors**

| Type of firm | Year | Firms | % of workforce | Emp. per firm | Estabs. per firm | CZs per firm |
|---|---|---|---|---|---|---|
| Goods, single estab. | 1980 | 623000 | 11% | 49 | 1.00 | 1.00 |
| Goods, single estab. | 2017 | 762000 | 7% | 38 | 1.00 | 1.00 |
| Goods, multi-estab. and single-CZ | 1980 | 6900 | 1% | 210 | 1.78 | 1.00 |
| Goods, multi-estab. and single-CZ | 2017 | 5400 | 1% | 230 | 1.03 | 1.00 |
| Goods, multi-estab. and multi-CZ | 1980 | 9900 | 19% | 8942 | 69.66 | 18.73 |
| Goods, multi-estab. and multi-CZ | 2017 | 11000 | 6% | 3535 | 48.22 | 11.50 |

Note: Data from the Census Bureau Longitudinal Business Database. The table includes statistics for firms with at least half of their wage bill in establishments in goods-producing sectors. Multi-estab. stands for a firm with at least two establishments, and multi-CZ stands for a firm with establishments in at least two 1990 commuting zone.

## E.2 Additional results on decompositions of wage inequality (Section 2.3)

### E.2.1 Comparison of rising inequality across workers and across establishments

I now establish that inequality across establishments accounts for most of the increase in overall inequality across workers in the economy. To this end, Table 10 compares the change in the variance of log wages between 1980 and 2017 across multiple data sources. The first row in Table 10 shows the change in this variance using data on individual-level reported wage earnings in the Decennial Census and the American Community Survey. Between 1980-2017, this variance has increased by 0.17 points. Similar magnitudes have been found by Song et al. (2019) using data from the U.S. Social Security Administration, and by Barth et al. (2016) using data from the BLS Current Population Survey (CPS). The second row in Table 10 repeats this exercise for establishment-level wages in the LBD, weighting different establishments by their employment. Evidently, both sources yield the same increase in the variance of log wages. Formally, let $w_i$ be the log of earnings for individual $i$, and let $e(i)$ be the establishment that employs individual $i$. By the law of total variance, the total change in variance equals to the sum of change in variance across establishments and the average change within establishments,

$$\underbrace{\Delta \text{Var}\left[w_i\right]}_{\text{Total change in variance}} = \underbrace{\Delta \text{Var}\left[\mathbb{E}\left[w_i | e\left(i\right)\right]\right]}_{\text{Change in variance across establishments}} + \underbrace{\Delta \mathbb{E}\left[\text{Var}\left[w_i | e\left(i\right)\right]\right]}_{\text{Change in variance within establishments}} .$$

Both the left-hand-side of this equation and the first term on its right-hand-side equal to approximately 0.17. Therefore, the last term that captures rising within-establishment variance is found to be close to zero.

The second Column in Table 10 repeats these findings for within-industry inequality, by demeaning detailed industry fixed-effects from log individual and establishment wages before computing the increase in the variance. Again, I find similar magnitudes across the Census/ACS and the LBD datasets, reinforcing the small role of within-establishment variance in the overall increase in inequality. Overall, rising within-industry wage dispersion accounts for slightly over half of the total increase in inequality.[50] Similar magnitudes for the role of within-industry trends are obtained when utilizing data from the CPS.[51]

[50] Haltiwanger et al. (2022) find that a slightly higher share – around 60% – of the rise in total inequality in the LEHD is due to differences between industries starting from the late 90s. Indeed, I find that the importance of differences between industries becomes larger in the latter part of the sample, reconciling the figures reported here with their findings.

[51] Comparing these moments to the Social Security Administration data as in Song et al. (2019) is infeasible due to the lack of high-quality industry identifiers in that data.

**Table 10: Rising variance of log wages in the U.S. economy**

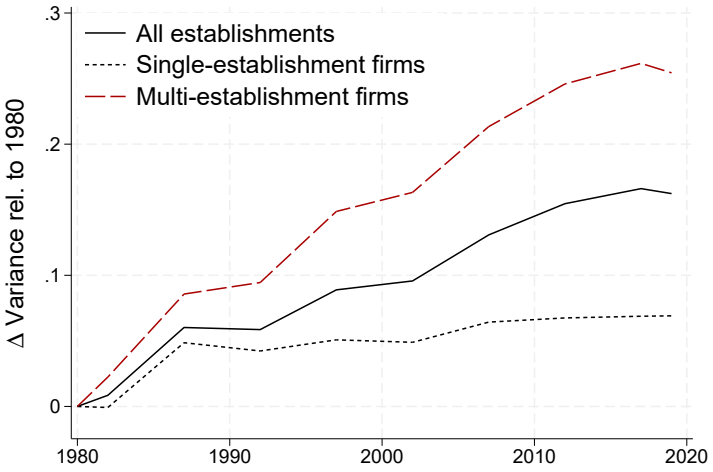
**Change in the variance of log wages (1980-2017)**

| Data | Total change | Within-industry change |
|---|---|---|
| Decennial Census / ACS - across U.S. workers | 0.17 | 0.1 |
| LBD - across all establishments | 0.17 | 0.09 |

Note: This table summarizes changes in the variance of log wages in the U.S. economy between 1980-2017. The first row computes this measure from individual-level reported wage earnings in the Decennial Census and the American Community Survey. The second row reports this measure for average establishment-level payroll in the LBD. The first column shows the overall change, and the second column reports this change after controlling for industry-level fixed effects.

### E.2.2 The role of multi-establishment firms in the rise of inequality – without industry fixed effects

**Figure 13:** The role of multi-establishment firms in the rise of inequality - without industry fixed effects




Note: This figure shows changes in the employment-weighted variance of log average payroll across establishments in the Longitudinal Business Dataset (LBD) in selected years relative to 1978: for all establishments (solid-black line), for multi-establishment firms (dashed-red line) and for single-establishment firms (dotted-gray line). Relative to Figure 2 in the paper, this figure shows this decomposition without first demeaning industry fixed effects.

### E.2.3 Occupational headquarters intensity and the rise of wage inequality

I now investigate how much of the overall increase in the variance of log wages in the economy can be accounted for by the distinction between headquarters and branch occupations. To this end, I decompose the overall increase in variance in the Census-ACS data, as in the above. I merge the headquarters-intensity measures from Section 2.1 to the Census-ACS data using workers' occupation codes,[52] and then compute how much of the

[52]For this exercise, I employ time-invariant classifications of occupations, utilizing the *occ1990dd* classification from Autor and Dorn (2013) and Autor et al. (2019). See also https://www.ddorn. net/data.htm.

**Table 11:** Decomposition of the rise in variance of individual earnings

| Component of log individual earnings | Δ Variance 1980-2017 | % of total Δ |
|---|---|---|
| HQ-intensity of the occupation | 0.08 | 46% |
| Commuting zone and college attainment | 0.03 | 18% |
| Covariance | 0.03 | 20% |
| Residual | 0.03 | 16% |
| **Total** | **0.17** | **100%** |

Note: this table shows a variance decomposition of individual wage earnings from the 1980 Decennial Census and 2015-2019 American Community Survey. Individual earnings are regressed on the interaction of commuting-zone and college attainment fixed effects, and on the continuous measure of occupational headquarters intensity from Section 2.1 of the paper. The table shows changes over time in the variance for the predicted part of earnings based on each of these components, and the residual change in variance.

rise in variance is explained by the increasing importance of this measure between 1980-2017. Specifically, I project log individual wage earnings on my HQ-intensity measure and other selected individual characteristics, and compute the share of the overall rise in variance this measure accounts for. I include in this projection a series of fixed effects capturing individual educational attainment, geography (commuting zones), and their interactions, in order to highlight the strength of my HQ-intensity measures on top of these other factors. Results from this variance decomposition can be seen in Table 11. The overall increase in the variance of log wages in the economy over the covered period stands at 0.17 points, similarly to the increase across the universe of establishments in the LBD. Out of this increase, 46% can be accounted for by my occupational HQ-intensity measure; 18% by the combination of education and geography; and additional 20% by their covariance. Of course, this is only an accounting decomposition, and the substantial role for my HQ-intensity measures could be reflecting other mechanisms that are not related to firm expansion. However, it provides further evidence that the HQ-branch distinction seem to be quantitatively important for the question of rising wage inequality.

I merge my headquarters intensity measures to this classification using a concordance between the 158 occupational codes in the Lightcast data and the *occ1990dd* codes.

## E.3 Classifying O*NET tasks (Section 4)

The following prompt is used as an input to a large language model (LLM) to classify the outputs of O*NET tasks as non-rival, rival, or ambiguous.

You are analyzing work tasks from ONET in the context of a firm with many branches. You need to determine whether the output of each ONET task is non-rival across branches. Non-rivalry occurs when a task output can be used simultaneously across multiple firm locations without diminishing its value, availability, or functionality at any single location. Consider outputs like standardized procedures, software code, training materials, analytical frameworks, or digital templates. In contrast, rival outputs include location-specific services, physical products, or tasks that require dedicated resources that cannot serve multiple locations simultaneously. For example, the output of a manager can be non-rival if it is dedicated to developing training program that can be deployed across all branches, while it can be rival if it is dedicated to a conference call with a team from one branch.

Classification Instructions: For each work task description provided, classify the output using exactly one of these categories: NON-RIVAL, RIVAL, or AMBIGUOUS (insufficient information in the task description to make a clear determination about rivalry characteristics).

## E.4 Additional results for Section 4.2

### E.4.1 Spatial disparities

In this section, I provide the regression tables with OLS and 2SLS results for the analysis from Section 4.2 on firm expansion and spatial disparities.

The model yields several implications for how the expansion of locally headquartered firms affects regional outcomes. Markets where such firms expand more should see a relative increase in expenditure on headquarters jobs within the services sector. These markets should also exhibit increased specialization in services relative to goods production, implying a decline in the manufacturing employment share. Finally, because this shift amounts to a positive local labor demand shock, both regional wages and total employment should rise.

Table 12 presents evidence consistent with these predictions. The key explanatory variable is the log number of commuting zones (CZs) served by firms headquartered in a given market in 1980. For example, because Walmart Corporation was headquartered in Benton County, AR in 1980, its subsequent expansion into additional commuting zones between 1980 and 2017 contributes to Benton's average measure of local firm expansion. Columns 1–4 correspond to the following outcomes, all in logs: the headquarters share of total payroll in services, constructed from the payroll share in the Decennial Census or American Community Survey of the most headquarters-intensive occupations (as defined

in Section 2.1); the manufacturing employment share from the County Business Patterns; average wages from the BLS Quarterly Census of Employment and Wages (QCEW); and total employment from the QCEW. The sample includes all commuting zones in the contiguous U.S. with at least 10,000 residents in 1980, in order to reduce measurement error and avoid comparisons with very rural areas. The analysis covers five benchmark years: 1980, 1990, 2000, 2010, and 2017. For the regression of the headquarters payroll share, I exclude outliers where the measure equals 1, equals 0, or is missing.

The OLS results (Panel A) show that local markets with greater expansion of locally headquartered firms experienced reallocation to headquarters activity within the services sector, reallocation away from manufacturing, and an increase in regional wages and employment, all consistent with the model. All results are quantitatively meaningful and statistically significant at standard confidence levels.

To isolate a causal effect of local firm expansion on such labor market outcomes, I instrument the expansion of locally headquartered firms using ex-ante exposure to national industry expansion trends. Specifically, for each labor market area $i$ and period $t$, I construct the following shift-share measure of exposure to firm expansion:

$$EE_{it} \equiv \sum_{k \in \text{Industries}} \left( \frac{N^{HQ}_{ik,1980}}{\sum_{k'} N^{HQ}_{ik',1980}} \right) \times \log \left( \frac{\overline{\text{locations}}_{k,t}}{\overline{\text{locations}}_{k,1980}} \right).$$

Here, $N^{HQ}_{ik,1980}$ is the number of firm headquarters located in region $i$ for industry $k$ in 1980 (from the Dun & Bradstreet dataset), and $\overline{\text{locations}}_{k,t}$ is the national average number of locations per firm in industry $k$ at time $t$. The logic is that regions hosting headquarters of industries that expanded more nationally after 1980 were more "exposed" to spatial firm expansion by their locally headquartered firms, regardless of other local conditions. This approach follows the shift-share (Bartik) tradition (Bartik, 1991).

The 2SLS results (Panel B) confirm the OLS patterns: expansion of locally headquartered firms significantly increased headquarters payroll shares, reduced manufacturing employment shares, and raised both wages and employment. The first-stage KP-Wald F-statistics, reported at the bottom of the table, are around conventional thresholds, indicating relevance of the instrument. The magnitudes are higher, in line with the expected attenuation bias resulting from measuring expansion of locally headquartered firms using the Dun & Bradstreet data.

Finally, Panel C repeats the 2SLS analysis when controlling for time-varying regional employment shares in the information and financial sectors (NAICS 51 and 52), ruling out concerns that the results are driven by exceptional regional growth in these industries.

Overall, Table 12 provides strong evidence consistent with the model's predictions: regions more exposed to expansion of their locally headquartered firms experienced a reallocation toward headquarters services, away from manufacturing, accompanied by

regional gains in wages and employment.

**Table 12:** Expansion of locally-headquartered firms and regional labor market outcomes

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Outcome: (log) | HQ share | Manuf share | Average wage | Total employment |
| **Panel A: OLS** | | | | |
| Log # CZs for locally HQed firms | 0.0909*** | -0.0917*** | 0.0476*** | 0.0707*** |
| | (0.0197) | (0.0254) | (0.0113) | (0.0225) |
| **Panel B: 2SLS** | | | | |
| Log # CZs for locally HQed firms | 0.613** | -0.942*** | 0.391*** | 1.136*** |
| | (0.264) | (0.292) | (0.134) | (0.328) |
| First-Stage F-stat (KP-Wald) | 7.761 | 14.38 | 14.38 | 14.38 |
| **Panel C: 2SLS + Controls** | | | | |
| Log # CZs for locally HQed firms | 0.596** | -0.890*** | 0.378*** | 1.175*** |
| | (0.265) | (0.285) | (0.134) | (0.347) |
| First-Stage F-stat (KP-Wald) | 7.586 | 13.39 | 13.39 | 13.39 |
| Year and CZ fixed effects | Y | Y | Y | Y |
| Observations | 1925 | 2499 | 2499 | 2499 |

*Notes:* Local labor market outcomes against expansion of locally-headquartered firms. Each column shows the estimated coefficient from a regression of a different labor market outcome for a given local market against the log number of locations that are operated by the firms that were headquartered there in 1980. Columns 1-4 include as outcome variables the logs of headquarters share of total payroll in services, manufacturing share of total employment, average wages, and total employment, respectively. Observations include five years (1980,1990,2000,2010,2017) and commuting zones with at least 10,000 employees in 1980. Observations number is smaller in column 1 since missing values or values that equal 1 or 0 are dropped. The other columns rely on administrative data. Panel A shows the OLS results, and Panel B instruments the expansion of locally-headquartered firms with the instrument described in Section E.4. Panel C shows the 2SLS results when also controlling for the time-varying regional employment share in the information and financial sectors (NAICS 51 and 52, respectively). Standard errors are clustered at the commuting zone level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

### E.4.2 Long-run occupational and industry dynamics

In this section, I test whether long-run wage growth has been concentrated in more headquarters-intensive occupations, and whether this effect is particularly pronounced in sectors that underwent greater spatial expansion. To do so, I draw on the measure of occupational headquarters intensity introduced in Section 2.1. I combine this measure with long-term changes in occupation-level wages using data from the Census Bureau's Decennial Census and the American Community Survey. The empirical exercise consists of regressions of changes in occupational outcomes at the industry–occupation level between

1980 and 2017 on the headquarters intensity measure.[53]

Table 13 reports the results. Column (1) shows that relatively more headquarters-intensive occupations experienced larger increases in total payroll, implying greater relative sectoral expenditure on headquarters labor once sector fixed effects are accounted for. Columns (2) and (3) reveal similar positive associations for log average wages and for residual wages (controlling for individual demographics and location), indicating greater wage dispersion between headquarters and branches. Columns (4)–(6) extend the analysis by interacting headquarters intensity with an indicator for sectors in the top quartile of spatial expansion. Consistent with the model's predictions (Propositions 1–3), sectors that expanded more geographically exhibit larger relative increases in both wages and residual wages of headquarters-intensive occupations. Moreover, consistent with substitutability between headquarters and branch labor ($\eta > 1$), these sectors also display greater reallocation of labor costs toward headquarters-intensive work.

**Table 13: HQ-branch dispersion based on occupational classifications**

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Outcome: 1980-2017 changes in - | Pay share | Log wage | Log r. wage | Pay share | Log wage | Log r. wage |
| Occupational headquarters intensity | 0.0912*** | 0.519*** | 0.0809** | 0.0401** | 0.362*** | -0.0325 |
| | (0.0263) | (0.0507) | (0.0390) | (0.0195) | (0.0593) | (0.0491) |
| Headquarters intensity × high expansion | | | | 0.0963** | 0.295*** | 0.214*** |
| | | | | (0.0467) | (0.0936) | (0.0780) |
| Industry fixed-effects | Y | Y | Y | Y | Y | Y |
| Number of observations | 13325 | 13325 | 13325 | 13325 | 13325 | 13325 |
| R-squared | 0.206 | 0.422 | 0.253 | 0.213 | 0.426 | 0.256 |

Note: Changes in occupational wages and payroll-shares between 1980-2017 as a function of headquarters intensity. Each cell shows the estimate from a regression of changes in industry-occupation level outcomes against the headquarters intensity of the occupation and against an interaction of this measure with an indicator for high-expansion industry. Each column captures a different industry-occupation-level outcome: changes in within-industry occupational payroll shares (columns 1 and 3); changes in log average wages (columns 2 and 4); and changes in log residual wages (columns 3 and 6). All changes are between 1980-2017. The measurement of occupational headquarters intensity is described in Section 2.1. High-expansion industries are those with the greatest increase in average establishments per firm (the top quartile of this distribution). Observations are weighted by total employment. All columns control for industry fixed effects. Occupations and industries are classified in accordance to the *occ1990dd* and *ind1990ddx* classifications from Autor and Dorn (2013) and Autor et al. (2019), which can be found in https://www.ddorn.net/data.htm. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

[53]For this exercise, I employ time-invariant classifications of industries and occupations, utilizing the *occ1990dd* and *ind1990ddx* classifications from Autor and Dorn (2013) and Autor et al. (2019). See also https://www.ddorn. net/data.htm.

### E.4.3 Pass-through from firm growth to HQ payroll and wages in the LBD

In this Section I study how the relative expenditure on headquarters labor and the relative wages in headquarters establishments change when firms expand in space using the LBD data. Recall that one key prediction in the model is that as firms expand, then, other things equal, they increase the relative wages in their headquarters jobs relative to branch-level jobs. Moreover, if $\eta \neq 1$, then also the relative expenditure on headquarters jobs rises. Here, I study these relationships using the Census LBD data, leveraging the alternative measurement of headquarters activity using the variation in establishments' industry classifications within firms (see Section 2.1 for details).

Specifically, the analysis regresses measures of relative headquarters intensity on the firm's log number of establishments, using variation both across firms and within firms over time. The outcome variables are the log ratio of headquarters to branch payroll (Column 1) and the log ratio of headquarters to branch average wages (Column 2). Columns 3 and 4 repeat the specifications with firm fixed effects, year and sector-by-year fixed effects, and controls for regional wage differences across the locations where firms operate, as required in the model (see Proposition 2). The sample is restricted to Economic Census years from 1997 to 2017, and standard errors are clustered at the firm level. Establishments in the LBD are classified as headquarters or branches based on their industry codes, as described in Section 2.1.

**Table 14: HQ-branch dispersion based on industry classifications**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Outcome: Log ratio between HQ and branches | Payroll ratio | Wage ratio | Payroll ratio | Wage ratio |
| Log # of establishments | 0.305*** | 0.0394*** | 0.153*** | 0.0403** |
| | (0.0152) | (0.00526) | (0.0325) | (0.0123) |
| Observations (rounded) | 29500 | 29500 | 29500 | 29500 |
| Year and firm FE | Y | Y | Y | Y |
| Year × sector FE | N | N | Y | Y |
| Controls | N | N | Y | Y |

Note: Relationship between firms' scope and within-firm headquarters-branch dispersion. Each observation is a firm at a particular year. Column 1 shows the estimated coefficient from a cross-sectional regression of the log of the firm's payroll ratio between its headquarters and branch establishments, as defined in Section 2.1, against the firm's log number of establishments. Column 2 replaces the outcome variable with the log ratio of average payroll per worker between headquarters and branch establishments. Columns 3 and 4 repeat columns 1 and 2 when controlling for firm fixed effects, year fixed effects, year×sector fixed effects, and controls for differences in average regional wages across firms' locations. Standard errors are clustered at the firm level. The sample includes years from Economic-Census years starting from 1997 (i.e. 1997,2002,2007,2012,2017). $^{*}$ $p < 0.1$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$.

The results reveal a positive relationship between firm scope and within-firm headquarters-branch payroll and wage dispersion. In the cross-sectional specifications, a one log-point increase in the number of establishments is associated with a 0.31 log-point increase in

the payroll ratio between headquarters and branches, and a smaller but still significant 0.04 log-point increase in the wage ratio. When adding firm and sector-year fixed effects, the payroll ratio estimate is attenuated but remains positive and significant at 0.15, while the wage ratio estimate stays close to 0.04.

## E.5 Additional information for the quantified model

### E.5.1 Skill groups and skill intensities

**Definition of skill groups.** I define four skill groups by dividing workers into equally-sized bins in the 1980 Decennial Census based on the average skill requirement of their occupation. The skill requirement is measured using the share of college graduates in each occupation. I use the same allocation of occupations to skill groups when computing statistics for the skill groups in 2017. This division into skill-groups has a few advantages relative to categorizing individuals based on their educational attainment. First, the demographic composition of different education groups has changed dramatically over time; this is not the case for my four groups of occupations. Second, the relative size of these groups remained relatively constant over time, allowing me to abstract from changes in the aggregate supply of skill. Finally, it allows me to consider more skill groups. In support of this approach, the average college intensity of a worker's occupation is a stronger predictor of their wage than their own educational attainment.

**Skill intensities.** I also recover the skill elasticities in production $\alpha_{hs}, \alpha_{\ell s}, \alpha_{gs}$ directly from the data, by computing the payroll-share of each group in all three activities. To compute $\alpha_{hs}$ and $\alpha_{\ell s}$, I utilize the occupational headquarters-intensity measures from Section 2.1, and the fact that each occupation is classified into a single group, delivering the intensity of each skill group $s$ in headquarters and branch production. For $\alpha_{gs}$, I take the payroll share of each group in goods-producing sectors. The values for all skill intensities can be found in Table 15.

| Skill group | $\alpha_{h,s}$ | $\alpha_{\ell,s}$ | $\alpha_{g,s}$ |
|---|---|---|---|
| 1 | 0.07 | 0.16 | 0.37 |
| 2 | 0.11 | 0.19 | 0.19 |
| 3 | 0.22 | 0.26 | 0.19 |
| 4 | 0.59 | 0.39 | 0.25 |

**Table 15:** Calibrated values for skill-intensities in the headquarters labor bundle ($\alpha_{h,s}$), branch-level labor bundle ($\alpha_{\ell,s}$), and tradable-goods labor bundle ($\alpha_{g,s}$).

### E.5.2 Headquarters-branch gravity

In this Section I document the strong gravity pattern of headquarters-branch linkages and how it has changed over time. Denote by $M_{ij}$ the number of establishments in location $j$ that belong to $i$-headquartered firms. Table 16 reports the results from a gravity regression of $M_{ij}$ against log bilateral distance, controlling for origin and destination fixed effects, where each observation corresponds to a pair $ij$ of the 200 local markets in my quantification.[54] I also include a dummy variable for whether the two markets share state

[54] I implement this procedure using the Dun & Bradstreet data, which has the key advantage that every multi-establishment firm has a clear headquarters location. In contrast, in the LBD, headquarters

borders, to account independently for state border effects. Column (1) shows results for 1980 from an OLS regression that sets the log of $M_{ij}$ as the outcome variable, abstracting from $ij$ pairs with zero linkages. Column (2) estimates this relationship using the Poisson Pseudo-Maximum Likelihood estimator from Silva and Tenreyro (2006), allowing for zero bilateral linkages and encompassing all observations. In both specifications, the coefficient on distance is negative and statistically significant, at around unity. Notably, the coefficient on distance has declined over time as can be seen in Columns (3) and (4).

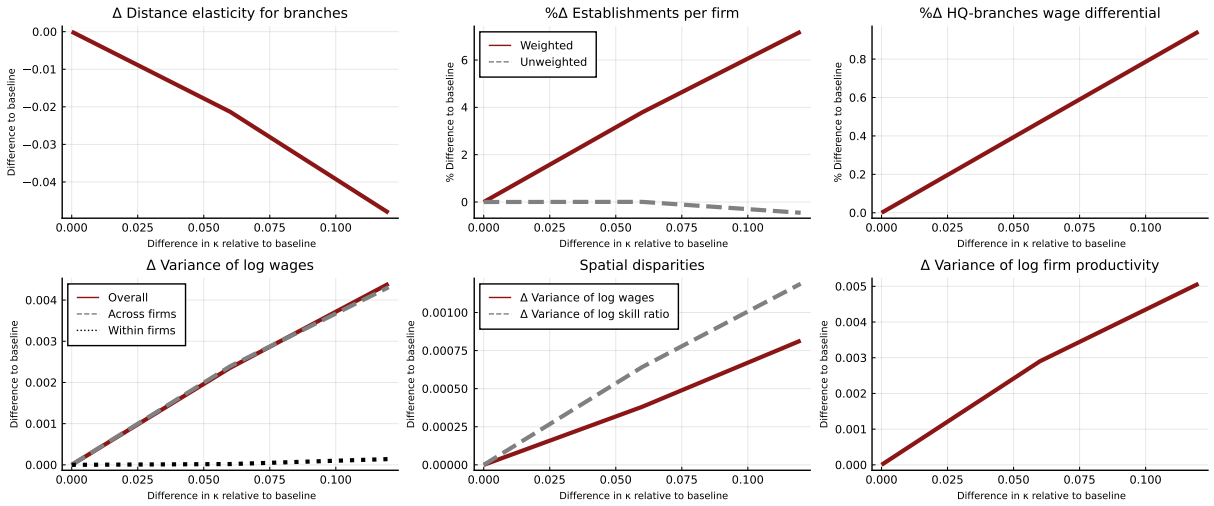
**Table 16: Gravity for cross-region firm linkages**

| | (1) | (3) | (2) | (4) |
|---|---|---|---|---|
| | 1980 | | 2017 | |
| Outcome: number of cross-region establishments | Logs | Levels | Logs | Levels |
| Log distance | -0.956*** | -0.981*** | -0.777*** | -0.608*** |
| | (0.0135) | (0.0322) | (0.0128) | (0.0165) |
| Different state | -1.132*** | -1.281*** | -1.032*** | -0.873*** |
| | (0.0263) | (0.0445) | (0.0246) | (0.0334) |
| Origin and destination fixed effects | Yes | Yes | Yes | Yes |
| Observations | 11451 | 39800 | 11451 | 39800 |
| R-squared | 0.688 | | 0.767 | |
| Pseudo R-squared | | 0.790 | | 0.838 |

Note: This table shows results from the regression of the number of bilateral headquarters-branch linkages across pairs of local markets against the distance between the markets and a dummy that equals 1 if the markets do not share state borders. Each observation is a pair out of 200 local U.S. labor markets as defined in Section 5.2. Columns (1) and (3) are estimated in logarithms by ordinary least squares (OLS), and Columns (2) and (4) are estimated in levels by Poisson pseudo maximum likelihood (PPML). All specifications include headquarters-location and branch-location (origin and destination) fixed effects. Columne (1) and (3) show results for 1980, Columns (2) and (4) show results for 2017. $^{*}$ $p < 0.1$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$.

location cannot be inferred for a large share of firms, especially in earlier periods. The Dun & Bradstreet data is well-suited for this purpose, since it puts great emphasis on documenting firm linkages, and since the spatial distribution of firms and establishments is one of the dimensions in which it compares well to administrative datasets. See Barnatchez et al. (2017) for additional details.

**Figure 14: Labor market implications of lower state-border effects**




Note: Distributional implications of lower state border effect. The x-axis depicts changes in the state border effect in the cost of expansion, $\kappa$, expressed in terms of the difference from the baseline equilibrium (with zero representing the baseline equilibrium). Positive values indicate lower state border effect in absolute value. The y-axis in each subplot captures changes in selected moments of interest relative to the baseline equilibrium, after re-computing the equilibrium for every new value of $\rho$: change in the gravity coefficient of headquarters-branch linkages (top-left panel); % change in average establishments per firm (employment-weighted and unweighted, top-middle); % change in mean wage dispersion between headquarters and branches (top-right); change in variance of log wages (overall, across firms, and within firms; bottom-left); change in the variance of log wages and the variance of high-skilled to low-skilled employment (bottom-middle); change in the variance of log firm productivity (bottom-right).