Opportunity & Inclusive Growth INSTITUTE

FEDERAL RESERVE BANK OF MINNEAPOLIS

INSTITUTE WORKING PAPER
No. 25

Urban-Biased Growth: A Macroeconomic Analysis

Revised September 2022

Fabian Eckert
University of California, San Diego

Sharat Ganapati
Georgetown University

Conor Walsh
Columbia University

DOI: https://doi.org/10.21034/iwp.25
Keywords: Urban growth; High-skill services; Technological change
JEL classification: J31, O33, R11, R12

Urban-Biased Growth: A Macroeconomic Analysis*

Fabian Eckert† Sharat Ganapati‡ Conor Walsh§

First Version: March 2020
This Version: September 2022

Abstract

Since 1980, US wage growth has been fastest in large cities. Empirically, we show that most of this urban-biased growth reflects wage growth at large Business Services firms, which are also the most intensive users of ICT capital in the US economy. We provide an explicit economic mechanism whereby ICT is more complementary with labor at larger firms. Quantitatively, we find that with such a complementarity, the observed decline in ICT prices alone can account for most of the urban-biased growth, since Business Services firms in big cities tend to be large.

Keywords: Urban Growth, High-skill Services, Technological Change
JEL Codes: J31, O33, R11, R12

*We thank Fabian Trottner for many in-depth discussions. We also thank Pol Antràs, David Autor, Costas Arkolakis, Gideon Bornstein, Laura Castillo-Martinez, Jonathan Dingel, Pierre-Olivier Gourinchas, Gordon Hanson, J. Bradford Jensen, Tom Kemeny, Chris Moser, Michael Peters, Esteban Rossi-Hansberg, and Steve Redding for insightful comments. 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 2193 (CBDRB-P2193-R8942, R9405, R9629, and R10013). Eckert and Walsh thank the International Economics Section at Princeton University where some of this work was completed. The present version of this paper supersedes "Skilled Scalable Services: The New Urban Bias in Economic Growth."
†University of California, San Diego; fpe@ucsd.edu
‡Georgetown University; sharat.ganapati@georgetown.edu
§Columbia University; caw2226@columbia.edu

INTRODUCTION

Since 1980, the US economy has experienced urban-biased growth, with wages in large cities rising substantially faster than wages in smaller cities and rural areas. The left panel of Figure 1 shows average wages across US commuting zones ordered by density. In 1980, workers in the cities with the highest population density (New York and Chicago) earned, on average, 34% more than workers in cities with the lowest population density. By 2015, the gap had risen to around 62%.

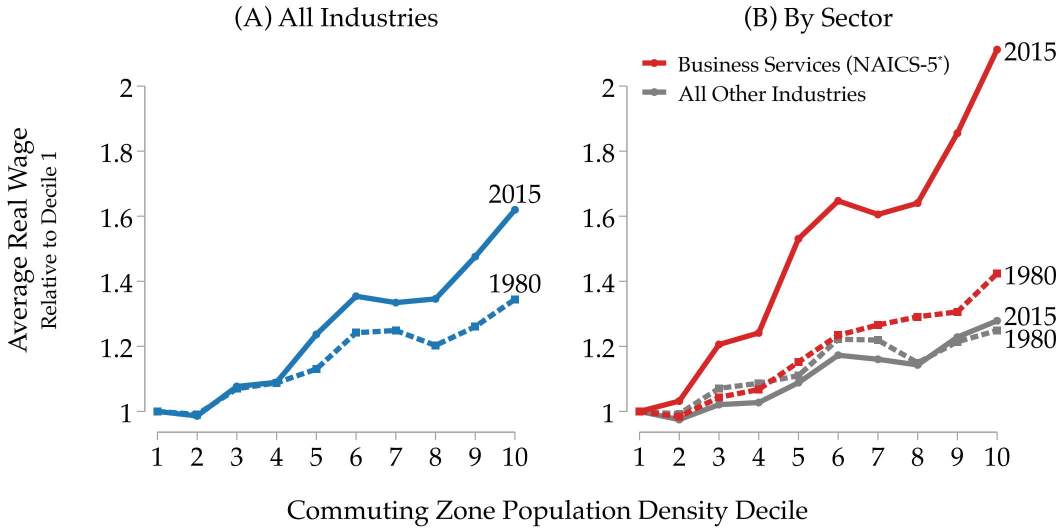
Urban-biased growth is related to many economic and societal challenges the US has faced in recent decades. It occurred alongside skyrocketing house prices in urban centers with high population density (Gyourko, Mayer, and Sinai, 2013), increasing political polarization between big cities and rural areas (Scala and Johnson, 2017), and rising income inequality (Piketty and Saez (2003)). However, a fundamental explanation of the origins of urban-biased growth has not been forthcoming. The contribution of this paper is to provide such an explanation using data and economic theory.

Empirically, we show that a single sector – Business Services – has been responsible for virtually all urban-biased growth since 1980. The right panel of Figure 1 shows average wages across commuting zones separately for the Business Services sector (NAICS-5) and the rest of the economy. In Business Services, in 1980, workers in cities with the highest population density earned on average 42% more than workers in cities with the lowest population density. Today, they make 111% more. At the same time, the relationship between wages and population density has changed little in the rest of the US economy.

Within the Business Services sector, average wage growth mainly reflects fast wage growth within education and occupation groups. Compositional changes are less important. While the share of college workers in the Business Services sector has risen faster than the aggregate college share, such "skill deepening" can explain only a small fraction of the sector's urban-biased wage growth. Notably, workers without a college degree in the Business Services sector *have* experienced urban-biased growth, whereas neither college nor non-college workers have experienced urban-biased wage growth outside the sector. Furthermore, the distribution of Business Service jobs across US cities has barely changed since 1980.

We interpret our findings as evidence of a labor demand shock emanating from

FIGURE 1: THE NEW URBAN BIAS



Notes: This figure shows average wages across commuting zones (Tolbert and Sizer (1996)) sorted into ten groups of increasing population density, relative to the group of commuting zones with the lowest population density. Each decile accounts for one-tenth of the US population in 1980. The first decile corresponds to 10 $people/mi^2$ and the tenth decile corresponds to 2300 $people/mi^2$. Data for average wages come from the US Census Bureau's Longitudinal Business Database. Business Services firms are firms in the NAICS-5 sector (excluding NAICS 56 and 53 for disclosure reasons). We define average wages as total payroll over total employment within a commuting zone and industry pair.

the Business Services sector of big US cities, starting around 1980. The shock had an important skill-biased component: while it raised all workers' wages, it tilted the sector's composition toward college-educated workers and raised their wages fastest. Recent work by Moretti (2013) supports our demand shock interpretation and suggests that "localized skill-biased technical change is a potential explanation, as long as it is enriched by a theory of why demand shocks occur in some cities and not in others." In the remainder of the paper, we provide such a theory.

A large literature links recent changes in the returns to skill in the US economy to advances in information and communication technologies (ICT) that particularly benefited skilled labor (see, e.g., Krueger (1993)). Krusell, Ohanian, Ríos-Rull, and Violante (2000) formalized this explanation as a "capital-skill complementarity" paired with declining capital prices. Through the lens of their neoclassical framework, observed declines in equipment (ICT) capital prices account for most of the increase in the college wage premium since 1980.[1]

[1] Krusell et al. (2000) use the decline of equipment capital prices of which ICT prices are a component; Eden and Gaggl (2019) show that the decline in equipment capital prices is in fact driven by the decline in the ICT prices on which we focus; see also Figure 5.

We argue that, with some modification, a capital-skill complementarity and the decline in ICT capital prices can also account for the urban-biased wage growth in Business Services.

The Business Services sector's intensive use of ICT capital in 1980 made it particularly exposed to changes in the price of ICT. Both in 1980 and in 2015, ICT capital accounted for a much larger share of value added in Business Services than in any other sector in the US economy. The use of ICT in Business Services also featured an urban bias. We use a firm-level survey to show that Business Services firms in big cities use much more ICT per worker than firms elsewhere and in other sectors.

In a cross-section of regions, a neoclassical capital-skill complementarity produces *slower* wage growth in locations with initially higher wages as ICT capital prices decline. The reason is that the capital-skill complementarity leads to lower capital cost shares wherever labor is expensive. As a result, high-wage locations are less exposed to capital price movements. When ICT becomes cheaper, *low-wage* locations see the largest cost savings and higher increases in labor demand. We refer to this mechanism as the "neoclassical channel"; its predictions are the opposite of the empirical patterns shown in Figure 1.

We argue that the neoclassical theory misses the role of ICT in helping large firms scale up their operations. Empirically, we show that large firms are more intensive ICT users than smaller firms and tend to locate in big cities. These patterns are substantially stronger in Business Services than in any other sector. We also show that large firms account for most Business Service employment in big cities.

We provide a one-parameter extension of the neoclassical theory that adds a firm-size-ICT complementarity. In our theory, locations differ in productivity. In equilibrium, more productive locations have higher wages, more employment, and larger firms. The neoclassical channel pushes for higher capital cost shares in low-productivity locations where wages are low. At the same time, the firm-size-ICT complementarity pushes for higher capital cost shares in more productive locations, since they host larger firms. If this "scale channel" dominates the "neoclassical channel," our theory predicts that capital cost shares are higher in locations with high wages. As a result, capital price declines generate faster labor demand growth in locations with high wages and large populations, in line with Figure 1.

We quantify the strength of our mechanism by calibrating a quantitative version of the model. The firm-size-ICT complementarity governs the relationship between firm size and ICT cost shares; without the complementarity, capital cost shares are the same among firms faced with the same input prices. Our calibration strategy exploits this and targets the empirical relationship between ICT usage and firm size to discipline the scale channel's strength. We discipline the strength of the neoclassical channel by ensuring that the model matches canonical estimates of the macro-elasticity of high-skill labor and capital.

We calibrate the model to the 1980 economy, choosing location productivities and amenities to match the cross-section of employment and wages across commuting zones. We then decrease the price of ICT capital in the model from its 1980 to its 2015 level and study the general equilibrium response of wages and employment across space, holding all other parameters fixed at their 1980 levels. We find that the observed ICT price decline can generate most of the urban-biased growth in the data. The model predicts that our mechanism is mainly active in the Business Services sector, since other sectors use much less ICT capital. The firm-scale-ICT complementarity is crucial in generating this result; without it, there would have been no urban-biased economic growth.

Literature Review Our paper makes an empirical and a theoretical contribution. On the empirical side, we use detailed data to show which sectors, firms, and worker types account, in a statistical sense, for the urban-biased growth of the US since 1980. A large, related literature studies *static* differences in skill types, industries, and occupations across locations (see Gaubert (2018), Davis and Dingel (2020), and Almagro and Domınguez-Iino (2019)). Another set of related papers studies the implications rather than the origins of urban-biased growth (see, e.g., Diamond (2016), Moretti (2013), and Hsieh and Moretti (2019)). Finally, a recent literature examines the "end of wage convergence" in the US, which is related to urban-biased growth since big cities had high wages to start with (e.g., Giannone (2022), Ganong and Shoag (2017), Berry and Glaeser (2005), Moretti (2012), and Rubinton (2019)), and the uneven growth of the college wage premium across US cities (see Beaudry, Doms, and Lewis (2010) and Eckert (2019)). In contrast to these papers, we highlight the essential role of Business Services in big city wage growth, and provide a formal statistical accounting of the sources of urban-biased growth.

A separate empirical contribution of our paper is to show that Business Services are the most intensive users of ICT capital and provide evidence that ICT use is increasing in firm and city size, especially in the Business Services sector.

A large set of papers study the role of increases in ICT usage for *skill-biased* wage growth in the US economy (see Krusell et al. (2000), Krueger (1993), and Lashkari, Bauer, and Boussard (2022)), yet our paper relates it to the *urban-biased* growth in recent decades. Beaudry et al. (2010) focus on the role of PC adoption in generating convergence in relative skill prices across cities.

On the theoretical side, we present an explicit economic mechanism for the origins of urban-biased growth and quantify its importance in the US economy. Our model can generate either wage convergence or divergence across regions depending on parameter values. An existing literature has studied wage convergence across regions (see Barro and Sala-i Martin (1992), Beaudry et al. (2010), and Desmet and Rossi-Hansberg (2009)); many fewer papers present mechanisms for wage divergence (e.g., Ganong and Shoag (2017)) and none of them quantifies their mechanism for the US economy since 1980. Technically, our paper embeds a non-homothetic production function (Sato (1977)) into the workhorse quantitative spatial model (Allen and Arkolakis (2014), Redding (2016), and Redding and Rossi-Hansberg (2017)) and highlights the interaction of the scale elasticity with firm size differences across regions.[2]

1. Urban-Biased Growth

In this section, we document the urban-biased growth of the US economy between 1980 and 2015.

1.1 Data Sources and Measurement

The main data source for our paper is the Longitudinal Business Database (LBD), a restricted-use administrative data set providing payroll and employment counts for the universe of private employer establishments in the United States based on tax records.

While the LBD provides the most detailed and reliable information on employment and wages by industry and county, it has two shortcomings that we overcome by introducing additional data. First, its restricted-use policy makes it difficult to reproduce results from the LBD, so we supplement it with the the Quarterly Census of Employment and Wages (QCEW), an administrative, publicly available data set published by the Bureau of Labor Statistics. The QCEW

[2]The non-homothetic CES function is mainly used as a utility aggregator (e.g., Comin, Lashkari, and Mestieri (2020)) and was only recently studied as a production function when scale effects are important (see Trottner (2019) and Lashkari et al. (2022) for an analysis of the relationship between ICT usage and firm scale).

provides county-level tabulations of payroll and employment based on the universe of unemployment insurance records.[3] Second, the LBD lacks information on the educational attainment of workers. Therefore, we supplement it with a third data set, the US Decennial Census and American Community Survey (Census), which is survey-based microdata publicly available from Ruggles, Sobek, Alexander, Fitch, Goeken, Hall, King, and Ronnander (2015). Relative to QCEW or LBD, the Census contains information on individual workers' characteristics. The Census' drawback is its survey nature, which can make industry identifiers unreliable.[4]

In all data sets, we aggregate observations to 722 commuting zones (see Tolbert and Sizer (1996)) covering the entirety of the continental United States and focus on private-sector, non-agricultural employment. We aggregate all our data to 1-digit NAICS sectors, which are designed to capture the principal functional differences between groups of industries. Before 1997, we employ the establishment SIC-NAICS concordance from Fort and Klimek (2016) to map the native SIC into NAICS codes. In the Census, we split workers into those with at least a college degree ("college") and those without ("non-college") and those in cognitive non-routine occupations (CNR) and all others (non-CNR) following Rossi-Hansberg, Sarte, and Schwartzman (2019). We define the average wage within a location-sector pair as the ratio of its total payroll to its total employment. We deflate all nominal numbers by the CPI and provide more details on all data construction in the Online Appendix.

1.2 Documenting Urban-Biased Growth

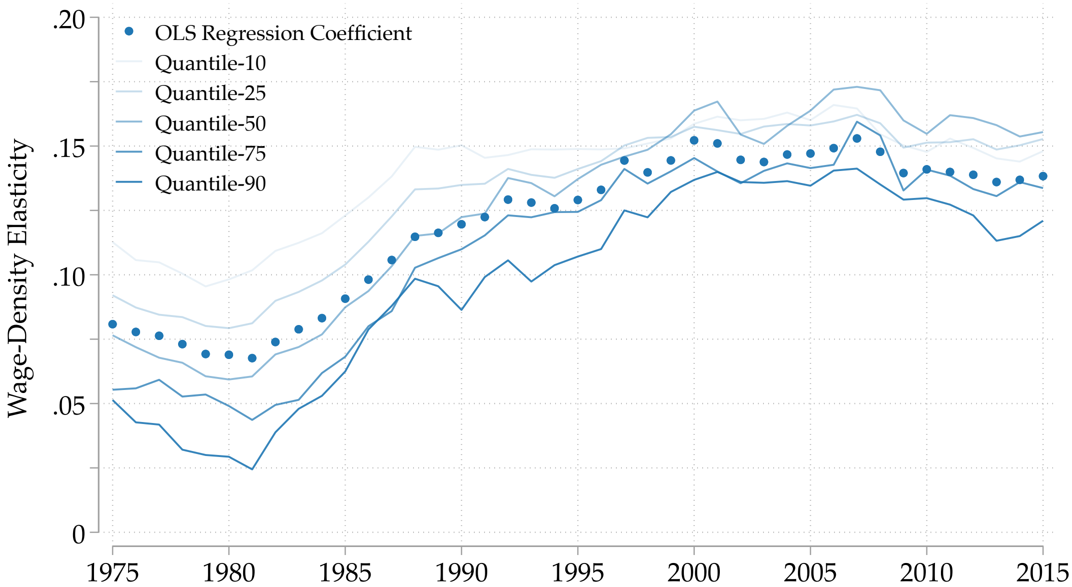
The observation that workers in locations with higher population density earn higher wages on average is one of the central facts of urban economics. Ahlfeldt and Pietrostefani (2019) show that a doubling of population density is associated with an 8% increase in average wages in a large cross-section of countries.

We show that the elasticity of wages to population density has *doubled* in the US economy since 1980. The blue dots in Figure 2 show yearly estimates of the elasticity of wages to population density in the cross-section of US commuting zones. The elasticity rose from 0.07 in 1980 to 0.14 in 2015, reflecting that large cities saw substantially faster wage growth compared to smaller cities between

[3] While the QCEW also has industry-level information, much of this data is suppressed due to data privacy restrictions.

[4] We discuss the shortcomings of the Census data in detail in the Supplemental Material on the authors' websites.

FIGURE 2: URBAN-BIASED GROWTH



Notes: This figure shows coefficients from a regression of log average wages on log population density run separately for each year between 1975 and 2015 across US commuting zones (blue dots), weighted by 1980 population. We use the US Bureau of Labor Statistics' Quarterly Census of Employment and Wages for wage data for private employers. We measure each commuting zone's population density in 1980 using US Census data. The lines show the coefficients from quantile regressions at the 10th, 25th, 50th, 75th, and 90th quantiles each year.

1980 and 2015.

Urban-biased growth was not uniform over time. The wage-density elasticity grew fastest between 1980 and 1990, then slightly slower between 1990 and 2000. The elasticity dropped and recovered following the dotcom bubble in 2000, but then dropped again somewhat during the 2007/08 financial crisis, and continued to stagnate at about twice its 1980 level.

The urban-biased growth patterns in Figure 2 are robust. The wage-density elasticity increased throughout the distribution of city wages, not just at the mean. The shaded blue lines in Figure 2 show the evolution of the coefficient at different conditional quantiles. We also replicate Figure 2 using the Census and LBD data, different measures of population density, and for counties instead of commuting zones in the Online Appendix. There, we also document that the wage-population-size elasticity follows a similar trend, reflecting that, empirically, population size and population density have a correlation coefficient of 0.9. As a result, we refer to size and density interchangeably throughout the paper.

Urban-biased growth is not only a feature of the US economy. We document comparable patterns of urban-biased growth for European cities over the same time period in the Online Appendix. The time series shows a similar pattern, with the steepest growth occurring before 2000, a decline during the dotcom bubble, followed by a brief increase and subsequent stagnation. The fact that urban-biased growth is not just a US phenomenon suggests that any explanation cannot be US-specific.

Urban-biased growth has been of an economically significant magnitude. Between 1980 and 2015, the wage gap between the average worker in the densest city (New York City, NY) relative to one in the median density city (Orlando, FL) grew by 22 percentage points. In the same time period, the gap in average wages between college and non-college workers increased by 54 percentage points. However, compared to the extensive macroeconomic literature studying skill-biased technological change, urban-biased growth has received less attention.

In the rest of the paper, we investigate the sources of urban-biased growth.

2. Accounting for Urban-Biased Growth

In this section, we present a detailed analysis of the role of the industrial, occupational, and demographic composition of different locations in explaining urban-biased wage growth. To keep the analysis succinct while adding these additional dimensions of heterogeneity, we aggregate commuting zones into a group of high- and low-density cities. High-density cities are the commuting zones with the highest population density jointly accounting for 50% of employment; low-density cities are all other locations.

2.1 The Role of Sectors

We index regions by r and sectors by s. We denote a location r's average wage by w_r, its average wage in sector s by $w_{r,s}$, and the share of its employment in sector s by $\mu_{r,s}$. We decompose changes in a location's average wage between two periods as follows:

$$\Delta w_r = w'_r - w_r = \sum_s \mu'_{r,s} w'_{r,s} - \sum_s \mu_{r,s} w_{r,s} = \sum_s \underbrace{\left(\mu'_{r,s} w'_{r,s} - \mu_{r,s} w_{r,s}\right)}_{\equiv \delta_{r,s}}, \tag{1}$$

where w_r precedes w'_r in time and $\delta_{r,s}$ denotes the contribution of sector s to wage growth in location r. The term $\delta_{r,s}$ reflects changes in employment shares *and* wages: average wage in location r can increase either because individual sectors experience wage growth or because employment shifts toward higher-wage sectors. Note that $\delta_{r,s}$ can be positive or negative.

We use equation (1) to study the industrial origins of the differential wage growth between high- and low-population-density cities indexed by $r = H$ and $r = L$, respectively:

$$(2) \qquad \Delta w_H - \Delta w_L = \sum_s (\delta_{H,s} - \delta_{L,s}) \, ; \qquad \phi_s \equiv \frac{\delta_{H,s} - \delta_{L,s}}{\Delta w_H - \Delta w_L},$$

where ϕ_s denotes the share of the differential wage growth between regions H and L accounted for by sector s. We refer to ϕ_s as the share of urban-biased growth accounted for by sector s.

Applying the decomposition to the LBD data reveals that the Business Services sector accounts for the vast majority of urban-biased wage growth. Figure 3 presents the share of urban-biased wage growth between 1980 and 2015 accounted for by each 1-digit NAICS sector in the economy. Changes in the wage and employment shares of the Business Services (or "NAICS-5") sector alone generate 98% of the observed urban-biased growth. The NAICS-5 sector accounted for 20% of national employment in 1980, 65% of which was concentrated in the group of high-population-density commuting zones. These numbers had changed to 26% and 61% by 2015.

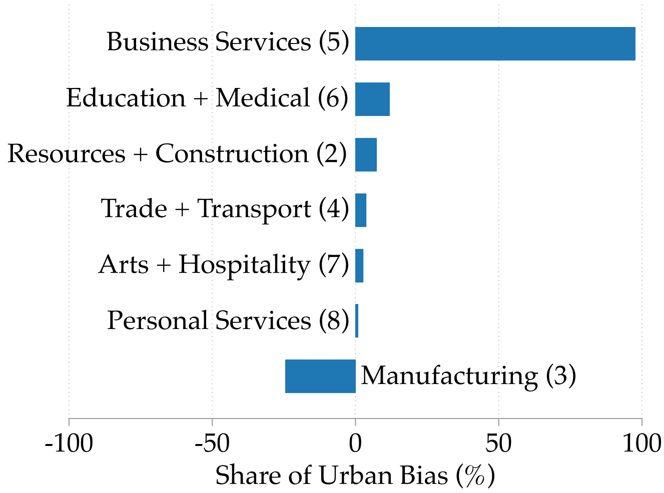
The Business Services sector comprises many industries common to high-population-density locations: professional services, finance and insurance, management of companies, information, administration and waste management services, and real estate services. In the Online Appendix, we apply the decomposition in equation (2) within the Business Services sector and find that all industries within the NAICS-5 sector experienced substantial urban-biased growth.[5] The industries contributing most to the sector's overall urban-biased growth are, in order of contribution, Professional Services, Finance, and Information.

Manufacturing is the sole NAICS-1 sector that exhibited negative employment growth since 1980, particularly in high-population-density cities. Since manufacturing jobs were high-paying on average, their disproportionate disappearance in these cities acted to depress the urban wage gradient.

Exposure versus Incidence. Business Services dominate big-cities' local econo-

[5]See Figure OA.2 in the Online Appendix, where we repeat the decomposition in (2) on the 2-digit level. Each Naics-1 sector accounts for a different fraction of national employment. In the Online Appendix, we present an alternative decomposition to control for differences in sector size; see Figure OA.3.

FIGURE 3: SECTORAL ORIGINS OF URBAN-BIASED WAGE GROWTH, 1980-2015



Notes: This figure shows the share of urban-biased wage growth between 1980 and 2015 accounted for by each NAICS-1 sector using the decomposition in equation (1). We compare wage growth between the commuting zones with the highest population density jointly accounting for 50% and all remaining commuting zones. The figure uses the Longitudinal Business Database.

mies. As a result, the Business Services sector's contribution to urban-biased growth may simply reflect Bartik-like "exposure differences" to the sector's fast aggregate wage or employment growth. We isolate an "exposure" component in each sector's contribution to local wage growth:

$$\delta_{r,s} = \underbrace{\mu_{r,s}\Delta\bar{w}_s + w_{r,s}\Delta\bar{\mu}_s}_{\text{Exposure}} + \xi_{r,s}, \tag{3}$$

where $\bar{x}_s$ denotes changes in variable x in sector s in the aggregate economy, and $\xi_{r,s}$ is a residual term.[6] Using equation 3, we can decompose each sector's contribution, ϕ_s, into an exposure term and a residual "incidence" term. The exposure term is the fraction of urban-biased growth due to aggregate shifts in wages and employment shares alone. The incidence term captures the fraction due to differences in the rate of change of sectoral wages and employment shares across locations. If all urban-biased growth reflected aggregate wage and employment share growth interacting with initial regional differences, the incidence term, $\xi_{r,i}$, would be zero.

Empirically, we find that exposure differences account for only a small share of urban-biased growth. The top panel of Figure 4 decomposes each sector's contribution into its exposure and incidence component. In the aggregate, expo-

[6]Formally, the residual term is: $\xi_{r,s} = \mu_{r,s}(\Delta w_{r,s} - \Delta\bar{w}_s) + w_{r,s}(\Delta\mu_{r,s} - \Delta\bar{\mu}_s) + \Delta w_{r,s}\Delta\mu_{rs}$.

sure differences alone explain only 13% of urban-biased growth. For Business Services, differences in incidence account for two-thirds of their contribution, suggesting that region-specific changes in wages and employment shares drive urban-biased growth in the sector.

Incidence is also the driver of urban-biased growth in most other sectors. The exception is the manufacturing sector. A full two-thirds of the sector's negative contribution to urban-biased growth reflects differences in exposure: the aggregate employment decline paired with the urban wage gradient within the manufacturing sector explains this result. The faster decline of manufacturing in big cities relative to the rest of the economy gives rise to the negative incidence component.

The key takeaway from the decomposition into exposure versus incidence is that exposure differences account for a small fraction of overall urban-biased growth in the aggregate economy and within the Business Services sector. As a result, explanations for the aggregate growth of the Business Services sector are insufficient to understand its spatial growth patterns. Instead, we must understand why wages and employment shares *within* the Business Services sector grew faster in high-density locations.

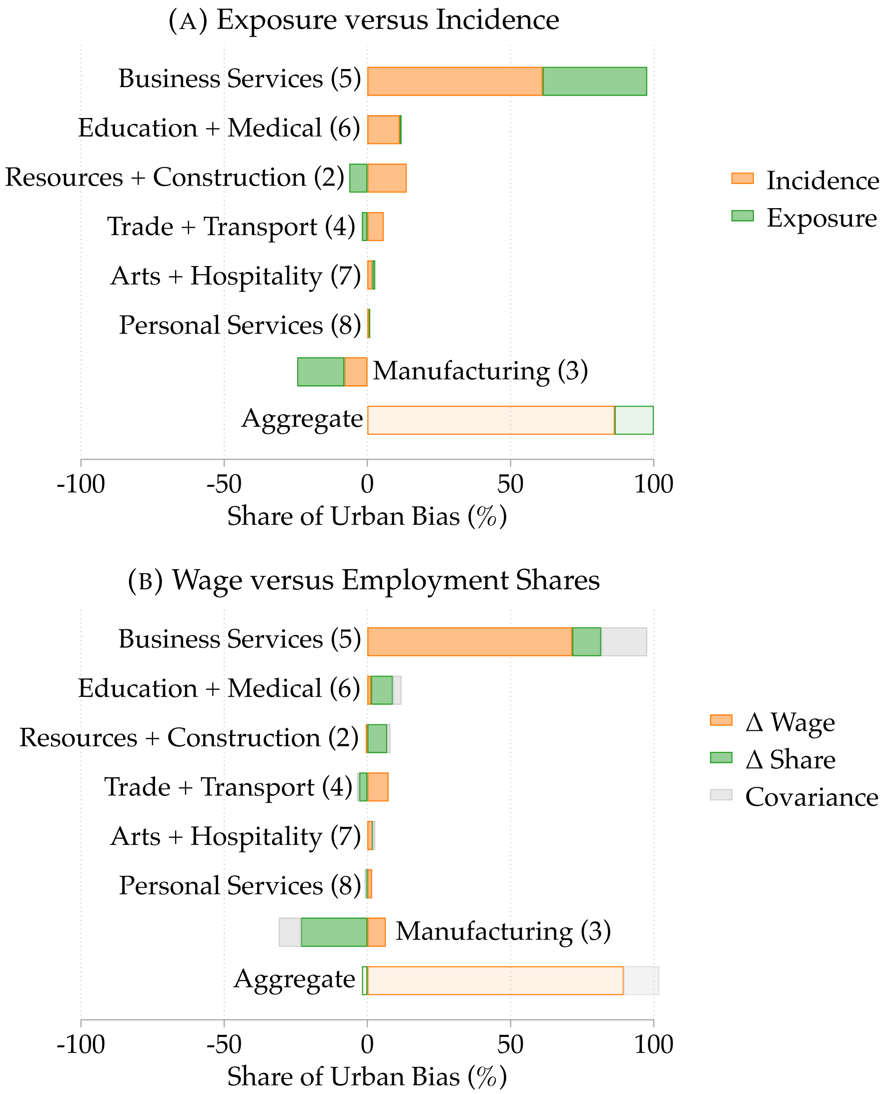
Wage Growth versus Employment Share Growth. We now explore whether urban-biased growth reflects wage growth or employment share growth differences across regions. Equation (1) showed that each sector's contribution to urban-biased growth can reflect differential changes in both wages and employment shares across regions. We decompose each sector's contribution, $\delta_{r,s}$, to differentiate these channels:

$$\delta_{r,s} = \underbrace{w_{r,s}\Delta\mu_{r,s}}_{\Delta\text{ Share}} + \underbrace{\mu_{r,s}\Delta w_{r,s}}_{\Delta\text{ Wage}} + \underbrace{\Delta\mu_{r,s}\Delta w_{r,s}}_{\text{Covariance}}. \tag{4}$$

The first term captures a location's wage growth due to changes in sectoral employment shares holding wages fixed; the second captures growth due to changes in sectoral wages holding employment shares fixed; and the third captures growth due to the co-movement of both sectoral employment shares and wages.

For Business Services, almost 75% is due to wage growth alone; another 15% due to the correlation between wage growth and employment share growth. The bottom panel of Figure 4 shows the share accounted for by each sector, ϕ_s, decomposed into the contributions of the margins in equation (4). Differen-

FIGURE 4: SOURCES OF URBAN-BIASED GROWTH, 1980-2015



Notes: The top panel decomposes the contribution of each sector to urban-biased growth into a component due to aggregate changes in sectoral wages and employment shares (see equation (4), "Exposure") and a component capturing urban-biased growth due to location-specific changes in sectoral wages and employment shares ("Incidence"). The bottom panel decomposes the contribution into a component holding sectoral employment shares fixed in 1980 while allowing wages to change as in the data ("Wage Growth"), a component that holds sectoral wages fixed and varies sectoral employment shares ("Sector Growth"), and their interaction ("Covariance"). The figure uses the Longitudinal Business Database. The parentheses behind sector description contain the sector's 1-digit NAICS code.

tial employment share growth alone explains only a small part of urban-biased growth.

There is heterogeneity in the relative importance of the three components across sectors. The "Education and Medical" and "Resources and Construction" sectors grew in an urban-biased way because they were more concentrated in big cities, near the centers of economic activity, while their wages were growing in a relatively balanced way across space. In the ”Trade and Transport” sector, wages grew faster in big cities, while employment grew faster elsewhere, making the employment share component negative. In manufacturing, collapsing employment, especially in big cities, led to a strongly negative sectoral employment share growth component. At the same time, manufacturing wages grew faster in big cities compared to small ones, leading to a positive wage and a negative covariance component.

Overall, 89% of urban-biased growth reflects wage growth differences across regions; it does not reflect big cities dramatically changing their sectoral composition.

2.2 The Role of Education and Occupation

The college share of employment in big cities increased sharply during the same period (see Diamond (2016)). Similarly, big cities are increasingly dominated by jobs in so-called cognitive non-routine occupations (see Rossi-Hansberg et al. (2019)). Such urban-biased compositional changes in the workforce may explain part of the observed urban-biased growth if it changes the composition of high-density cities toward higher-paying jobs. In particular, Business Services were already among the most skill-intensive industries in the US economy in 1980, and became even more skill-intensive by 2015.[7] In this section, we explore the role of such compositional changes.

We use the Census data, since the LBD lacks demographic information. Since the Census is a survey and sectors are self-reported, the fraction of urban-biased growth accounted for by each sector differs from the administrative data used in Figure 3.[8] The last column of Table 1 presents the share accounted for by each sector in the Census. The Business Services sector accounts for more than 100% of urban-biased growth, offsetting a more strongly negative contribution of the manufacturing sector.

We start with understanding the role of skill deepening. We index college workers and non-college workers by C and N, respectively, and decompose a sector's

[7] See Figure OA.1 in the Online Appendix.

[8] In particular, workers in high-skill service firms that also own manufacturing or retail establishments often misreport their sector as manufacturing or retail. As an example, in the Supplemental Material, we provide evidence that workers in Walmart's headquarters systematically report their sector as NAICS-44 (Retail) instead of the actual NAICS-55 (Management). As a result, the number of NAICS-55 workers in the Census microdata is substantially smaller than that reported in administrative data sources such as QCEW or LBD data.

contribution to local wage growth as follows:

$$\delta_{r,s} = \underbrace{\mu'_{r,s}(w^C_{r,s} - w^N_{r,s})\Delta\mu^C_{r,s}}_{\text{Deepening}} + \zeta_{r,s}, \tag{5}$$

where the term $\mu^C_{r,s}$ denotes the share of employment in sector s in region r accounted for by college-educated workers.[9]

When we apply the decomposition to college and non-college workers, the first term in equation (5) reflects the wage growth in location r and sector s that is the result of changes in the college share of employment between 1980 and 2015, holding the college wage premium fixed at its initial level. The residual term captures changes in college and non-college wages and changes in the sector's overall employment share. If all wage growth in sector s was due to changes in the composition of the workforce alone, the residual term would be zero. We again compute the difference of the term $\delta_{r,s}$ across our two groups of high- and low-density cities.

The left two columns of Table 1 present the results from the decomposition in equation (5) for college versus non-college workers. The changing composition of urban economies toward more educated workers explains about 28% of urban-biased growth. Across sectors, the importance of education-deepening varies. Education-deepening within Business Services explains only about 17% of all urban-biased growth. In other sectors, such as Manufacturing, Trade and Transport, and Education and Medical, education-deepening is responsible for a larger share of urban-biased growth within the respective sector, but only a very small fraction of overall urban-biased growth.

The left two columns of Table 1 reveals nothing about how wage growth is shared across education groups within each sector. To assess whether wage growth is broadly shared within each sector, we regress commuting zone level average wage growth on commuting zone population density, separately for college and non-college workers. Table 2 presents the results. Wage growth is faster in big cities for college and non-college workers conditional on working in Business Services, even if the urban bias in wage growth is stronger for more educated workers. College and non-college workers outside Business Services

[9] The residual term can be written in full as follows:

$$\zeta_{r,s} = \mu'_{r,s}\left((\mu^C_{r,s})'\Delta w^C_{r,s} + (\mu^N_{r,s})'\Delta w^N_{r,s}\right) + w_{r,s}\Delta\mu_{r,s}$$

TABLE 1: THE ROLE OF EDUCATION AND OCCUPATION

	Share of Urban-Biased Growth				
	Education		Occupation		
Sector	Deepening	Residual	Deepening	Residual	Total
Resources + Construction (2)	0.5	11.3	-0.2	12.0	11.8
Manufacturing (3)	4.2	-32.5	2.3	-30.5	-28.2
Trade + Transport (4)	3.4	-9.6	0.4	-6.5	-6.1
Business Services (5)	18.9	76.6	13.2	82.3	95.5
Education + Medical (6)	3.9	15.4	0.4	19.0	19.4
Arts + Hospitality (7)	1.4	3.5	-0.0	4.9	4.9
Personal Services (8)	0.5	2.2	0.2	2.5	2.7
Total	33.0	67.0	16.3	83.7	100.0

Notes: We use US Census data for 1980 and the American Community Survey data for 2015 from IPUMS and deflate all values by CPI-U. We compute average wages of full-time, prime-age workers within each commuting zone, sector, and either occupation or education group for both years. We only consider private non-agricultural employment. Not all observations have an occupational or industry code, and we omit those observations.

do not see substantially faster wage growth in big cities. College workers outside Business Services exhibit *less* urban-biased wage growth than non-college workers within Business Services. These results suggest that the urban-biased wage growth is a feature of the Business Services sector rather than of a particular education group.

Another recent line of work has studied the role of so-called cognitive non-routine (CNR) occupations in trends in aggregate and regional inequality (see, e.g., Rossi-Hansberg et al. (2019)). Columns 3 and 4 of Table 5 show that occupational shifts within Business Services explain about 9% of all urban-biased growth, whereas within-occupation wage growth explains the vast majority. Table 2 further shows that CNR and non-CNR workers within Business Services have seen strong urban-biased growth, whereas workers outside Business Services – regardless of occupation – have not.[10]

3. BUSINESS SERVICES, ICT, AND LARGE FIRMS

Fast wage growth for workers with college degrees in the Business Services sector is responsible for the vast majority of the urban-biased growth observed

[10]An example is doctors and nurses, which are CNR occupations outside of Business Services. These workers do earn slightly higher wages in cities, but have seen similar growth in these wages across space since 1980, leaving the medical urban wage gradient unchanged, and suggesting that this is not primarily a story of changes in the returns to cognitive work in cities.

TABLE 2: THE URBAN BIAS IN OCCUPATION AND EDUCATION WAGE GROWTH

	Growth in Average Commuting Zone Wage between 1980 and 2015			
	Business Services		All Other Sectors	
Education Group	College	Non-College	College	Non-College
Commuting Zone Population Density (1980, Logs)	0.0653*** (0.00992)	0.0195* (0.00769)	0.00321 (0.00424)	-0.0193* (0.00853)
adj. R^2	0.312	0.053	0.002	0.069
Occupation Group	CNR	Non-CNR	CNR	Non-CNR
Commuting Zone Population Density (1980, Logs)	0.0556*** (0.00843)	0.0903*** (0.00930)	0.00924* (0.00446)	-0.00705 (0.00934)
adj. R^2	0.283	0.460	0.023	0.007
N	722	722	722	722

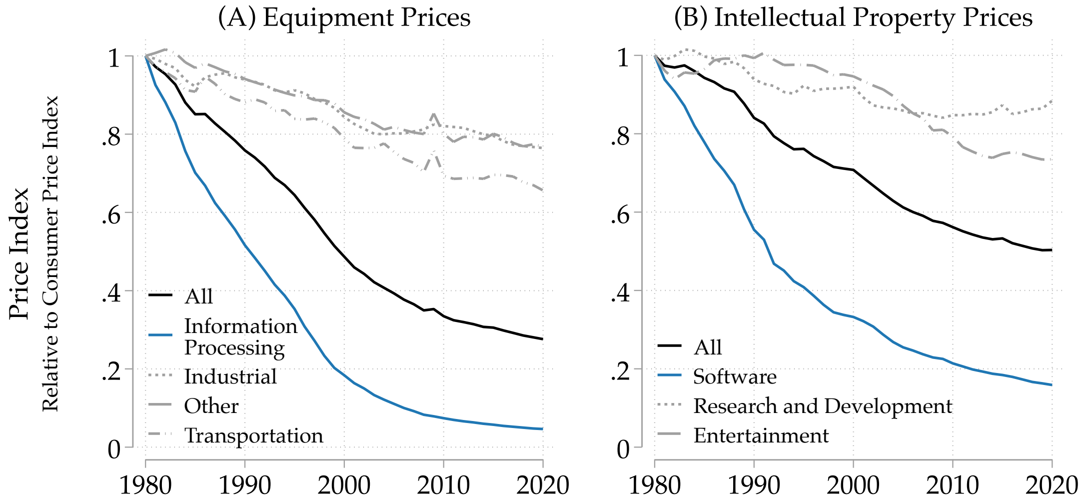
Notes: Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
We use US Census data for 1980 and the American Community Survey for 2015 from IPUMS and deflate all values by CPI-U. We compute average wages of full-time, prime-age workers within each commuting zone, sector, and either occupation or education group for both years.

in the US economy since 1980. Changes in the sectoral composition of cities, Bartik-type exposure differences across cities, and disproportionate skill deepening in big cities only explain a small fraction of urban-biased growth. We interpret our findings as evidence of a skill- and urban-biased labor demand shock originating in the Business Services sector.

A large literature links recent changes in the returns to skill in the US economy to advances in information and communication technologies (ICT) that particularly benefited skilled labor (see, e.g., (Krueger, 1993)). Krusell et al. (2000) modeled this as a "capital-skill complementarity" paired with falling capital prices. Their model can quantitatively account for the increase in the college wage premium since 1980 as a function of the observed decline in equipment (ICT) capital prices. We argue that, when allowed to vary across firms of different sizes, a capital-skill complementarity paired with the decline in ICT prices can also account for urban-biased wage growth.

Business Services and ICT. Figure 5 shows the overall decline in equipment capital and intellectual property prices using data from the BEA investment price indices. It also shows the price index for the narrower classes of "in-

FIGURE 5: TRENDS IN CAPITAL INVESTMENT PRICES



Notes: The left panel of this figure plots the price of equipment investment from 1980-2018 relative to the consumer price index. The right panel replicates that plot for intellectual property investment. The data used are the BEA asset price data and BLS CPI-U.

formation processing" and "software." Information processing equipment and software have played an outsized role in the decline of investment capital prices since 1980. For the rest of the paper, we define "ICT" as the sum of the "proprietary software," "pre-packaged software," and "computer hardware" capital categories in the BEA fixed asset tables and focus on the decline in their combined price index as an exogenous shock.[11]

The Business Services sector has been the most intensive user of ICT capital in the US economy since 1980. Figure 6 shows a proxy for the value added share of ICT capital by sector in 1980 and 2015.[12] In Business Services, ICT capital accounted for more than four times the share of value added compared to other sectors in both years.[13]

We run a set of firm-level regressions to show that Business Services firms in

[11]We take this decline to be exogenous to the spatial patterns we document, reflecting aggregate progress coming from Moore's law.

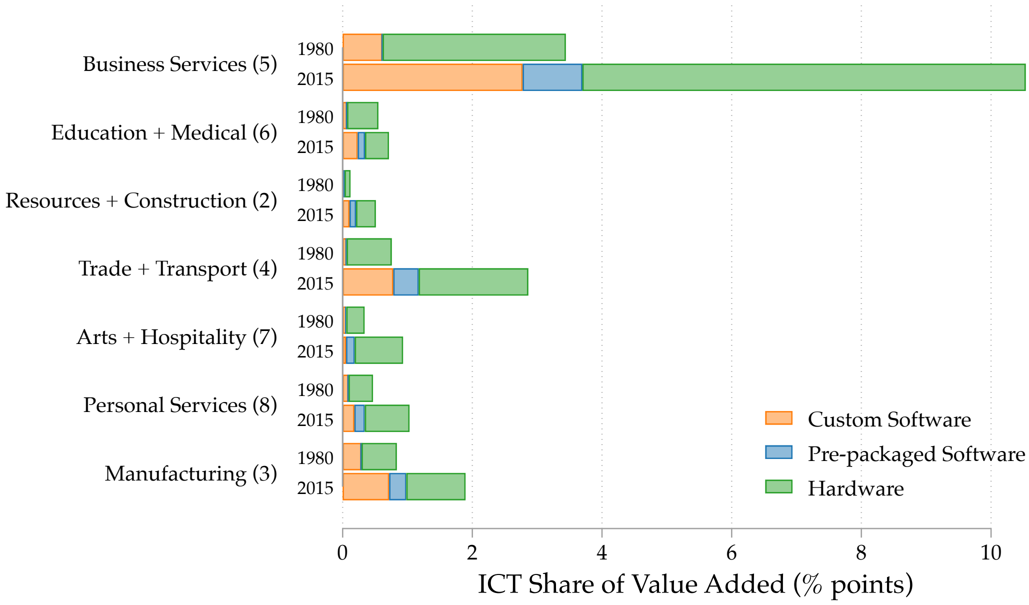
[12]We construct the value added share of capital type j in sector i at time t, $\Theta_{i,j,t}$, as follows:

$$\Theta_{i,j,t} = \underbrace{\frac{p_{i,j,t}K_{i,j,t}}{p_{i,t}K_{i,t}}}_{\text{Capital Type Share}} \times \underbrace{\left(\frac{VA_{i,t} - w_{i,t}L_{i,t}}{VA_{i,t}}\right)}_{\text{Non-Labor Share}},$$

where $K_{i,j,t}$ and $p_{i,j,t}$ are the stock and price of type j capital in industry i in year t. The terms $w_{i,t}$, $L_{i,t}$, and $VA_{i,t}$ denote average wages, total workforce, and total value added in sector i at time t.

[13]Figure OA.4 shows value added shares of ICT capital for more detailed, 2-digit NAICS industries within each sector. Almost all of the sub-industries of the Business Services sector are more intensive users of ICT capital than the average other industry in the US economy.

FIGURE 6: ICT VALUE ADDED SHARES ACROSS SECTORS



Notes: We show the share of real ICT of value added by industry in 2012 dollars from equation (12). Data is from the BEA. Prior to 1987, labor share uses data from the QCEW. Proprietary software refers to BEA codes ENS2 and ENS3, pre-packaged software refers to ENS1, and hardware to EP1A-EP31. Sectors are ordered by their contribution to urban-biased growth.

locations with high population density use ICT more intensively than their counterparts in lower density locations. We use data from the Annual Capital Expenditure Survey (ACES) of the US Census to measure ICT investments per employee at the firm level in 2013 and merge it onto the LBD.[14] For multi-establishment firms, we measure a firm's population density as the employment-weighted average density across establishments. We also define such firms' "Business Service employment share" as the fraction of their employment at establishments with a NAICS-5 code.

The first column of Table 3 shows that ICT expenditure per employee was substantially higher at firms in denser commuting zones. The second column shows that this was particularly true at firms with a large Business Service employment share. For a firm with only Business Service employment, doubling log population density raises ICT investments per capita by $550, compared to only $110 at a firm without any Business Service employment. To the extent that capital investments reflect capital stocks, Business Services firms in big cities are particularly intensive users of ICT capital.

[14]We do not have access to ACES data for earlier years. The BEA industry data used in Figure 6 are not available across firms or locations. In the ACES data, we cannot construct a proxy for value added shares, since we only observe ICT investments, total employment, and total payroll.

TABLE 3: ICT INVESTMENTS, SIZE, AND POPULATION DENSITY

	(1)	(2)	(3)	(4)	(5)	(6)
	ICT/Employee (x $1,000)					
Log(Density)	0.469***	0.155***			0.00140	0.101*
	(0.0299)	(0.0224)			(0.0520)	(0.0442)
Log(Employees)			0.352***	0.181***	-0.170**	0.167***
			(0.0158)	(0.0132)	(0.0607)	(0.0450)
Log(Employees) x Log(Density)					0.0889***	0.00201
					(0.0115)	(0.00848)
Business Services Emp. Share		-0.741		0.568**		1.696*
		(0.539)		(0.211)		(0.764)
x Log(Density)		0.651***				-0.182
		(0.0943)				(0.140)
x Log(Employees)				0.539***		-0.456*
				(0.0452)		(0.198)
x Log(Employees) x Log(Density)						0.163***
						(0.0346)

Notes: Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The independent variable "ICT investment per employee" is in thousands of 2013 dollars. Business Service Employment Shares represent the total share of employment at that firm in Business Service (NAICS-5) establishments. Density is the 1980 population density of the commuting zone for single establishment firms and the employment-weighted average across commuting zones for multi-establishment firms. The data come from the 2013 LBD and the ACES/ICTS survey. The regression reports robust standard errors. For conciseness, we do not report the regression constant.

Capital-Skill Complementarity in Space. As a point of departure, we study the implications of a general neoclassical capital-skill complementarity for wage growth across regions in a setting of declining capital prices.

Suppose that the economy consists of a set of locations r, each with a representative firm that uses a constant returns to scale technology to produce a homogeneous, freely traded good:

$$y = F_r(K, L) \quad \text{with} \quad \sigma_r \equiv \frac{d \log K/L}{d \log \frac{\partial F_r}{\partial L} / \frac{\partial F_r}{\partial K}} < 1 \tag{6}$$

where L and K denote skilled labor and capital. The production function is region specific; the productivity of factors can differ arbitrarily across regions. The assumption $\sigma_r(K, L) < 1$ corresponds to a global capital-skill complementarity whose strength can differ across locations arbitrarily. Capital is supplied elastically at a national price p, whereas labor markets clear locally with a wage w_r.[15] We abstract from low-skill labor for the moment to build intuition for the cross-sectional implications of the interaction of a capital-skill complementarity

[15]The exogenous capital price can be micro-founded by assuming that the final good can be converted into capital with some efficiency that varies with technical progress.

and declining capital prices.[16] We refer to this setup as the *neoclassical baseline* and provide a detailed discussion in the Online Appendix.

The constant returns to scale assumption combined with the capital-skill complementarity leads to sharp predictions for the cross-sectional response of wages to declines in the price of capital. In particular, for an exogenous decline in the price of capital (due to aggregate technical progress):

$$\frac{d\log w_r}{d\log p} = -\frac{\Theta_r}{1-\Theta_r} < 0 \tag{7}$$

where Θ_r is the cost share of capital.[17] Equation (7) shows that for a decline in the price of capital the general equilibrium response of wages is positive in all locations, but stronger in locations with a high capital cost share. The assumptions imply the following cross-sectional relationship between capital cost shares and wages:

$$\frac{d\log \frac{\Theta_r}{1-\Theta_r}}{d\log w_r} = \sigma_r - 1, \tag{8}$$

so that as long capital and labor are complements ($\sigma_r < 1$), locations with lower wages have higher capital cost shares.[18] Taken together equations 7 and 8 make a strong prediction: in the neoclassical baseline, wages *always* converge across regions as capital prices fall. Empirically, wages are higher in high-density locations even after controlling for observables (see, e.g., Ahlfeldt and Pietrostefani (2019)). In that context, the neoclassical baseline predicts rural- not urban-biased growth as capital becomes cheaper.[19]

The intuition for this *neoclassical channel* is simple. A decline in the price of capital reduces production costs and raises labor productivity more in locations where capital accounts for a larger share of costs. However, the capital cost share is lower in high-wage locations since capital and labor are complements. The neoclassical channel corresponds to the classic intuition that "capital flows to where labor is cheap."

[16] The classic CES production function with capital-labor complementarity is nested by our formulation: $y = Z_r(\lambda_r K^{\frac{\sigma-1}{\sigma}} + (1-\lambda_r)L^{\frac{\sigma-1}{\sigma}})^{\frac{\sigma}{\sigma-1}}$.

[17] Cost shares are defined as $\Theta_r \equiv pK_r/(w_r L_r + pK_r)$.

[18] For our result, it does not matter why wages are high in a location. Wages could be high in a location due to scarce labor supply, high labor productivity, or high general productivity. By implication, the result holds for any assumption on labor supply.

[19] The neoclassical channel also holds in a monopolistic competition setting, such as Krugman (1991).

Many papers have argued that local productivities are a function of the local population or population density. In the neoclassical baseline, as long as they are not internalized, such spillovers only strengthen the wage convergence implied by the neoclassical channel. As capital prices fall, workers move out of initially high-wage locations, endogenously raising productivity in initially low-wage locations.

Overall, wages converge across locations as capital prices fall in the neoclassical baseline. Such wage convergence occurred in the US economy for many decades before 1980, as documented by Barro and Sala-i Martin (1992). However, it is at odds with the patterns we document from 1980 onward, when wage growth started to be urban-biased, and ICT became an important factor of production.

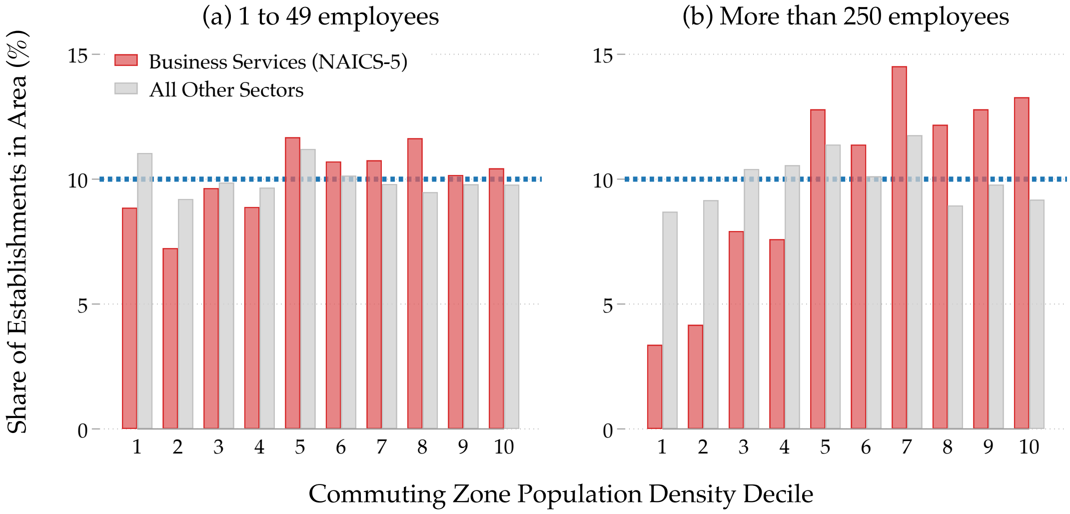
The Role of Large Firms. We argue that the neoclassical baseline misses a key point: large firms are more ICT-intensive than small firms, and large firms tend to locate in big cities, especially in the Business Services sector. Table 3 provides direct empirical evidence. Column 3 shows that ICT investments per employee are increasing in firm size as measured by the number of employees. Column 4 shows a strong positive interaction between firm size and the Business Services employment share. Column 6 shows the full interaction between population density, employment size, and the Business Service share of employment: large Business Services firms in big cities bought the most ICT per employee in 2013.

For example, Table 3 suggests that the average 10,000 person Business Services firm in New York bought 8 times more ICT capital per employee than the average 10 employee firm outside Business Services. The same 10,000 person New York firm also invested 68% more ICT per employee compared to a similar-sized firm in Orlando and 11 times more than a 10 person, non-Business Services firm in Orlando.

Not only is investment in ICT higher in large Business Services firms in big cities, but there is also a greater number of such firms. Figure 7 shows the share of small and large establishments located in one of ten commuting zone groups ordered by density, by sector. Large Business Services establishments are overrepresented in high-density commuting zones. Small Business Service establishments and establishments of any size in other sectors are evenly distributed across locations.

Figure 8 provides direct evidence that large firms in the Business Services sector have been crucial in driving the urban-biased growth documented in Section 1

FIGURE 7: ESTABLISHMENT SIZES ACROSS COMMUTING ZONE DENSITIES



Notes: We plot the share of all establishments in Business Services and all other sectors in each commuting zone decile, ordered by 1980 population density. This figure uses the 2015 US Census CBP.

above. It shows the total wage change between 1980 and 2015 in each of the ten groups of commuting zones ordered by increasing population density from Figure 1. The figure decomposes the wage growth in each group of commuting zones into the parts occurring at large and small firms in Business Services and the rest of the economy. Wage growth at large Business Services firms accounts for most of the difference in wage growth across locations.[20]

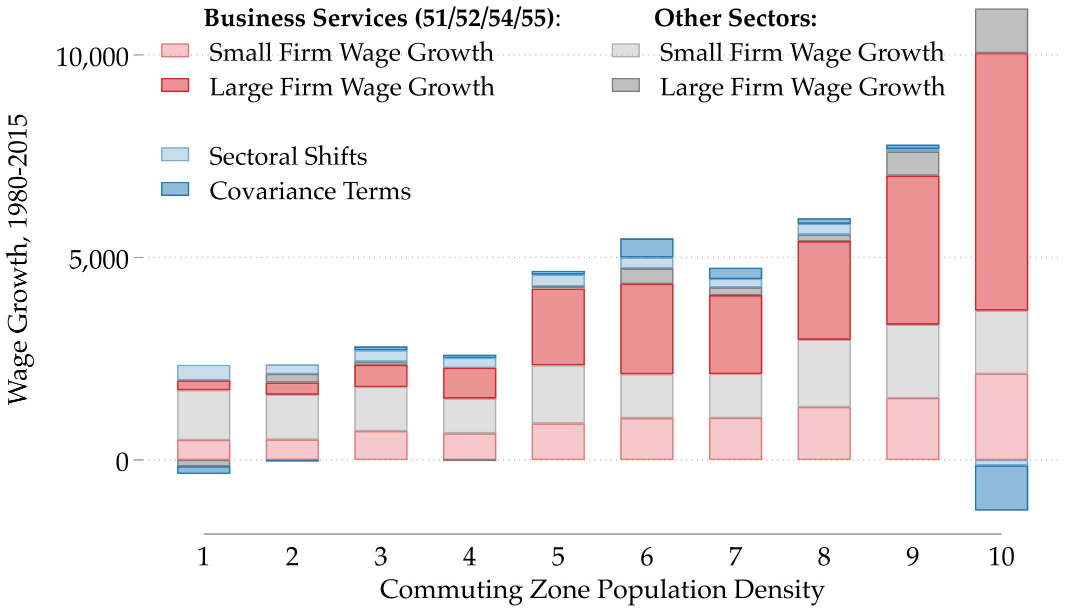
In the next section, we augment the neoclassical baseline with a capital-scale complementarity so that firms operating at a larger scale use capital more intensively. We show that if the capital-scale complementarity is strong enough, it can overcome the neoclassical channel discussed above and lead to larger capital cost shares in high-wage locations. In our augmented model, declines in ICT capital prices can generate faster wage growth in high-wage locations, which empirically are also high-density locations. In other words, declines in ICT prices can generate urban-biased growth.

[20]Formally, the wage change in region r can be decomposed as follows:

$$\Delta w_r = \underbrace{\mu^L_{r,O}\Delta w^L_{r,O}}_{\text{OL}} + \underbrace{\mu^S_{r,O}\Delta w^S_{r,O}}_{\text{OS}} + \underbrace{\mu^L_{r,N5}\Delta w^L_{r,N5}}_{\text{N5L}} + \underbrace{\mu^S_{r,N5}\Delta w^S_{r,N5}}_{\text{N5O}} + \underbrace{\sum_{f,s} w_{r,s,f}\Delta\mu_{r,s,f}}_{\text{S}} + \underbrace{\sum_{f,s}\Delta\mu_{r,s,f}\Delta w_{r,s,f}}_{\text{C}},$$

where $N5$ and O denote the Business Services sector and other sectors, and L and S index large and small firms. OL and OS refer to wage growth at large and small firms in the other sector, and similarly for N5L and N5S in Business Services. The term S is the sectoral shift component, and the term C is the covariance component.

FIGURE 8: URBAN-BIASED GROWTH AND LARGE FIRMS



Notes: We compute wage growth for each decile of commuting zones. We classify large firms as those with at least 1,000 employees. Such firms account for roughly 50% of US employment. Business Services firms are those with employment at establishments coded as NAICS 51, 52, 54, 55. Due to disclosure reasons we omit NAICS 53 and 56 here. We deflate all wages by the BLS CPI-U.

4. THEORY

In this section, we embed a one-parameter extension of a neoclassical production function with capital-labor complementarity into the workhorse spatial model. Section 5 presents a quantitative version of the model with richer detail, including low- and high-skill labor, firm heterogeneity, and multiple sectors.

4.1 The Model

Setup. The economy consists of a set of discrete regions indexed by r. Locations differ in an exogenous labor demand shifter, "productivity," and an exogenous labor supply shifter, "amenities." In each region, there is a unit supply of residential land. There is a final consumption good composed of varieties, which intermediate input firms produce using capital and labor. Input and final good markets are competitive, while intermediate input markets are monopolistically competitive. The final good, intermediate varieties, and capital are freely traded across locations. The model is static.

Technology and Market Structure. A representative firm produces the final consumption good by combining intermediate inputs with a constant elasticity of substitution ι. The price of the final good serves as the numeraire. As a result, an intermediate input firm's revenue as a function of output y is given by $\mathcal{D}y^{\zeta}$, where $\zeta = 1 - 1/\iota \in (0,1)$, and $\mathcal{D}$ is a measure of aggregate demand.

Intermediate input firms in location r produce their output, y, with a non-homothetic CES production technology ("NHCES"):

$$\left(\frac{l}{y}\right)^{\frac{\sigma-1}{\sigma}} + \left(\frac{k}{y^{1+\epsilon}}\right)^{\frac{\sigma-1}{\sigma}} = Z_r^{\frac{1-\sigma}{\sigma}},$$

where l and k denote the firm's choices for labor and capital, and the parameter σ indexes their substitutability. The term Z_r denotes the location-specific productivity.[21] The "non-homotheticity" parameter ϵ is a central parameter in our theory. As long as $\epsilon \neq 0$, the value of a firm's marginal product of capital depends on its level of output, y.[22] If $\epsilon = 0$ the production technology collapses to the standard CES production function in which each factor's marginal product is independent of the scale of production.

Importantly, the marginal rate of substitution between labor and capital is

$$\frac{\partial y/\partial l}{\partial y/\partial k} = (\frac{k}{l})^{\frac{1}{\sigma}} y^{-\frac{1-\sigma}{\sigma}\epsilon}. \tag{9}$$

As long as $\epsilon > 0$ and $\sigma < 1$, the marginal rate of substitution is decreasing in firm output. In other words, capital and labor are more complementary at firms operating at larger scale.[23] As a result, we refer to ϵ as the "scale elasticity." For the rest of the paper, we assume that capital and labor are complements, and that this complementarity is stronger at larger firms.

Assumption 1. *Capital and labor are complements and this complementarity is increasing in the level of firm output, i.e.,* $\sigma < 1$ *and* $\epsilon > 0$.

[21] Many papers offer micro-foundations for productivity differences across locations in terms of agglomeration spillovers; see Duranton and Puga (2004) for an extensive review. We do not explicitly micro-found these differences here, and take Z_r as fixed. However, in the quantitative extension we endogenize these differences as a function of local population.

[22] With the NHCES function, the elasticity of substitution continues to be constant at different ratios of input prices, but now varies across firms producing different levels of output at a given ratio of input prices (see Sato (1977)).

[23] This implies that a large firm that seeks to increase its labor force by 10% needs to increase its capital stock per worker by more than smaller firms to keep workers' marginal product constant.

Cost minimization of the firm gives rise to the following cost function:

$$c_r(y) \equiv c_r(y; Z_r, w_r, p) = Z_r^{-1} \left((w_r y)^{1-\sigma} + \left(p y^{1+\epsilon} \right)^{1-\sigma} \right)^{\frac{1}{1-\sigma}},$$

where, as before, w_r is the market price of a unit of labor in location r and p is the national price of capital.

Given the demand system for intermediate goods, firms solve the following profit maximization problem:

$$\pi^\star(Z_r, w_r, p, \mathcal{D}) = \max_y \left[\mathcal{D} y^\zeta - c(y; Z_r, w_r, p) \right].$$

To enter location r, firms must pay a fixed cost $\mathcal{E}$ denoted in units of local labor. Firms enter freely in each location until profits equal the fixed entry cost:

$$\mathcal{E} w_r = \pi^\star(Z_r, w_r, p, \mathcal{D}). \tag{10}$$

We denote the total number of firms that enter location r in equilibrium by N_r.

A representative capital-producing firm transforms the final good into capital at a constant rate $\mathcal{Z}$. Since the price of the final good serves as the numeraire, the price of a unit of capital is $p = 1/\mathcal{Z}$.

Preferences and Endowments. The economy is populated by a mass 1 of identical workers who inelastically supply one unit of labor. Workers spend a fraction α of their income on residential land and the remainder on the final good.[24] They choose their location to maximize utility, which is the product of consumption utility and a location-specific amenity term A_r. We assume that all regions have the same land supply and normalize it to 1.[25]

In equilibrium, utility is equalized across space, which yields an upward-sloping labor supply curve in each location:

$$L_r = A_r^{1/\alpha} w_r^{\frac{1-\alpha}{\alpha}} \mathcal{G} \quad \text{where} \quad \mathcal{G} \equiv (\sum_r A_r^{1/\alpha} w_r^{\frac{1-\alpha}{\alpha}})^{-1} \tag{11}$$

where the term $(1-\alpha)/\alpha$ is the local labor supply elasticity and $\mathcal{G}$ is a general

[24]This is isomorphic to a model in which consumers demand housing services, which are produced using local land and the final good. Both approaches yield the same upward-sloping labor supply curves in each location, with slightly different interpretations of the labor supply elasticity parameter.

[25]The labor supply elasticity in each location is independent of the local supply of land.

equilibrium shifter common to all locations. We refer to L_r interchangeably as the population size or population density of location r since all locations have the same land supply.

All rental proceeds from the land market accrue to a class of absentee capitalists who spend all their income on the final good.

4.2 General Equilibrium

An equilibrium in the economy is defined as follows:

Definition. *An equilibrium is a set of wages, rental rates, worker allocations, and number of firms, $\{w_r, r_r, L_r, N_r\}_r$, and a price of capital, p, such that (i) consumer location choices maximize utility, (ii) firm output choices maximize profit given prices, (iii) profits are equal to the entry cost in each location, (iv) and capital, labor, final good, and intermediate goods markets clear.*

A crucial object in our analysis is the capital cost share of firms in location r, which we denote as $\theta_r(y) \equiv pk(y)/c_r(y)$. The theory implies that capital cost shares satisfy:

$$\frac{\theta_r(y)}{1-\theta_r(y)} = w_r^{\sigma-1} p^{1-\sigma} y^{\epsilon(1-\sigma)} \tag{12}$$

For $\epsilon = 0$, we recover the standard neoclassical result that capital cost shares are independent of firm size, and declining in the local wage. However, with $\epsilon > 0$, capital cost shares are higher in locations in which firms produce at larger scale y. We denote the average capital cost share in location r by Θ_r.[26]

The following proposition characterizes equilibrium allocations across space.

Proposition 1. *In general equilibrium, in the cross-section of locations, (i) wages, w_r, and (ii) firm scale, y_r, are increasing in local productivity, Z_r. If $\epsilon > \zeta$, (iii) capital cost shares, Θ_r, are also increasing in Z_r.*

The first two results are intuitive. More productive locations pay higher wages; if they did not, firms in these locations would make higher profits than firms elsewhere, violating the free-entry condition. In particular, in the cross-section of locations, the following holds for wages:

$$\frac{d \log w_r}{d \log Z_r} = \frac{\zeta}{1+\Theta_r(\epsilon-\zeta)}, \tag{13}$$

[26]For now, firms in each location are identical so that $\Theta_r = \theta_r$; however, the distinction becomes relevant in the quantitative model below.

which is always positive since $\zeta < 1$ and $\epsilon > 0$.[27] Equation (13) highlights the role of ζ. As ζ approaches 1, firm varieties are more substitutable. As a result, a larger fraction of varieties is produced in the most productive locations, driving up their wage, and making the wage-productivity gradient steeper.

Firms have a lower marginal cost in more productive locations; if they did not, they would not be willing to pay higher wages to be there, given that trade is free and capital has a constant price across locations. As a result, demand for the varieties produced by firms in productive locations is higher, and they operate at a larger scale, y, in equilibrium.

The third part is more subtle. Equation (12) makes clear that two opposing forces shape the relationship of capital cost shares and productivity in the cross-section of locations.

First, since capital and labor are complements, a higher wage decreases the capital cost share. Since locations with a higher productivity have higher wages in equilibrium, this force pushes for a negative relationship between capital cost shares and local productivity. This is the "neoclassical" channel from Section 3; its strength increases in ζ, since the wage-productivity gradient is increasing in ζ (see equation (13)).[28]

Second, due to the positive scale elasticity ϵ, capital and labor are more complementary at firms operating at a larger scale (see equation (9)). Since firms in more productive locations are larger, this "scale" channel pushes for a higher capital cost share in more productive areas. If the scale channel dominates the neoclassical channel in equilibrium, capital cost shares increase in location productivity. The firm-level scale elasticity ϵ indexes the strength of the scale channel. The parameter restriction $\epsilon > \zeta$ is a sufficient condition for the scale channel to dominate the neoclassical channel.

In particular, the model permits the following expression for the relationship between capital cost shares and productivity in the cross-section of locations:

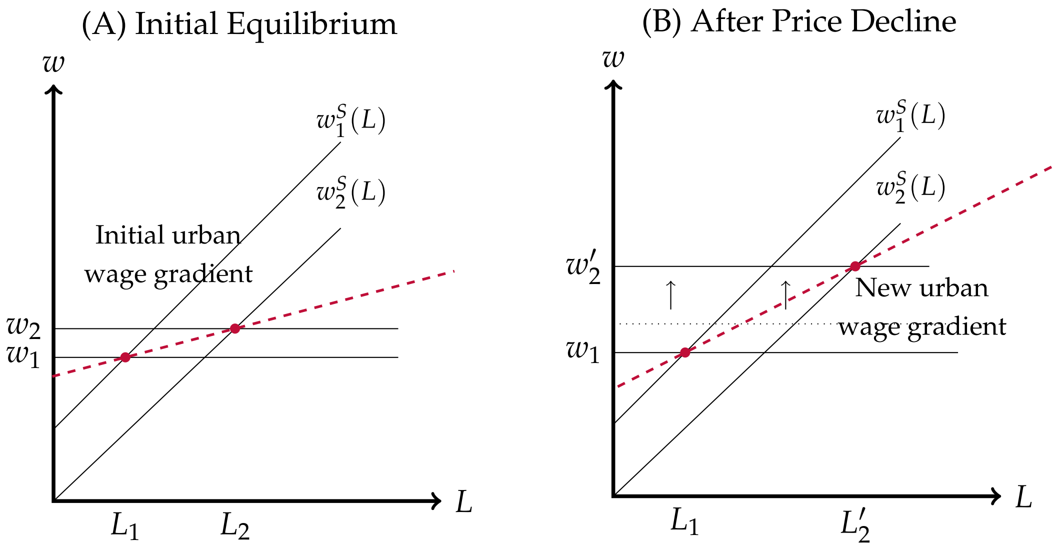
$$\frac{d \log \Theta_r}{d \log Z_r} = g(\Theta_r)\left(\underbrace{-(1+\epsilon-\zeta)\frac{d \log w_r}{d \log Z_r}}_{\text{Neoclassical}} + \underbrace{\epsilon}_{\text{Scale}} \right)$$

where $g(\cdot)$ is a positive, decreasing function. Note how for the case where $\epsilon =$

[27] Locations can be ranked in terms of their local productivity Z_r. This derivative and others should be understood as moving up the Z_r ranking.

[28] See Appendix B.2 for a detailed discussion.

FIGURE 9: THE WAGE-DENSITY GRADIENT IN EQUILIBRIUM



Notes: The axes are on a log-scale. L denotes employment and w denotes wages. The left panel of this figure shows the labor market equilibrium in the model with two locations. The right panel shows how the equilibrium changes after a capital price decrease. There is an adjustment to the general equilibrium constant $\mathcal{G}$ that we do not show, since it shifts both labor supply functions downward by the same amount.

0, the second term, the scale channel, disappears and the neoclassical channel implies that more productive locations have a *lower* capital cost share.

The cross-sectional relationship between local productivity and total employment depends on the correlation between productivity and amenities. The left panel of Figure 9 shows the equilibrium in a version of our model with just a high- and a low-productivity location. The figure shows the labor demand and supply curves on log scales (black lines) and the locus of the equilibrium points (dotted red line). The slope of the equilibrium locus is the wage-density gradient in the model. Labor demand curves are flat, since wages are independent of labor supply. The more productive location's labor demand curve is shifted up relative to the less productive location. Labor supply curves are log-linear and shifted by the amenity term A_r (see equation (11)).

For the equilibrium to feature a positive wage-density gradient – as in the data – the labor supply curve of the less productive location cannot be too far to the right relative to the labor supply curve of the more productive location. A simple sufficient condition for a positive wage-density gradient in equilibrium is for high-productivity locations to have higher amenities than low-productivity locations. This intuition extends to a setting with multiple locations.

4.3 The Effect of Changes in the Price of Capital

This section shows how a decline in the nationwide price of capital can cause urban-biased wage growth. The comparative static we consider is a uniform decline in the price of capital p caused by an increase in the productivity of capital production, $\mathcal{Z}$.

Proposition 2. *Under Assumption 1, and if $\epsilon > \zeta$, a decline in the price of capital, p, raises wages faster in higher wage locations.*

A decline in capital prices leads to larger cost savings at firms for which capital constitutes a larger part of their costs. Larger cost savings lead to larger increases in output and labor demand. If $\epsilon > \zeta$, then capital cost shares are higher in locations with higher wages (see Proposition 1). As a result, in contrast to the neoclassical model in Section 3, a decline in the price of capital raises wages most in locations with high initial wages.

In the Online Appendix, we derive the following expression for wage growth across regions in response to a decline in the price of capital:

$$\frac{d\log w_r}{d\log p} = -\frac{\zeta}{1+\Theta_r(\epsilon-\zeta)}\left(\Theta_r - (1+\Theta_r\epsilon)\frac{d\log\mathcal{D}}{d\log p}\right).$$

As long as $\epsilon > \zeta$, the expression is always negative, since increases in the price of capital lower aggregate demand. Importantly, the expression on the left is more negative the larger the average capital cost share of a location. As a result, as the price of capital declines, wages increase most in the location with high capital cost shares, high wages, and high underlying location productivity.

However, faster wage growth in high-wage locations only represents urban-biased growth if these locations also have larger populations in equilibrium. The right panel of Figure 9 shows the new labor market equilibrium in the economy after a capital price drop. Figure 9 clarifies that for a decline in the capital price to increase the wage-population gradient, the labor supply curve in the less productive location has to cross its labor demand curve to the left of the high-wage region's labor supply curve. In other words, productivity and amenities need to be positively correlated across locations for the decline in ICT prices to increase the wage-density gradient.

4.4 Discussion of the Mechanism

At the heart of our mechanism is that firms in more productive locations are larger and pay higher wages. This observation implies that if capital and labor are complements, capital cost shares are lower in high-wage locations, and declines in capital prices lead to wage convergence. However, if large firms produce with more capital per unit of labor than smaller firms, firms in more high-wage locations may have higher capital cost shares. We find this assumption highly plausible: doubling firm size at a 1-employee firm may not require extra investments in ICT infrastructure, whereas doubling firm size at a 10,000-employee firm likely requires a large expansion of ICT infrastructure.

We capture the increased need for ICT capital at larger firms in a reduced-form way through the non-homothetic CES production function. The advantage of the non-homothetic CES formulation is that a single parameter ϵ comes to summarize the "scale" channel parsimoniously. Trottner (2019) presents several micro-foundations for the non-homothetic CES production function.

Due to the non-homotheticity, our model is not scale-invariant. The size of aggregate demand, $\mathcal{D}$, matters for allocations. In the Online Appendix, we show that an expansion in aggregate demand, for example, through increased trade, leads to urban-biased wage growth under the same condition as a decline in capital prices. In the next section, we focus on quantifying the fraction of urban-biased growth explained by a decline in capital prices alone.

The most important determinant of the strength of the scale channel is the scale elasticity ϵ. The elasticity governs the relationship between firm size and ICT cost shares. Table 3 showed that ICT per employee is increasing in firm size, even when controlling for population density differences that capture factor prices across space. In a model without the non-homotheticity, capital cost shares do not vary across firms faced with the same input prices. Below, we exploit this fact to identify ϵ.

The strength of the scale channel also depends on the productivity differences across locations and the productivity of ICT capital in the firm's production function. Both of these factors are likely to differ across sectors. In particular, Business Services' high ICT value added shares suggest that ICT is much more productive in Business Services than in other sectors. Moreover, the large differences in Business Services employment shares across cities of different sizes suggest that productivity differences across space are particularly large in this sector. How much stronger the mechanism is for Business Services is ultimately an empirical question, which we now turn to answering.

5. Quantifying the Mechanism

We now embed the economic mechanism introduced in Section 4 into a richer quantitative model of the US economy. We use the calibrated model to quantify the importance of the mechanism in explaining the urban-biased growth observed in the US economy since 1980.

5.1 Extending the Model

We introduce three important additions to the model: multiple worker types, multiple sectors, and firm heterogeneity within location.

Technology, Commercial Land Markets, and Sectoral Good Aggregation. We follow Krusell et al. (2000) in introducing a second type of labor that is less complementary with ICT capital. We refer to labor highly complementary to capital as type-H labor and to labor less complementary as type-L labor. We index these types by e.

Intermediate input firms now produce using the following non-homothetic extension of the neoclassical production function in Krusell et al. (2000):

$$y = z\left(\left(\alpha^K_{r,s}(y)k^{\frac{\sigma-1}{\sigma}} + \alpha^H_{r,s}(y)h^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{1-\sigma}\frac{\kappa-1}{\kappa}} + \alpha^L_{r,s}(y)l^{\frac{\kappa-1}{\kappa}}\right)^{\frac{\kappa}{1-\kappa}} \tag{14}$$
$$\text{where} \quad \alpha^K_{r,s}(y) \equiv y^{\epsilon_K/\sigma}\phi^K_s Z^H_{r,s}, \quad \alpha^H_{r,s}(y) \equiv Z^H_{r,s}, \quad \alpha^L_{r,s}(y) \equiv y^{\epsilon_L/\sigma} Z^L_{r,s},$$

$Z^H_{r,s}$ and $Z^L_{r,s}$ are location- and sector-specific productivity terms for workers, and ϕ^K_s is a sector-specific productivity term for capital. The firm's demand for type-H labor, type-L labor, and capital is denoted by h, l, and k. Firms now also differ in a firm-specific efficiency term z and operate in one of two different sectors s, Business Services ($s = N5$) and Other ($s = O$).

The parameter σ parameterizes the elasticity of substitution of type-H labor and capital, and the parameter κ captures the elasticity of substitution between the bundle of type-H labor and capital, and type-L labor. The production function in equation (14) collapses to the production function in the simple model above for $Z^L_{r,s} = 0$, $\phi^K_s = 1$, and $z = 1$.

Intermediate input firms draw their efficiency z from a Pareto distribution with tail parameter ν after paying the fixed cost of entry. As a result, there is no selection on entry, and the shape of the distribution of fundamental firm efficiencies is the same across locations, in line with the evidence presented in Combes, Duranton, Gobillon, Puga, and Roux (2012). Nonetheless, average firm productivity differs across locations due to the location- and sector-specific productivity shifters, $Z^e_{r,s}$, operating on each type of labor.

To enter a location, intermediate input firms must purchase a building. A local construction sector combines commercially zoned land, $U_{r,s}$, and the final good, $X_{r,s}$, to create commercial buildings for sector s firms, $T_{r,s}$, using the following technology:

$$T_{r,s} = \tau_s X_{r,s}^{1-\eta_s} U_{r,s}^{\eta_s}$$

where τ_s is a sectoral cost shifter. We assume that commercially zoned land for each sector, $\bar{U}_{r,s}$, is in fixed supply.[29] As a result, the price of buildings within each sector rises with city size, and entry becomes more costly in more productive locations in equilibrium.

A class of atomistic capitalists own the claims to all intermediate input firms' profits and all returns to commercial land and spend their income on the final good.

In each sector, a final good firm aggregates sectoral intermediate input varieties with elasticity ζ_s, and these sectoral bundles into one final CES bundle with elasticity ζ_F. The homogeneous final good is traded freely across locations and serves as the numeraire.

Worker Heterogeneity, Location Decisions, and Sectoral Choice. In the simple model, workers derive utility from the consumption of the final good, the consumption of residential land, and local amenities. In the quantitative model, workers additionally receive idiosyncratic preference shocks for locations and sectors. We assume that workers draw their location- and sector-specific shocks from Fréchet distributions with inverse scale parameters A^e_r and $A^e_{r,s}$ and shape parameters ϱ^e_r and ϱ^e_s.[30] Workers first learn their location-specific shocks and only learn about their sectoral preferences upon arriving in a location.

These assumptions yield familiar expressions for the fraction of agents choosing to live in location r and for the fraction of workers choosing to work in sector s,

[29] Allowing a common land market between sectors introduces an element of competition for land between sectors from which we abstract for simplicity. For our purposes, the role of land markets is to generate an upward-sloping firm supply function in each sector and location.

[30] The amenity term in the simple model is now absorbed in A^e_r.

conditional on moving into a location r:

$$\lambda_r^e = \frac{A_r^e(\bar{v}_r^e)^{\varrho_r^e}}{\sum_r A_r^e(\bar{v}_r^e)^{\varrho_r^e}} \text{ where } \bar{v}_r^e \equiv R_r^{-\alpha}(\sum_s A_{r,s}^e(w_{r,s}^e)^{\varrho_s^e})^{\frac{1}{\varrho_s^e}};\ \mu_{r,s}^e = \frac{A_{r,s}^e(w_{r,s}^e)^{\varrho_s^e}}{\sum_s A_{r,s}^e(w_{r,s}^e)^{\varrho_s^e}}$$

where the terms A_r^e and $A_{r,s}^e$ play the role of type-specific amenity terms for regions and sectors, respectively, and R_r is the rental rate of local housing. We denote the aggregate supply of type e workers in the economy by $\bar{L}^e$ and the equilibrium quantities of type e workers in region r and sector s by $L_{r,s}^e$, so that $L_{r,s}^e = \lambda_r^e \mu_{r,s}^e \bar{L}^e$.

5.2 Calibrating the Model

We calibrate our model to data from the US economy in 1980. Since we require information on the education status of workers, we use the US Decennial Census data (see Ruggles et al. (2015)). We map locations in the model to data on 722 commuting zones covering the entire continental United States. Following Krusell et al. (2000), we define type-H workers as those with at least a four-year college degree and type-L workers as all others; we refer to type-H and type-L workers as college and non-college workers from now on. We estimate most of the parameters using a method of moments estimator and calibrate others from external sources. While most parameters are estimated jointly, we discuss the calibration strategy of each parameter in terms of its most informative empirical moment. Table 4 provides an overview of the calibrated parameters.

Productivities, Amenities, and Housing Supply. We follow the recent quantitative spatial literature in choosing all regional productivity and amenity terms to match regional data in 1980 exactly, given all other parameters (see Redding and Rossi-Hansberg (2017)). In particular, we infer productivities ($Z_{r,s}^e$) and amenities ($A_r^e, A_{r,s}^e$) as structural residuals to ensure the model matches average annual wages and employment counts for all worker types, sectors, and locations exactly.

The residential housing supply in each location, $\bar{H}_r$, is fixed. To calibrate $\bar{H}_r$, we construct a hedonic rent price index for US commuting zones for 1980.[31] We choose housing supply $\bar{H}_r$ in each location to match the rent price index exactly. We assume that commercial land supply in each location, $\bar{U}_{r,s}$, is proportional to residential housing supply, $\bar{H}_r$.[32]

[31] We use gross rents from the Decennial Census to construct the index; see the Online Appendix for more details.

[32] We are interested in the change in the distribution in economic activity in response to a

We set the productivity of capital production, $\mathcal{Z}$, to 1 in 1980. Given this normalization, we choose the productivity of ICT capital in each sector, ϕ_s^K, to match the aggregate share of ICT value added in the BEA asset tables in 1980 as displayed in Figure 6 above.

Scale Elasticities ϵ^K and ϵ^L. An important implication of our theory is that factor ratios vary with firm output. As a result, we use moments relating to the variation of factor input ratios across firms of different sizes to inform the non-homothetic scale elasticities.

In the quantitative model, factor input ratios satisfy the following two equations:

$$\frac{k}{h} = (\frac{p}{w_{r,s}^H})^{-\sigma} y^{\epsilon_K} \quad \text{and} \quad \frac{h}{l} = (\tilde{w}_{r,s}^H)^{-\sigma(1-\sigma)} (\frac{\tilde{w}_{r,s}^H}{w_{r,s}^L})^{-\kappa} y^{-\epsilon_L} (Z_{r,s}^L)^{-1}$$

where $(\tilde{w}_{r,s}^H)^{1-\sigma} \equiv (w_{r,s}^H)^{1-\sigma} (Z_{r,s}^H)^{\sigma} + p^{1-\sigma} (Z_{r,s}^H)^{\sigma} y^{\epsilon_K}$. The ratio of capital to college labor within a firm varies with firm output with an elasticity ϵ_K. Conditional on ϵ_K, the parameter ϵ_L governs the ratio of college to non-college workers within each firm.

We choose ϵ_K to match the difference in average ICT investment per worker between the smallest firms (fewer than 10 employees) and the largest firms with 1000 employees or more, in data from the 2013 Annual Capital Expenditure Survey available within the US Census. We calibrate ϵ_L to match the difference in the college to non-college worker ratio between the smallest firms (fewer than 10 employees) and the largest firms (with 1000 employees or more) in data from the 1990 Current Population Survey (CPS).[33]

We use data on factor shares from more recent years since data for earlier years are unavailable. Note, however, that since we target *relative* investments and skill ratios between small and large firms, their level (which varies over time as the economy becomes more skill- and capital-intensive) is less relevant.

Substitution Elasticities σ and κ. Given the scale elasticities, the substitution elasticities σ and κ determine the ease of substituting between capital, college, and non-college labor at the firm level. We choose these elasticities to ensure that our calibrated model matches canonical estimates of the macro substitution elasticities between these factors from Krusell et al. (2000).

decline in the price of ICT capital. The stock of commercial land supply does not influence this response since it does not influence the land price elasticity. The constant of proportionality is absorbed into the sectoral shifter τ_s.

[33] The resulting fit for all firm-size bins is shown in Figure OA.10.

In particular, we follow Burstein and Vogel (2017) and define the following macro elasticities:

$$\tilde{\kappa} \equiv \frac{d\log\left(\bar{w}^H/\bar{w}^L\right)}{d\log(\bar{L}^H/\bar{L}^L)} \quad \text{and} \quad \tilde{\sigma} \equiv \frac{d\log\left(\bar{w}^H/p\right)}{d\log(\bar{L}^H/\bar{L}^L)}$$

where $\bar{w}^e \equiv \sum_{r,s} w^e_{r,s}(L^e_{r,s}/\bar{L}^e)$ is the economy-wide average wage for type e workers. Krusell et al. (2000) estimate $\tilde{\sigma} = 0.66$ and $\tilde{\kappa} = 1.67$, which we use as our calibration targets.

Recall that our calibration matches worker allocation and wages across locations in 1980 exactly. To calculate the $\tilde{\kappa}$, we compare the college wage premium in the baseline economy to that in an alternative economy with the same parameters but an aggregate ratio of college to non-college workers, which is 1% higher. To compute the model-implied $\tilde{\sigma}$, we compare the baseline economy to an alternative one in which the ratio of college to non-college workers changes enough to cause an endogenous adjustment in K such that the ratio of capital to college workers increases by 1%.[34]

Other Parameters. The Fréchet dispersion parameters of the location preference shocks, ϱ^e_r, function as spatial labor supply elasticities. We use values for these elasticities from Diamond (2016), whose estimates are based on Census data for US metropolitan areas in the same period as our model. We obtain the sectoral labor supply elasticities, $\varrho^e_{r,s}$, from Artuç, Chaudhuri, and McLaren (2010). We follow Eckert and Kleineberg (2019) in setting the Cobb-Douglas share associated with housing, α, to 0.3.

We set the firm-level demand elasticity, ζ_s to 4 in both sectors, following Garcia-Macia, Hsieh, and Klenow (2019), and the sectoral demand elasticity, ζ, to 1.2. Doing so allows us to match the aggregate payroll share going to the Business Services sector in our counterfactual exercise.[35] We choose the tail coefficient of the firm efficiency distribution, ν, to match the tail index of the US firm size distribution reported in Axtell (2001).[36]

We choose the cost shifter of commercial land services, τ_s, to match the av-

[34] The resulting estimate of the micro elasticity κ is significantly higher than the corresponding macro elasticity. The difference reflects the fact that increases in the college ratio induce ICT capital investment due to the complementarity with college labor, which partially offsets the decrease in the college wage premium.

[35] We explore the sensitivity of our results to different values of this sectoral elasticity in the Supplementary Material on the authors' websites.

[36] The quantitative results are largely insensitive to this choice, and matching the higher tail index of the establishment size distribution instead leaves our results unchanged.

erage establishment size by sector in the County Business Patterns data from 1980. The cost share of commercial land in commercial building construction, η_s, acts as an entry elasticity, governing the responsiveness of the number of firms to expected local profits. We estimate η_s to match the gradient of average establishment size to local population density in each sector in data from the QCEW.[37]

5.3 Discussion of Calibrated Parameters

Table 4 shows the calibrated parameters with their associated moments. As discussed in Section 4, our mechanism relies on larger ICT cost shares in higher-density locations. The left panel of Figure 10 shows capital cost shares across commuting zones in 1980 in our baseline calibration. Two observations are crucial for the strength of our mechanism. First, capital cost shares in Business Services are higher than those in the rest of the economy. Second, the cost shares in Business Services firms in high-density locations are much higher than in smaller locations. The positive cost share density gradient in the rest of the economy is much less pronounced. Together, these two facts imply that declines in the price of ICT raise labor productivity most in Business Services in high-density locations.

The key parameter driving the difference in ICT cost shares across sectors is the sector-specific productivity of ICT capital, ϕ_s^K, calibrated to match differences in sector cost shares in 1980. Since productivities of college and non-college labor differ across locations and sectors, the parameter ϕ_s^K cannot directly be compared across sectors. However, all else equal, our estimates imply that ICT capital is more productive in Business Services than in the rest of the economy.

The key parameters driving the difference in cost shares across locations are the scale elasticities, the substitution elasticities, and differences in location productivities. The calibrated regional productivity terms are strongly increasing in population density. The productivity-density gradient is steepest for college workers in Business Services.[38] As a result, in the calibrated model, firms in high-density location are larger, since these locations are more productive.[39]

[37]See Figure OA.6 and Figure OA.8 in the Online Appendix for average establishment sizes across commuting zones in the data and the model.

[38]The empirical moment identifying these productivity differences is the strong positive wage-density gradient for college and non-college workers in the data. See Table OA.1 in the Online Appendix for the correlation of population density with the different location- and type-specific productivity terms.

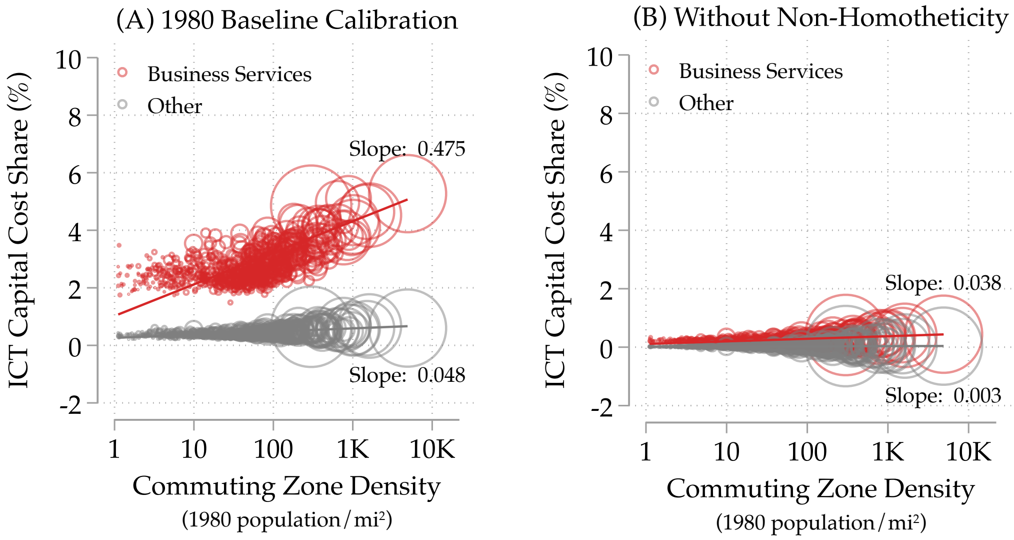
[39]See Figure OA.8 in the Online Appendix.

TABLE 4: OVERVIEW OF MODEL PARAMETERIZATION

	Estimated Structural Parameters	Value	Description of Moment	Moment: Model/Data
ϵ^K	Capital Scale Elasticity	0.31	Ratio of capital p.w. between 10 and 1000 emp firms	3.8/3.8
ϵ^L	Unskilled Scale Elasticity	-0.74	Skill ratio gap between 10 and 1000 emp firms, Figure OA.10	10% /10%
σ	EoS: Skilled Labor and Capital	0.65	Capital-skill macro elasticity of Krusell et al. (2000)	0.61/0.67
κ	EoS: Skilled and Unskilled Labor	2.12	Skilled-unskilled macro elasticity Krusell et al. (2000)	1.27/1.67
η_s	Firm Supply Elasticity	(0.25,0.13)	Elasticity of size to population density, Figure OA.6	(0.19,0.15)/(0.19,0.15)
τ_s	Entry Cost Level	(1.07,1.94)	Average establishment size by sector (QCEW)	(11.7,14.4)/(11.7,14.4)
ϕ_s^K	Sectoral ICT Capital Productivity	(67.2,1.8)	ICT share of value added by sector, Figure 6	(3.5%,0.5%)/(3.5%,0.5%)
	External Structural Parameters	Value	Source	
α	Housing share in final consumption	0.3	Eckert and Kleineberg (2019)	
$\varrho^{r,e}$	Spatial Labor Supply Elasticity (type-L, type-H)	(4.2,5.5)	Diamond (2016)	
$\varrho^{s,e}$	Sectoral Labor Supply Elasticity (type-L, type-H)	(0.21,0.21)	Artuç et al. (2010)	
ζ_s	Intermediates CES Aggregator	4	Garcia-Macia et al. (2019)	
ζ_F	Final Good CES Aggregator	1.2	N/A	
ν	Pareto Shape Parameter	1.1	Axtell (2001)	
	Productivities, Amenities, and Land Supply	Value	Data Matched	
$Z^e_{r,s}$	Location Productivity Shifter	Various	1980 employment and wages	
A^e_r	Location Amenities	Various	1980 CZ employment and wages	
$A^e_{r,s}$	Sectoral Amenities	Various	1980 CZ employment and wages	
$\bar{H}_r$	Residential Land Supply	Various	1980 local rents	
$\bar{U}_{r,s}$	Industrial Land Supply	Various	Proportional to $\bar{H}_r$	
$\mathcal{Z}$	Productivity of ICT Production in 1980	1	Normalized	

Notes: This table shows the baseline parameterization of the model. The location fundamentals vary across regions and by education group and sector, so their values are not listed. Where two values appear for a parameter (representing the value for the two sectors), the value for Business Services is first.

FIGURE 10: CAPITAL SHARES IN THE MODEL



Notes: This figure plots model-implied ICT cost shares, Θ. Panel (a) is the calibrated model, Panel (b) shows capital cost shares in a homothetic version of the model where $\epsilon = 0$.

To fit the variation in ICT investment per capita and the college to non-college ratio across firms of different sizes, we estimate scale elasticities of $\epsilon^K = 0.31$ and $\epsilon^L = -0.74$. For the substitution elasticities, we find $\sigma = .65$ and $\kappa = 2.12$. These findings imply that college labor and ICT capital are strong complements and that complementarity is increasing in firm scale. As a result, ICT cost shares are higher in high-density locations, particularly for Business Services firms.

The right panel of Figure 10 shows ICT cost shares across regions when we set the scale elasticities to zero, i.e., in a homothetic version of our model. We re-calibrated regional and sectoral fundamentals to ensure the model still matches the same cross-section of average wages and employment. All other parameters remain fixed at their calibrated levels. It shows that the positive gradient across regions disappears almost entirely, since the firm size differences across locations no longer affect the marginal products of capital and the different types of labor.[40]

[40]In Section 3, we showed that with just one worker type complementary to capital, capital cost shares are lower in high-wage (i.e., high-density) locations. In the quantitative model, the presence of non-college workers, which are substitutes for capital, implies that the high wages of non-college workers in cities push for higher capital cost shares in high-wage cities. These forces partially offset one another, leading to a slightly upward-sloping gradient in the right panel of Figure 10.

Table 5: Wage-Density Coefficient in Data and Model

	1980	2015 Data	2015 Baseline	2015 Homothetic	2015 A Spillover	2015 Z Spillover
Business Services	0.067	0.151	0.133	0.072	0.122	0.130
Other Sectors	0.056	0.068	0.049	0.049	0.031	0.046
Aggregate	0.060	0.099	0.105	0.059	0.109	0.120
Δ Aggregate		0.039	0.045	-0.001	0.049	0.059

Notes: Gradients computed use the ACS/Census for 1980 and 2015, weighting by 1980 population shares. Homothetic sets $\epsilon = 0$, but keeps all other parameters the same. A-Spillovers allow for amenity spillovers from college-educated workers, as in Diamond (2016). Z-spillovers allow for agglomeration economies from college-educated individuals as in Rossi-Hansberg et al. (2019).

5.4 Counterfactual Exercise

We now quantify the ability of our mechanism to generate the urban-biased growth observed in the US economy since 1980.

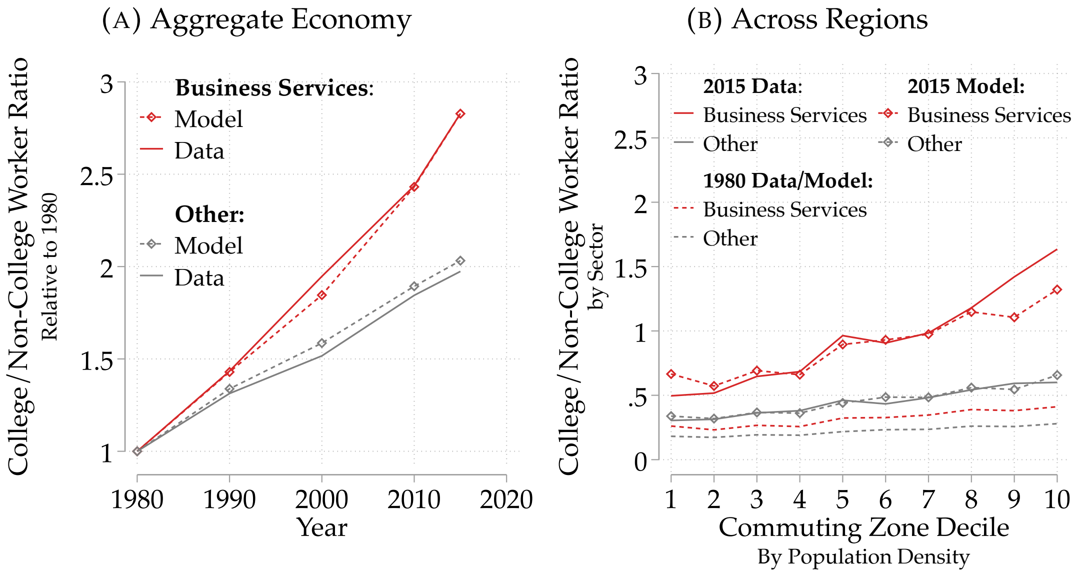
The first two columns of Table 5 present the wage-density gradient in the data in 1980 and 2015, for each sector and in the aggregate. Since we use Decennial Census data for the calibration, the wage-density gradients in the data in both years differ somewhat from that shown in Figure 2. However, as in the LBD, the gradient roughly doubled between 1980 and 2015. As in the LBD, most of the increase in the gradient occurs in the Business Services sector, as rows 1 and 2 suggest. Since we calibrated our model to match the 1980 cross-section of data, the first column of the table is also the 1980 wage-density gradient in the model.

To understand the role of our mechanism in explaining the change in the wage-density gradient, we take the calibrated model and increase the productivity of capital production, $\mathcal{Z}$, to match the decrease in the price of ICT capital in the BLS data between 1980 and 2015.[41] We hold all local amenities, land supplies, and productivity parameters fixed at their calibrated values but adjust the relative aggregate supply of college and non-college workers to evolve as in the data.

Column 3 shows the resulting wage-density coefficient in our baseline model, by sector, and in the aggregate. Declining ICT prices alone can explain almost all of the increase in the wage-density gradient in the data. The model also

[41] ICT capital combines information processing equipment and software-related capital; we match their joint price index over time. Figure 5 shows the price index for each category separately.

FIGURE 11: SKILL DEEPENING IN MODEL AND DATA



Notes: The left panel of this figure shows the growth in the ratio of college-educated to non-college workers in both the model and the Decennial Census data by year and sector. The right panel of this figure shows this ratio in 2015 in both the model and the data by sector across the commuting zone groups of increasing density used throughout the paper.

captures that virtually all of the increase in the gradient occurs in the Business Services sector. The model generates these large differences across sectors because of the differences in the level and gradient in cost-shares across sectors shown in Figure 10.[42]

The increase in the aggregate gradient reflects both the changes in the gradient in each sector and the reallocation of workers across sectors. The left panel of Figure 11 shows the ratio of college to non-college workers by sector in the model and the data. In the right panel of Figure 11, we show the ratio of college to non-college workers across commuting zones. The model generates a slightly flatter gradient in the ratio across locations, suggesting an additional role for adjustments in amenities or productivities from their 1980 levels (see, e.g., Diamond (2016), Almagro and Domınguez-Iino (2019) and Rossi-Hansberg et al. (2019)). Note that while we adjust the *aggregate* share of college workers as in the data, all sorting of workers across sectors and locations is endogenous. [43]

The non-homotheticity of the production function is essential in generating

[42]In Figure OA.9 in the Online Appendix, we also show the non-parametric relationship between wages and population density in the model in 2015.

[43]Because the model replicates the changes in worker stocks across regions, it also generates much of the changes in the rent price index across space observed in the data. See Figure OA.7 in the Online Appendix.

urban-biased growth in response to the decline in the price of ICT. In Column 4 of Table 5, we present the wage-density gradient in 2015 in a version of the model with scale elasticities set to 0 and fundamentals re-calibrated to continue to exactly match the 1980 data on wages and employment. Without the scale elasticities, the wage-density gradient does not meaningfully increase as ICT prices fall, reflecting the flat cost shares across regions in the homothetic case, shown in Figure 10.

Finally, the decline in the ICT price also generates a significant increase in the aggregate college wage premium. The aggregate college premium rose from 1.45 in 1980 to 1.79 in 2015, compared to 1.89 in the Decennial Census data. Krusell et al. (2000) showed that the decline in equipment capital prices could explain the increase in the college premium between 1980 and 2000. We match the increase in the college premium so closely because our production function is a direct non-homothetic extension of that in Krusell et al. (2000), our price series is nested in theirs, and we target the same macro elasticities of substitution between capital, college and non-college labor.

In our baseline calibration, we abstracted from "spillover" effects in either amenities or local productivity. In the Online Appendix, we describe how we integrate such spillovers in two extensions of our model. We follow Diamond (2016) and model positive spillovers in amenities for college workers from the presence of other college workers, and we follow Rossi-Hansberg et al. (2019) in allowing productivity to increase both in total employment in a location and in the college share of employment. We use the elasticity estimates presented in those papers.

All else equal, higher amenities lead to lower wages in spatial equilibrium, offsetting some of the urban-biased wage growth that would otherwise occur. Column 5 shows that relative to the baseline calibration. The wage-density gradient increases less in Business Services in response to a decline in the ICT price when amenities are endogenous since the expansion of Business Services draws college workers into cities, which raises their amenities.

Column 6 of Table 5 presents the resulting wage-density gradients in 2015 with productivity spillovers. The model predicts a slightly larger increase in the wage-density gradients once productivities endogenously adjust. The reason is that endogenous growth in productivity in big cities draws a larger fraction of workers into the Business Services sector, which has a much larger wage-density gradient in 2015.

Overall, we conclude that the proposed mechanism is quantitatively important in explaining the urban-biased wage growth in the US economy.

References

Ahlfeldt, G. M. and E. Pietrostefani (2019): "The Economic Effects of Density: A Synthesis," *Journal of Urban Economics*, 111, 93–107.

Allen, T. and C. Arkolakis (2014): "Trade and the Topography of the Spatial Economy," *The Quarterly Journal of Economics*, 129, 1085–1140.

Almagro, M. and T. Dominguez-Iino (2019): "Location Sorting and Endogenous Amenities: Evidence from Amsterdam," *Working Paper*.

Artuç, E., S. Chaudhuri, and J. McLaren (2010): "Trade Shocks and Labor Adjustment: A Structural Empirical Approach," *American Economic Review*, 100, 1008–45.

Autor, D. and D. Dorn (2013): "The Growth of Low-skill Service Jobs and the Polarization of the US Labor Market," *American Economic Review*, 103, 1553–97.

Axtell, R. L. (2001): "Zipf Distribution of US Firm Sizes," *Science*, 293, 1818–1820.

Barro, R. J. and X. Sala-i Martin (1992): "Convergence," *Journal of Political Economy*, 100, 223–251.

Beaudry, P., M. Doms, and E. Lewis (2010): "Should the Personal Computer be Considered a Technological Revolution? Evidence from US Metropolitan Areas," *Journal of Political Economy*, 118, 988–1036.

Berry, C. R. and E. L. Glaeser (2005): "The Divergence of Human Capital Levels Across Cities," *Papers in Regional Science*, 84, 407–444.

Burstein, A. and J. Vogel (2017): "International Trade, Technology, and the Skill Premium," *Journal of Political Economy*, 125, 1356–1412.

Combes, P.-P., G. Duranton, L. Gobillon, D. Puga, and S. Roux (2012): "The Productivity Advantages of Large Cities: Distinguishing Agglomeration from Firm Selection," *Econometrica*, 80, 2543–2594.

Comin, D. A., D. Lashkari, and M. Mestieri (2020): "Structural Change with Long-run Income and Price Effects," *Econometrica*.

Davis, D. R. and J. I. Dingel (2020): "The Comparative Advantage of Cities," *Journal of International Economics*, 123, 103291.

Desmet, K. and E. Rossi-Hansberg (2009): "Spatial growth and industry age," *Journal of Economic Theory*, 144, 2477–2502.

DIAMOND, R. (2016): "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000," *American Economic Review*, 106, 479–524.

DURANTON, G. AND D. PUGA (2004): "Micro-foundations of Urban Agglomeration Economies," in *Handbook of Regional and Urban Economics*, Elsevier, vol. 4, 2063–2117.

ECKERT, F. (2019): "Growing Apart: Tradable Services and the Fragmentation of the US Economy," *Working Paper*.

ECKERT, F. AND T. KLEINEBERG (2019): "Can We Save the American Dream?" in *2019 Meeting Papers*, Society for Economic Dynamics, 1197.

EDEN, M. AND P. GAGGL (2019): "Capital Composition and the Declining Labor Share," *CESifo Working Paper*.

EHRLICH, M. V. AND H. G. OVERMAN (2020): "Place-based policies and spatial disparities across European cities," *Journal of Economic Perspectives*, 34, 128–49.

FORT, T. C. AND S. D. KLIMEK (2016): "The Effects of Industry Classification Changes on US Employment Composition," *Tuck School at Dartmouth mimeo*.

GANONG, P. AND D. SHOAG (2017): "Why has Regional Income Convergence in the US Declined?" *Journal of Urban Economics*, 102, 76–90.

GARCIA-MACIA, D., C.-T. HSIEH, AND P. J. KLENOW (2019): "How Destructive is Innovation?" *Econometrica*, 87, 1507–1541.

GAUBERT, C. (2018): "Firm Sorting and Agglomeration," *American Economic Review*, 108, 3117–53.

GIANNONE, E. (2022): "Skilled-biased Technical Change and Regional Convergence," *Working Paper*.

GYOURKO, J., C. MAYER, AND T. SINAI (2013): "Superstar Cities," *American Economic Journal: Economic Policy*, 5, 167–99.

HSIEH, C.-T. AND E. MORETTI (2019): "Housing Constraints and Spatial Misallocation," *American Economic Journal: Macroeconomics*, 11, 1–39.

KRUEGER, A. B. (1993): "How Computers Have Changed the Wage Structure: Evidence from Microdata, 1984–1989," *The Quarterly Journal of Economics*, 108, 33–60.

KRUGMAN, P. (1991): "Increasing Returns and Economic Geography," *Journal of Political Economy*, 99, 483–499.

KRUSELL, P., L. E. OHANIAN, J.-V. RÍOS-RULL, AND G. L. VIOLANTE (2000): "Capital-skill Complementarity and Inequality: A Macroeconomic Analysis," *Econometrica*, 68, 1029–1053.

LASHKARI, D., A. BAUER, AND J. BOUSSARD (2022): "Information Technology and Returns to Scale," *Available at SSRN 3458604*.

MORETTI, E. (2012): *The New Geography of Jobs*, Houghton Mifflin Harcourt.

——— (2013): "Real Wage Inequality," *American Economic Journal: Applied Economics*, 5, 65–103.

PIKETTY, T. AND E. SAEZ (2003): "Income inequality in the United States, 1913–1998," *The Quarterly Journal of Economics*, 118, 1–41.

REDDING, S. J. (2016): "Goods Trade, Factor Mobility and Welfare," *Journal of International Economics*, 101, 148–167.

REDDING, S. J. AND E. ROSSI-HANSBERG (2017): "Quantitative Spatial Economics," *Annual Review of Economics*, 9, 21–58.

ROSSI-HANSBERG, E., P.-D. SARTE, AND F. SCHWARTZMAN (2019): "Cognitive Hubs and Spatial Redistribution," Tech. rep., National Bureau of Economic Research.

RUBINTON, H. (2019): "The Geography of Business Dynamism and Skill-Biased Technical Change," *Working Paper*.

RUGGLES, S., M. SOBEK, T. ALEXANDER, C. A. FITCH, R. GOEKEN, P. K. HALL, M. KING, AND C. RONNANDER (2015): "Integrated Public Use Microdata Series: Version 6.0," Tech. rep., IPUMS.

SATO, R. (1977): "Homothetic and Non-homothetic CES Production Functions," *The American Economic Review*, 67, 559–569.

SCALA, D. J. AND K. M. JOHNSON (2017): "Political Polarization Along the Rural-urban Continuum? The Geography of the Presidential Vote, 2000–2016," *The ANNALS of the American Academy of Political and Social Science*, 672, 162–184.

TOLBERT, C. M. AND M. SIZER (1996): "US Commuting Zones and Labor Market Areas: A 1990 Update," Tech. rep., United States Department of Agriculture, Economic Research Service.

TROTTNER, F. (2019): "Who Gains from Scale?" *Working Paper*.

ONLINE APPENDIX FOR "URBAN-BIASED GROWTH" BY FABIAN ECKERT, SHARAT GANAPATI, AND CONOR WALSH

A. ADDITIONAL FIGURES AND TABLES

A.1 Details on the Urban Wage Gradient Over Time

In Figure 2, we showed that the wage-density coefficient computed across US commuting zones using QCEW data and 1980 population density numbers roughly doubled since 1980. In this section, we show that the same result holds in a number of different data sets, with different population density definitions, across counties, and in other countries.

Alternative Data Sources. In Figure OA.11a, we show the wage-density coefficient for each year computed in the QCEW, the LBD, the Decennial Population Census/American Community Survey (Census/ACS), and the County Business Patterns (CBP). The CBP is derived from the Business Register (BR) and excludes many public employees, as well as those in the agricultural sector. The QCEW, CBP, and LBD are all broadly similar and exhibit similar levels and trends. The point estimates from the Census/ACS data are somewhat lower, but exhibit similar time trends, with a sharp rise from 1980-2000 and a leveling off from 2000-2015.

Alternative Density and Size Measures. In Figure OA.11b, we show the wage-density coefficient in the QCEW using different measures of commuting zone density. First, we re-compute commuting zone population density in each year by dividing commuting zone total population by commuting zone total area. Second, we use the 1980 population density of a commuting zone. Third, we use the 1980 tract-weighted density of a commuting zone. In constructing this density, we consider the density of each census tract and create an aggregate commuting zone density by taking the population-weighted mean across tracts; this de-emphasizes rural tracts and empty land, e.g., the edges of the LA commuting zone. Finally, we show the wage-population elasticity instead of the wage-density elasticity, using 1980 commuting zone populations. Broadly, all coefficients exhibit similar trends.

Alternative Spatial Resolution. In Figure OA.11c, we show the wage-density coefficient in the QCEW estimated across counties instead of commuting zones. The wage-density coefficient estimated on county data is lower but shows a trend to that of the commuting zone estimates over time.

Alternative Countries. In Figure OA.11d, we show the wage-density coefficient computed across regions within the EU-15 countries. Instead of wages, the outcome variable is GDP per worker, and region size is measured in employment rather than population density. Europe shows trends similar to what is observed in the US; the GDP-region size elasticity doubles from .04 in 1980 to .08 in 2010.

A.2 College to Non-College Ratio by Sector

Figure OA.1 shows the college to non-college worker ratio by NAICS-1 sector for 1980 and 2015. The Education and Medical sectors have the highest ratio, largely because almost all teachers have college degrees. Business Services have the second highest ratio of college to non-college workers in both years, and have a very similar ratio to Education and Medical in 2015.

A.3 Disaggregated Industry Detail within Sectors

In the body of the paper, we present all results on the level of 1-digit sectors. Here, we present key results at the 2-digit NAICS level instead.

Figure OA.2 replicates Figure 3 in the main part of the paper on the 2-digit NAICS level. The industries within Business Services that are contributing most to the urban bias are in descending order: Professional Services, Finance, Information, Admin and Waste, Management, and Real Estate.

Our baseline decomposition is silent on the role of sector size. If an industry contributes a lot of employment in every location, a small amount of differential wage growth across regions translates into a large contribution to urban-biased growth. We conduct another decomposition to understand which industry's contribution is due to this "large industry effect."

We ask what the contribution of each industry would be if it accounted for the same fraction of national employment, i.e., $1/S$ of national employment where S is the number of industries. We decompose local wage growth into a component that captures local growth if all sectors had the same aggregate size and a "residual." Industries for which this residual is large contribute more to urban-biased growth primarily because they are large.

(OA.1)

$$\Delta w_r = \underbrace{\sum_i \phi'_{r,i} \frac{1}{S} w'_{r,i} - \phi_{r,i} \frac{1}{S} w_{r,i}}_{\text{"Sizeless" Growth}} + \underbrace{\sum_i (\mu'_{r,i} - \phi'_{r,i} \frac{1}{S}) w'_{r,i} - (\mu_{r,i} - \phi_{r,i} \frac{1}{S}) w_{r,i}}_{\text{Role of Size}},$$

where $\phi_{r,i}$ is the fraction of total sector i employment accounted for by region r (sums to 1 across regions within industry). The first term in the last line is local growth in an economy in which all sectors have the same aggregate size. The second term captures the role of differences in sectoral size.

Figure OA.3 presents the results. The size adjustments make the Business Services sector even more important in contributing to urban-biased growth The intuition for this result is that the sector is relatively small and so contributed a lot of urban-biased growth "per worker."

Figure OA.4 shows ICT usage for 2-digit industries within each sector. Almost all sub-industries within the Business Services sector are more intensive users of ICT than any other industry in the US economy.

A.4 Moments Used in the Model Calibration

Rent Index. We construct a commuting zone rent index. We use the 1980 version of the index to calibrate housing supply in the model and the 2015 version to compare with our model predictions for 2015. To construct the index, we use microdata on reported gross rents from the US Census and American Community Survey, and regress them on the age of the building, the number of rooms, and commuting zone fixed effects, separately in 1980 and 2015. The commuting zone fixed effects serve as our index. They can be interpreted as the price of a unit of observationally equivalent housing in each commuting zone. Figure OA.5 shows the rent index across commuting zones for 1980 and 2015.

Average Establishment Size Differences. In our theory, firms are larger in larger locations to finance the increased entry cost. We use data on differences in establishment size across commuting zones to discipline how entry costs vary with population size, i.e., to calibrate τ_s, the entry cost shifter, and η_s, the entry elasticity. The data for average establishment size by sector come from the QCEW data. In Figure OA.6, we plot average establishment size against population density, for Business Services establishments and establishments in all other sectors.

A.5 Additional Model Results

In Figure OA.7, we show the predictions for changes in the rent gradient between 1980 and 2015 as a result of the counterfactual price change. The rent-density gradient increased markedly between the two years as a result of the decline in ICT prices.

In Figure OA.8, we show the predictions for average establishment size in the model. The establishment size gradient steepens slightly in both sectors.

Figure OA.9 shows average wages in the model and data across US commuting zones ordered by density in 1980 and 2015. Overall, in the model, the average wages are very similar to those in the data in all deciles of commuting zone density. The decline in ICT prices does not explain the entire increase in the wage-density gradient in the Business Services sector.

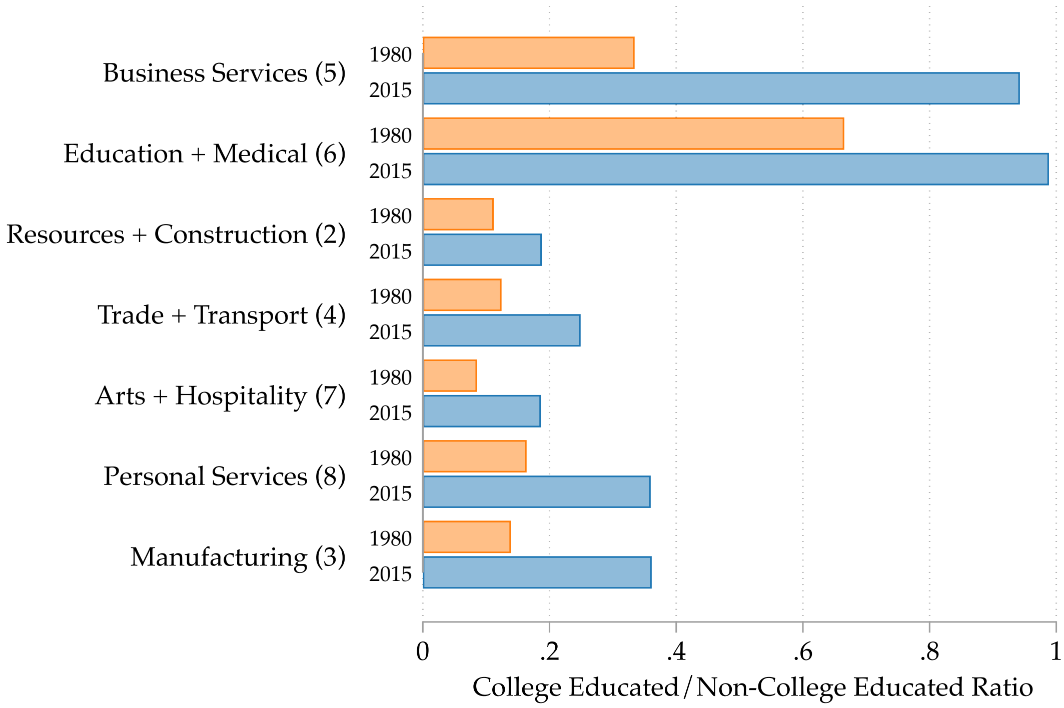
Figure OA.10 plots the fit of the model for ICT per employee and the college share of employment across different firm size bins against data from the ACES (Panel (a)) and the CPS (Panel (b)).

The first row of Table OA.1 shows how the calibrated location productivity terms vary with population density, for each sector and worker type. Column 1 shows the correlation for college productivities in Business Services, Column 2 for non-college productivities in Business Services, Column 3 for college productivities outside Business Services, and Column 4 for college productivities outside Business Services. Business Services sector productivity for college workers is increasing the most in population density suggesting that high population density locations have a distinct comparative advantage in college-level Business Services work.

The second row of Table OA.1 shows how the calibrated location amenity terms vary with population density, for each worker type. Column 1 shows the correlation for college location. amenities and Column 2 for non-college location amenities. Amenities for college-educated workers are increasing more in population density than amenities for non-college workers.

Model Robustness. Table OA.2 shows robustness of the main results in Table 5 obtained by varying the sectoral elasticity of substitution in final good production, ζ_F. Higher values lead to far too much value added accruing to Business Services, while the opposite is true for lower values. Table OA.2 serves as a justification for our "Baseline" choice of $\zeta_F = 1.2$.

FIGURE OA.1: COLLEGE/NON-COLLEGE WORKER RATIOS BY SECTOR



Notes: This figure shows the ratio of college educated workers to non-college educated workers in both 1980 and 2015. The data are from the US Census/ACS.

A.6 Endogenous Local Fundamentals

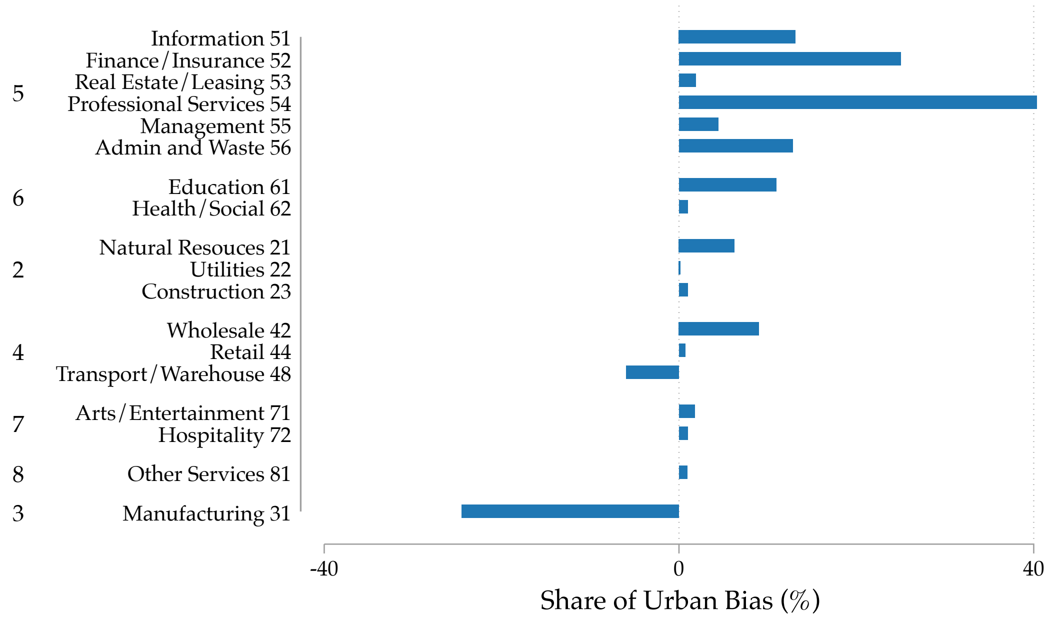
A long literature suggests that local productivities and amenities may be endogenous functions of the size and composition of a location's workforce. In our main calibration, we abstracted from such "spillover" effects. We investigate their qualitative role in affecting the strength of our mechanism.

Diamond (2016) provides direct evidence that the number of amenities for high-skill workers is an increasing function of the share of high-skill workers in a location. We change the location amenity term for high-skill workers in our model to incorporate that channel by setting $A_r^H = \bar{A}_r^H (L_r^H / L_r^L)^\chi$. We borrow the parameter χ from Diamond (2016). Note that we do not need to re-calibrate our model; we can simply decompose the calibrated amenities into an endogenous and an exogenous part ($\bar{A}_r^H$). Column 5 of Table 5 presents the resulting wage-density gradients in 2015.

Rossi-Hansberg et al. (2019) provide estimates for productivity spillovers. We change the specification of productivities in our model as follows:

$$Z_{r,s}^e = \bar{Z}_{r,s}^e L_r^{\omega_1^s} (L_r^H / L_r)^{\omega_2^s}. \quad \text{(OA.2)}$$

FIGURE OA.2: SECTORAL ORIGINS OF URBAN-BIASED WAGE GROWTH ACROSS NAICS-2 INDUSTRIES



Notes: This figure shows the share of urban-biased wage growth between 1980 and 2015 accounted for by each NAICS-2 industry using the decomposition in equation (1). We compare wage growth between the commuting zones with the highest population density jointly accounting for 50% and all remaining commuting zones. The figure uses the Longitudinal Business Database.

and use their parameter estimates by sector. Column 6 of Table 5 presents the resulting wage-density gradients in 2015.

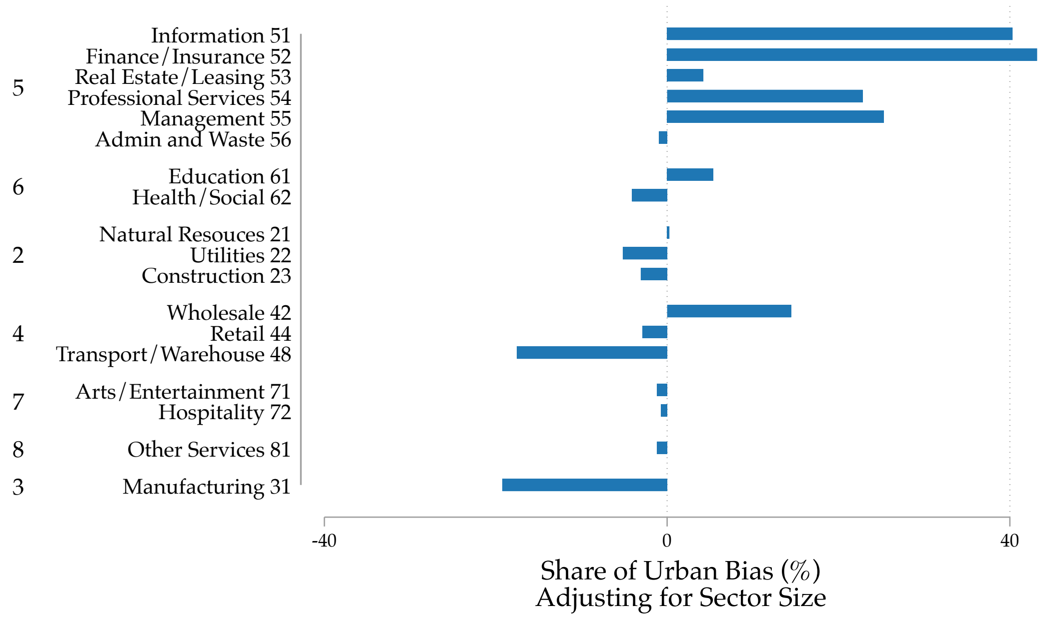
TABLE OA.1: LOCATION FUNDAMENTALS AND EMPLOYMENT DENSITY

Productivity Term	$\log Z^H_{r,N5}$	$\log Z^L_{r,N5}$	$\log Z^H_{r,O}$	$\log Z^L_{r,O}$
Log Density	0.169	0.116	0.121	0.0750
R^2	0.570	0.625	0.539	0.555

Amenity Term	$\log A^H_r$	$\log A^L_r$
Log Density	1.218	1.060
R^2	0.718	0.739

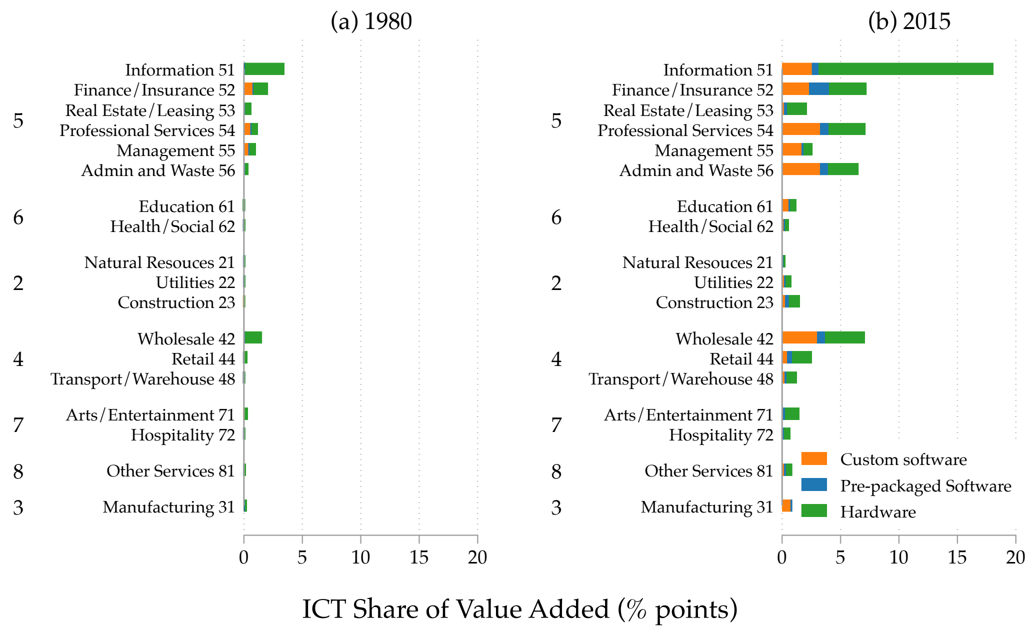
Notes: This table presents six regressions of calibrated model objects for 1980 on commuting zone density. The top panel regresses (log) density on the underlying productivity ($\log Z$) across commuting zones r for four different groups of workers, those in Business Services (N5) and those in other sectors (O), separately for college-educated (H) and non-college-educated (L) workers. The bottom panel shows the correlation between location amenities and population density for college and non-college workers.

Figure OA.3: Sectoral Origins of Urban-Biased Wage Growth across NAICS-2 industries with Industry Size Adjustment



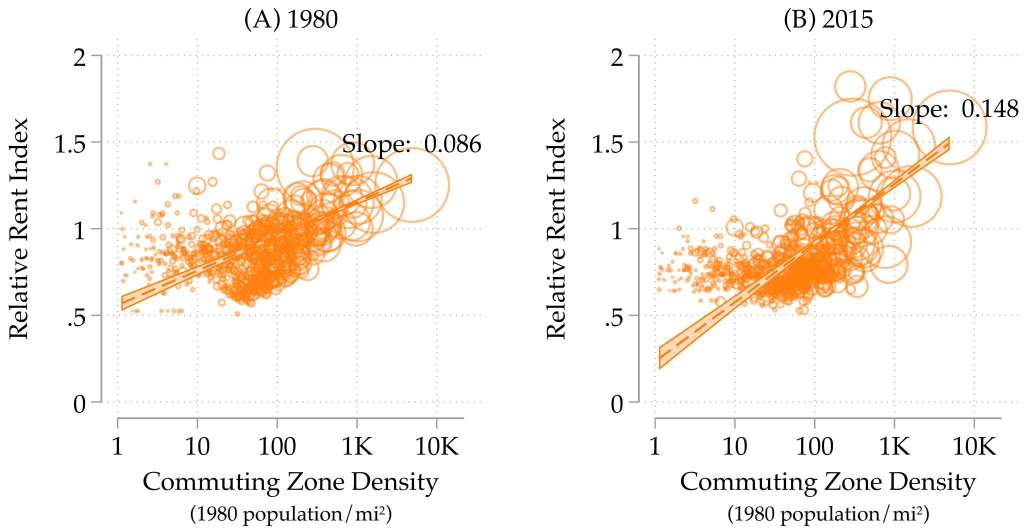
Notes: This figure shows the share of urban-biased wage growth between 1980 and 2015 accounted for by each NAICS-2 industry using the sizeless growth decomposition shown in equation OA.1. We compare wage growth between the commuting zones with the highest population density jointly accounting for 50% and all remaining commuting zones. The figure uses the Longitudinal Business Database.

Figure OA.4: ICT Value Added Shares Across NAICS-2 Industries



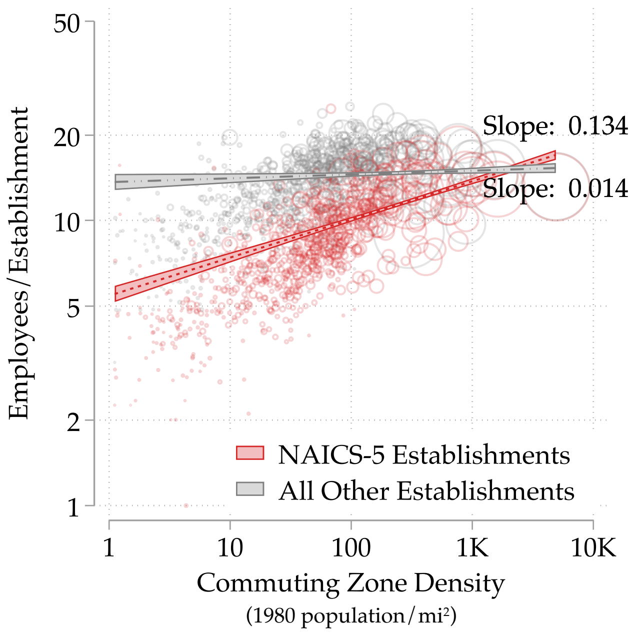
Notes: We show the share of real ICT of value added by industry in 2012 dollars from equation (12). Data are from the BEA. Prior to 1987, labor share uses data from the QCEW. Proprietary software refers to BEA codes ENS2 and ENS3, pre-packaged software refers to ENS1, and hardware to EP1A-EP31. Sectors ordered by contribution to urban-biased growth.

FIGURE OA.5: COMMUTING ZONE RENT INDEX



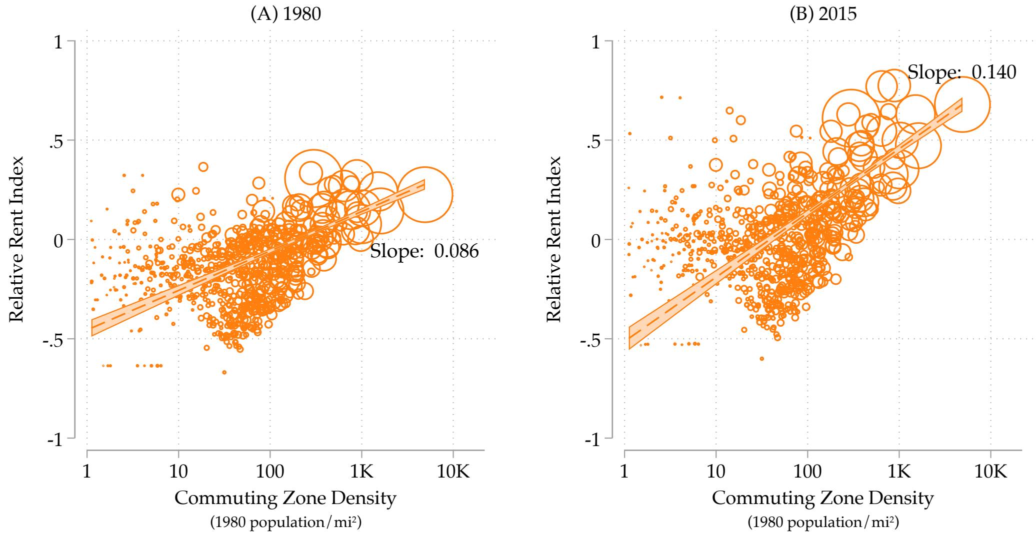
Notes: This figure plots relative rent indexes against commuting zone population density. Mean rent is normalized to 1. Data are from the US Census and American Community Survey.

FIGURE OA.6: AVERAGE ESTABLISHMENT SIZE ACROSS SPACE



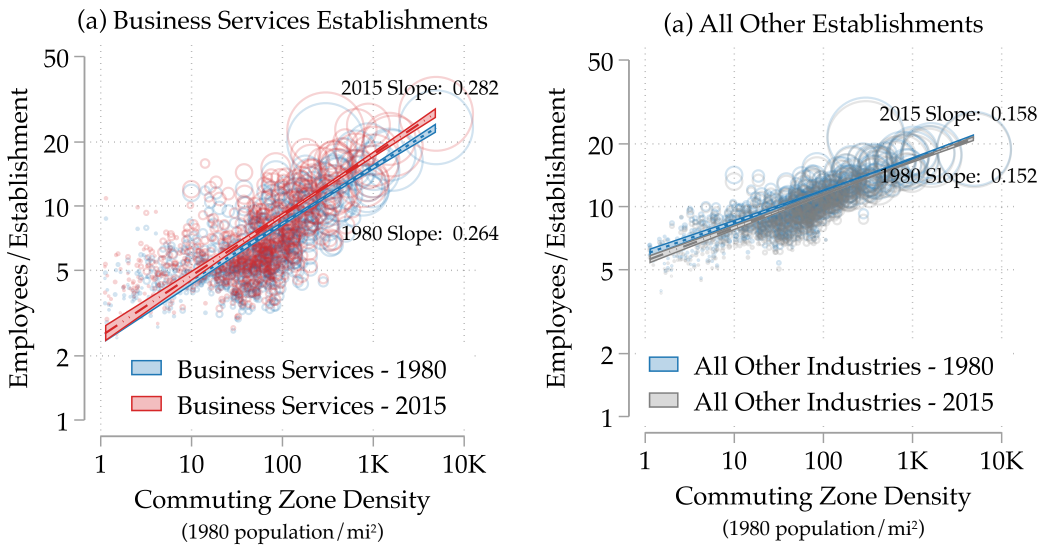
Notes: This figure plots average establishment size against population density using data from the 2015 QCEW. The regression coefficient is weighted by commuting zone population.

Figure OA.7: Commuting Zone Rent Index in the Model



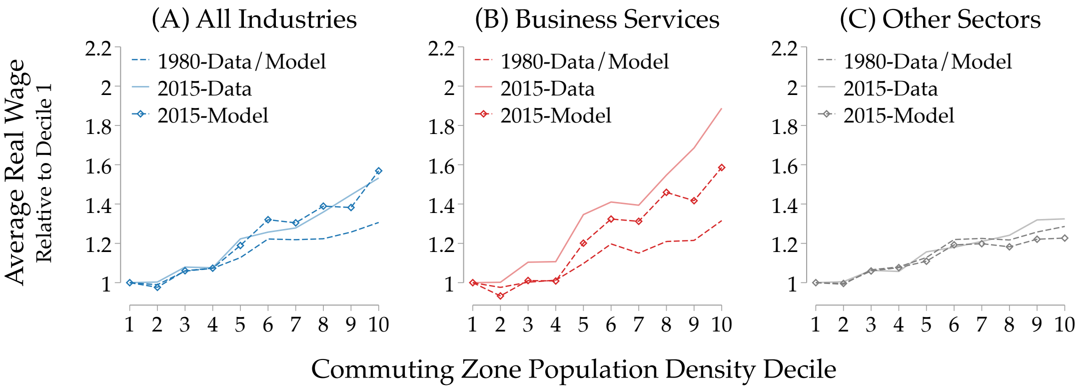
Notes: This figure plots relative rent indexes against commuting zone population density. Mean rent is normalized to 0.

Figure OA.8: Average Establishment Size Across Space in the Model



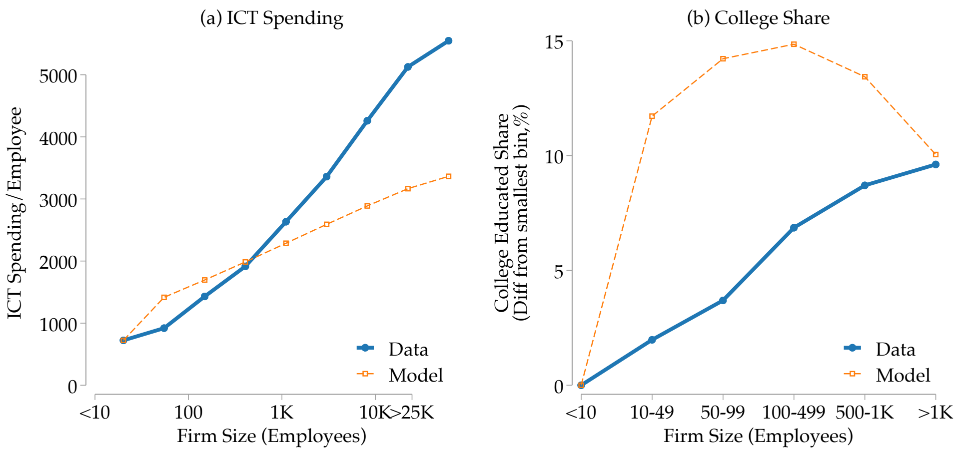
Notes: This figure plots average establishment size against population density in the calibrated model for both 1980 and 2015.

FIGURE OA.9: URBAN-BIASED GROWTH – MODEL AND DATA



Notes: This figure compares the counterfactual wage outcomes across space by sector with the data. The 1980 data and model are identical by construction.

FIGURE OA.10: FACTOR RATIOS ACROSS THE FIRM-SIZE DISTRIBUTION



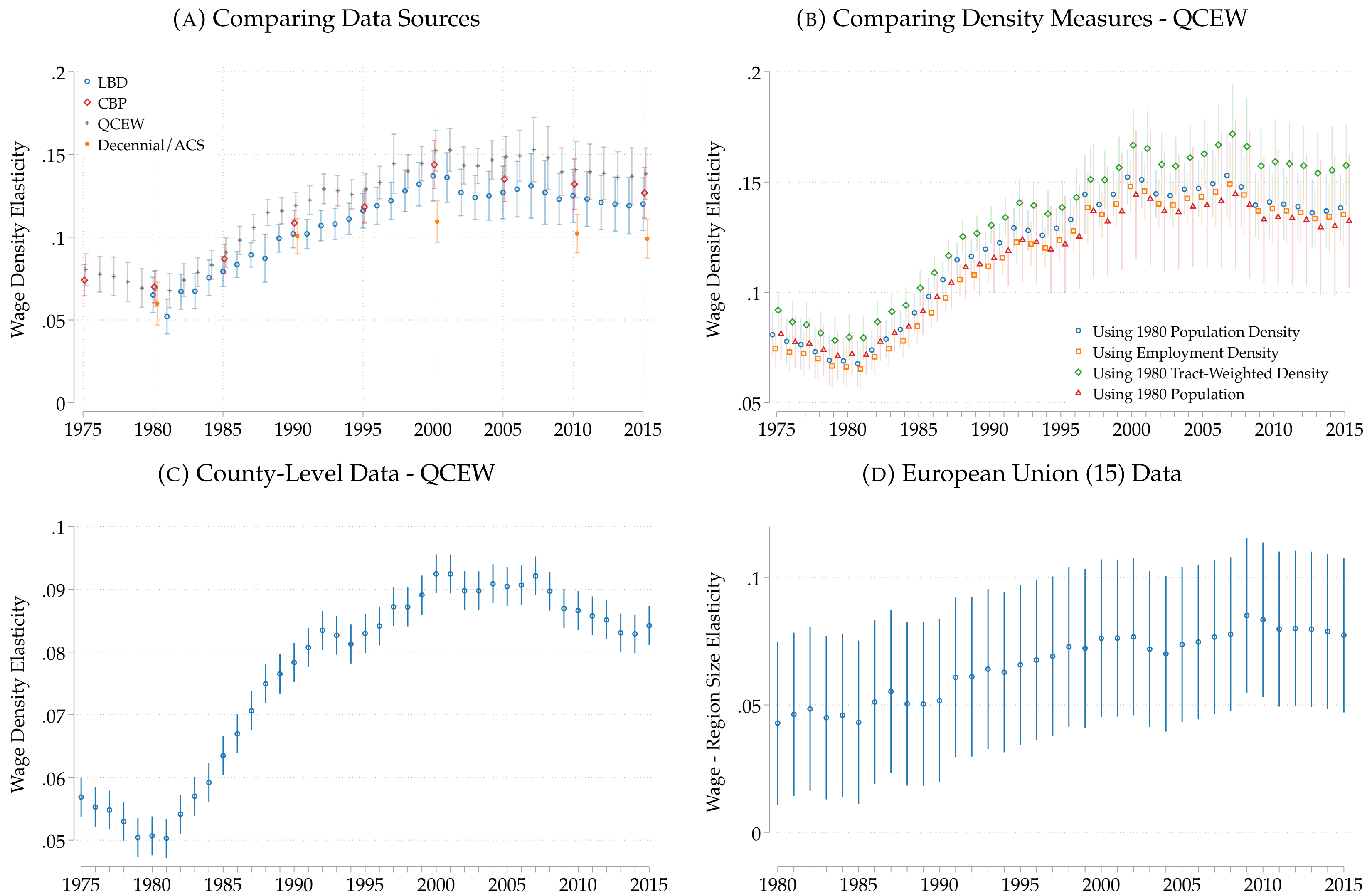
Notes: The left panel plots ICT spending per worker using the 2013 ACES/ICT survey matched to the LBD from the US Census. The right panel plots the difference in college educated worker share compared with the firms with under 10 employees in the 1992 Current Population Survey (CPS) from the US Census.

TABLE OA.2: WAGE-DENSITY COEFFICIENT IN DATA AND MODEL: ROBUSTNESS

		2015				
	1980	Data	Baseline	$\zeta = 1.4$	$\zeta = 1.1$	Fixed Labor
Urban Wage Gradient						
Business Services	0.067	0.151	0.133	0.138	0.100	0.056
Other Sectors	0.056	0.068	0.048	0.052	0.038	0.050
Aggregate	0.060	0.099	0.107	0.168	0.062	0.050
Payroll Shares						
Business Services	0.155	0.270	0.342	0.563	0.220	0.144
Other Sectors	0.845	0.730	0.658	0.437	0.780	0.856
Employment Shares						
Business Services	0.143	0.195	0.166	0.177	0.157	0.150
Other Sectors	0.857	0.805	0.834	0.823	0.843	0.850

Notes: Gradients computed use the ACS/Census for 1980 and 2015, weighting by 1980 population shares.

FIGURE OA.11: WAGE-DENSITY COEFFICIENTS



Notes: This figure shows the wage-density gradient coefficients β_t across commuting zone, r, for each year from the regression $\ln w_{rt} = \alpha + \phi_t + \beta_t \ln density_r + \epsilon_{rt}$. Panel (D) replicates Ehrlich and Overman (2020) for all years from 1980-2015. The sample covers EU-15 countries and reports the coefficients β_t across regions, r, for each year from the regression $\ln w_{rt} = \alpha + \phi_t + \beta_t \ln employment_r + \epsilon_{rt}$. $\ln employment_r$ refers to the size of the workforce in region r.

B. Theory Appendix

B.1 The Baseline Model

We first show how to solve an individual intermediate input firm's problem and then derive a set of additional results.

B.1.1 The Firm's Problem

An individual firm produces with the following production technology:

$$\left(\frac{Z_r l}{y}\right)^{\frac{\sigma-1}{\sigma}} + \left(\frac{Z_r k}{y^{1+\epsilon}}\right)^{\frac{\sigma-1}{\sigma}} = 1,$$

where l denotes labor, k denotes capital, and σ indexes the substitutability of these factors in production.

We can write the firm's cost minimization problem:

$$\min_{k,l} pk + wl + \lambda\left(1 - \left(\frac{Z_r l}{y}\right)^{\frac{\sigma-1}{\sigma}} - \left(\frac{Z_r k}{y^{1+\epsilon}}\right)^{\frac{\sigma-1}{\sigma}}\right)$$

The resulting first order conditions are given by:

$$w = \lambda\frac{\sigma-1}{\sigma}\left(\frac{Z_r l}{y}\right)^{-\frac{1}{\sigma}}\frac{Z_r}{y} \quad \text{and} \quad p = \lambda\frac{\sigma-1}{\sigma}\left(\frac{Z_r k}{y^{1+\epsilon}}\right)^{-\frac{1}{\sigma}}\frac{Z_r}{y^{1+\epsilon}}$$

Plugging the first order conditions back into the cost function:

$$C = \lambda\frac{\sigma-1}{\sigma}\left[\left(\frac{Z_r l}{y}\right)^{\frac{\sigma-1}{\sigma}} + \left(\frac{Z_r k}{y^{1+\epsilon}}\right)^{\frac{\sigma-1}{\sigma}}\right] = \lambda\frac{\sigma-1}{\sigma}$$

Combining the expression for the cost function with the first order conditions for the individual factors yields expressions for factor demands:

$$l = C^{\sigma}w^{-\sigma}\left(\frac{y}{Z_r}\right)^{1-\sigma} \quad \text{and} \quad k = C^{\sigma}p^{-\sigma}\left(\frac{y^{1+\epsilon}}{Z_r}\right)^{(1-\sigma)} \tag{OA.3}$$

Next, we define the central object in our analysis, the firm's cost share of capital:

$$\theta_r \equiv \frac{pk}{C} = C^{\sigma-1}p^{1-\sigma}\left(\frac{y^{1+\epsilon}}{Z_r}\right)^{(1-\sigma)} \tag{OA.4}$$

which satisfies

$$\text{(OA.5)} \qquad \frac{\theta_r}{1-\theta_r} = w_r^{\sigma-1} p^{1-\sigma} y^{\epsilon(1-\sigma)}.$$

These factor demands also give rise to an expression for the cost function:

$$\text{(OA.6)} \qquad C(y) = \left(w_r^{1-\sigma} \left(\frac{y}{Z_r} \right)^{1-\sigma} + p^{1-\sigma} \left(\frac{y^{1+\epsilon}}{Z_r} \right)^{1-\sigma} \right)^{\frac{1}{1-\sigma}}$$

The cost function in equation OA.6 is the cost of producing a quantity y given optimally chosen input quantities. Note how the non-homotheticity ($\epsilon \neq 0$) makes the marginal cost curves rise with output.

The final good firm aggregates firm varieties with an elasticity of substitution ι into a final good for consumers. Standard arguments for CES utility functions imply that the demand for an individual firm's output can be written:

$$y = p^{-\iota} P^{\iota-1} Y \Rightarrow p = y^{-\frac{1}{\iota}} P^{\frac{\iota-1}{\iota}} Y^{\frac{1}{\iota}} \equiv y^{-\frac{1}{\iota}} \mathcal{D}$$

where P is the usual price index and Y is aggregate consumer spending. We let P be the numeraire, and let $\mathcal{D}$ be an index of demand, so that revenue is

$$py = yy^{-\frac{1}{\iota}} \mathcal{D} \equiv y^{\zeta} \mathcal{D},$$

where $\zeta = 1 - 1/\iota \in (0, 1)$.

With an expression for the demand curve in hand, we can write the firm's profit maximization problem as follows:

$$\text{(OA.7)} \qquad \max_y \pi(y) = \max_y \left(\mathcal{D} y^{\zeta} - C(y) \right)$$

where we denote the firm's profit function by $\pi(\cdot)$. The first order condition with respect to output is given by:

$$\begin{aligned}
\zeta \mathcal{D} y^{\zeta-1} &= (1+\epsilon) C^{\sigma} p^{1-\sigma} Z_r^{\sigma-1} y^{(1+\epsilon)(1-\sigma)} y^{-1} + C^{\sigma} w^{1-\sigma} \left(\frac{y}{Z_r} \right)^{1-\sigma} y^{-1} \\
&= \left[(1+\epsilon) C \theta_r + C(1-\theta_r) \right] y^{-1} \\
&= y^{-1} C \epsilon \theta_r + y^{-1} C
\end{aligned}$$

But then we have that for the profit-maximizing output quantity the following holds:

$$\zeta \mathcal{D} y^{\zeta} = C\left(\epsilon\theta_r + 1\right) \tag{OA.8}$$

Note that optimal output is only implicitly defined since C and θ are functions of y. Finally, we present some useful expressions that employ the envelope theorem

$$\frac{d\pi_r}{dw_r} = -\frac{\partial C}{\partial w_r}$$

Thus we have

$$\begin{aligned}\frac{d\pi_r}{dw_r}\frac{w}{\pi_r} &= -\left(p^{1-\sigma}\left(\frac{Z_r}{y^{1+\epsilon}}\right)^{\sigma-1} + w_r^{1-\sigma}\left(\frac{Z_r}{y}\right)^{\sigma-1}\right)^{\frac{1}{1-\sigma}-1} w_r^{1-\sigma}\left(\frac{Z_r}{y}\right)^{\sigma-1}/\pi_r \\ &= -\frac{C^{\sigma} w_r^{1-\sigma}\left(\frac{Z_r}{y}\right)^{\sigma-1}}{\pi_r} = -(1-\theta_r)\frac{C}{\pi_r}\end{aligned}$$

Now from the definition of profit

$$\frac{\pi_r}{C} = \frac{\mathcal{D}y^{\zeta}}{C} - 1.$$

From (OA.8) we have

$$\zeta\mathcal{D}\frac{y^{\zeta}}{C} = \epsilon\theta_r + 1,$$

and so we can write

$$(1-\theta_r)\frac{C}{\pi_r} = \frac{1-\theta_r}{\frac{1}{\zeta}(\epsilon\theta_r+1)-1} = \zeta\frac{1-\theta_r}{\epsilon\theta_r+1-\zeta}. \tag{OA.9}$$

Similarly for the effect of productivity on profit

$$\frac{d\pi_r}{dZ_r}\frac{Z_r}{\pi_r} = \frac{C}{\pi_r} = \zeta\frac{1}{\epsilon\theta_r+1-\zeta}. \tag{OA.10}$$

B.1.2 Useful Lemmas

Lemma 1. *The total derivative of the cost function is given by:*

$$d\log C = \theta_r d\log p + (1-\theta_r)d\log w_r + (1+\epsilon\theta_r)d\log y - d\log Z_r$$

Proof. Taking the total derivative of equation (OA.6), we obtain:

$$\begin{aligned}(1-\sigma)d\log C = &(\sigma-1)d\log Z_r \\ &+ (1-\sigma)\theta_r d\log p \\ &+ (1-\sigma)(1-\theta_r)d\log w_r \\ &+ \left[\theta_r(1+\epsilon)(1-\sigma)+(1-\theta_r)(1-\sigma)\right]d\log y\end{aligned}$$

or

$$d\log C = \theta_r d\log p + (1-\theta_r)d\log w_r + (1+\epsilon\theta_r)d\log y - d\log Z_r$$

□

Lemma 2. *The total derivative of the optimal output equation is given by:*

$$\zeta d\log y = d\log C + \frac{\epsilon\theta_r}{\epsilon\theta_r+1}d\log\theta_r - (\epsilon\theta_r+1)\,d\log\mathcal{D}$$

Proof. Taking the total derivative of equation (OA.8), we obtain:

$$\zeta^2\mathcal{D}y^{\zeta}d\log y + \zeta\mathcal{D}y^{\zeta}d\log\mathcal{D} = (\epsilon\theta_r+1)\,Cd\log C + C\epsilon\theta_r d\log\theta_r$$

But then we can use the expression for optimal output in (OA.8):

$$\begin{aligned}\zeta C\,(\epsilon\theta_r+1)\,d\log y + C\,(\epsilon\theta_r+1)\,d\log\mathcal{D} &= (\epsilon\theta_r+1)\,Cd\log C + C\epsilon\theta_r d\log\theta_r \\ \zeta\,(\epsilon\theta_r+1)\,d\log y + (\epsilon\theta_r+1)\,d\log\mathcal{D} &= (\epsilon\theta_r+1)\,d\log C + \epsilon\theta_r d\log\theta_r \\ \zeta d\log y + (\epsilon\theta_r+1)\,d\log\mathcal{D} &= d\log C + \frac{\epsilon\theta_r}{\epsilon\theta_r+1}d\log\theta_r\end{aligned}$$

□

Lemma 3. *The total derivative of the capital cost share is given by:*

$$\frac{1}{(1-\theta_r)(1-\sigma)}d\log\theta_r = \epsilon d\log y + d\log p - d\log w_r$$

Proof. Taking the total derivative of equation (OA.5), we obtain:

$$\begin{aligned}\frac{1}{(1-\theta)^2}d\theta_r = &\left(\frac{p}{w_r}\right)^{1-\sigma}y^{\epsilon(1-\sigma)-1}\epsilon(1-\sigma)dy \\ &+\left(\frac{p}{w_r}\right)^{1-\sigma}y^{\epsilon(1-\sigma)}(1-\sigma)p^{-1}dp - \left(\frac{p}{w_r}\right)^{1-\sigma}y^{\epsilon(1-\sigma)}(1-\sigma)w_r^{-1}dw_r\end{aligned}$$

But then:

$$\frac{1}{(1-\theta_r)(1-\sigma)} d\log\theta_r = \epsilon d\log y + d\log p - d\log w_r.$$

□

B.1.3 Proof of Proposition 1

We can totally differentiate the free entry condition to find

$$d\log w_r = \frac{\partial \log \pi(w_r, Z_r, p)}{\partial \log w_r} d\log w_r + \frac{\partial \log \pi(w_r, Z_r, p)}{\partial \log Z_r} d\log Z_r$$

noting that prices p and demand $\mathcal{D}$ are constant across locations. Using the expressions in (OA.9) and (OA.10) we can write

$$\text{(OA.11)} \qquad \frac{d\log w_r}{d\log Z_r} = \frac{\frac{\zeta}{\epsilon\Theta_r+1-\zeta}}{1+\zeta\frac{1-\Theta_r}{\epsilon\Theta_r+1-\zeta}} = \frac{\zeta}{(\epsilon-\zeta)\Theta_r+1},$$

which is always positive. As such, for given amenities, locations with a higher productivity Z_r will have higher wages w_r. Similarly, since the labor supply curve slopes upward, these locations will have larger populations L_r.

As for capital cost shares θ_r, combining Lemma 1 and Lemma 2 gives

$$(\zeta - 1 - \Theta_r\epsilon) d\log y = -d\log Z_r + (1-\Theta_r) d\log w_r + \frac{\Theta_r\epsilon}{(1+\Theta_r\epsilon)} d\log\Theta_r,$$

and plugging this into Lemma 3

$$\frac{1}{(1-\Theta_r)} d\log\Theta_r = \frac{\epsilon(1-\sigma)}{(\zeta-1-\Theta_r\epsilon)}\left[-d\log Z_r + (1-\Theta) d\log w_r + \frac{\Theta_r\epsilon}{(1+\Theta_r\epsilon)} d\log\Theta_r\right]$$
$$- (1-\sigma) d\log w_r.$$

Then using (OA.11) we get

$$\left[\frac{(1+\Theta_r\epsilon-\zeta)}{(1-\Theta_r)} - \frac{\epsilon(1-\sigma)\Theta_r\epsilon}{(1+\Theta_r\epsilon)}\right]\frac{d\log\Theta_r}{d\log Z_r} = \epsilon(1-\sigma) - (1-\sigma)[1+\epsilon-\zeta]\frac{\zeta}{1+\Theta_r(\epsilon-\zeta)}$$

or

$$\frac{d\log\Theta_r}{d\log Z_r} = \frac{\epsilon - \frac{\zeta[1+\epsilon-\zeta]}{1+\Theta_r(\epsilon-\zeta)}}{\frac{(1+\Theta_r\epsilon-\zeta)}{(1-\Theta_r)(1-\sigma)} - \frac{\epsilon^2\Theta_r}{(1+\Theta_r\epsilon)}}.$$

Note that the denominator is always positive since

$$\begin{aligned} 1+\Theta_r\epsilon-\zeta > 1 > \epsilon^2\theta_r \\ (1-\Theta_r)(1-\sigma) < 1 < (1+\theta_r\epsilon). \end{aligned}$$

But then we have that the gradient is positive as long as

$$\epsilon - \frac{\zeta[1+\epsilon-\zeta]}{1+\Theta_r(\epsilon-\zeta)} > 0$$

which always holds if $\epsilon > \zeta$.

B.1.4 Proof of Proposition 2

Totally differentiating the free entry condition in (10), we find

$$\mathcal{E}dw_r = \frac{\partial \pi^\star}{\partial w_r}dw_r + \frac{\partial \pi^\star}{\partial p}dp + \frac{\partial \pi^\star}{\partial \mathcal{D}}d\mathcal{D},$$

which can be written

$$\mathcal{E}dw_r = -\frac{C}{w_r}(1-\Theta_r)dw_r - \frac{C}{p}\Theta_r dp + \frac{d\mathcal{D}}{\mathcal{D}}C(1+\Theta_r\epsilon),$$

and substituting in the free entry condition yields

$$\left(\frac{1}{\zeta}C\,(1+\Theta_r\epsilon)-C\right)d\log w_r = -C(1-\Theta_r)d\log w_r - C\Theta d\log p + C(1+\Theta_r\epsilon)d\log\mathcal{D}$$

As a result we have

$$\frac{d\log w_r}{d\log p} = -\frac{1}{\frac{1}{\zeta}+\Theta_r(\frac{\epsilon}{\zeta}-1)}(\Theta_r-(1+\Theta_r\epsilon)\frac{d\log\mathcal{D}}{d\log p}) \tag{OA.12}$$

It can be shown that aggregate demand $\mathcal{D}$ falls as the capital price rises, so that $\log w_r$ is decreasing in $\log p$. It also follows from this expression that $\log w_r$ is falling faster in places with higher θ.

B.1.5 Changes in Aggregate Demand

Similar to before we totally differentiate the free entry condition to obtain:

$$\mathcal{E}w_r d\log w_r = -C(1-\Omega)d\log w_r + C(1+\Omega\epsilon)d\log\mathcal{D}$$

We combine the equilibrium expression for maximized profits with the free entry condition to obtain:

$$\mathcal{E}w_r = \frac{1}{\zeta}C\left[1+\Omega\epsilon\right] - C$$

Plugging this into the totally differentiated free entry condition, we obtain:

$$[\frac{1}{\zeta}(1+\Theta_r\epsilon) - \Theta_r]d\log w_r = (1+\Theta_r\epsilon)d\log\mathcal{D},$$

which we can re-arrange to yield:

$$\frac{d\log w_r}{d\log\mathcal{D}} = \frac{\zeta(1+\Theta_r\epsilon)}{(1+\Theta_r\epsilon)-\zeta\Theta_r},$$

which is always positive since $\epsilon > \zeta$. As a result, an increase in aggregate demand raises wages in all locations, always. Taking the partial derivative of this expression with respect to the capital cost share yields:

$$\partial\frac{d\log w_r}{d\log\mathcal{D}}/\partial\Theta_r = \frac{\zeta^2}{(1+\Theta_r\epsilon-\zeta\Theta_r)^2} > 0,$$

so that locations with a higher capital cost share see faster wage growth than locations with smaller capital cost shares.

B.2 The Neoclassical Baseline

In this subsection, we embed a neoclassical production function with capital-labor complementarity into a regional setting. We show that this setup always produces wage convergence in response to falling capital prices, at odds with the recent US experience of regional wage divergence.

Consider an economy with discrete locations indexed by r each host to a representative firm producing the same homogeneous good using the following technology:

$$y = F_r(K, L) \quad \text{with} \quad \sigma_r \equiv \frac{d \log K/L}{d \log \frac{\partial F_r}{\partial L} / \frac{\partial F_r}{\partial K}} < 1. \tag{OA.13}$$

The homogeneous good is traded freely across locations and all input and output markets are competitive. The price of the final good serves as numeraire. Capital is produced by a national representative firm that transforms the final good at a constant rate $\mathcal{Z}$ into capital. As a result, the price of capital is $p = 1/\mathcal{Z}$. There is a unit mass of workers who supply labor to each region with an arbitrary labor supply function, such that

$$L_r = M_r(w_r)$$

with the restriction that $\sum_r M_r(w_r) = 1$ for all vectors of regional wages $\{w_r\}$. This labor supply function nest the formulation of our model in Section 4, but permits many more general formulations. Labor markets clear in each location, and the capital market and final goods market clears nationally.

We now derive equations (7) and (8) in the main text. First, we totally differentiate the production function to obtain:

$$dy = \frac{\partial F_r}{\partial K} dK + \frac{\partial F_r}{\partial L} dL \tag{OA.14}$$

Since production is constant returns to scale and the output market is perfectly competitive, there are zero profits and $y = C_r(K, L)$, so that:

$$p = \Theta_r \frac{y}{K} \quad \text{and} \quad w_r = (1 - \Theta_r) \frac{y}{L}, \tag{OA.15}$$

where $\Theta_r = pK/(w_r L + pK)$. Combine this with the first order condition of the firm,

$$\frac{\partial F_r}{\partial K} = p \quad \text{and} \quad \frac{\partial F_r}{\partial L} = w_r,$$

and substitute it into equation OA.14, to obtain:

$$d \log y = \Theta_r d \log K + (1 - \Theta_r) d \log L. \tag{OA.16}$$

Now totally differentiate the expression $y = C(K, L)$:

$$yd \log y = pKd \log K + pKd \log p + w_r Ld \log L + w_r Ld \log w_r$$

Using the definition of the cost shares:

$$yd \log y = \Theta_r d \log K + \Theta_r d \log p + (1 - \Theta_r) d \log L + (1 - \Theta_r) d \log w_r$$

Finally, plugging in equation OA.16 and re-arranging:

$$\frac{d \log w_r}{d \log p} = -\frac{\Theta_r}{1 - \Theta_r} \tag{OA.17}$$

Next, divide the two expressions in equations OA.15:

$$\frac{w_r}{p} = \frac{1 - \Theta_r}{\Theta_r} \frac{K}{L}$$

Now totally differentiate to obtain:

$$d \log \frac{w_r}{p} = d \log \frac{1 - \Theta_r}{\Theta_r} + d \log \frac{K}{L}$$

Finally, re-arranging:

$$\frac{d \log \frac{\Theta_r}{1 - \Theta_r}}{d \log \frac{w_r}{p}} = \frac{d \log \frac{K}{L}}{d \log \frac{w_r}{p}} - 1 = \sigma_r - 1$$

Since capital markets clear nationally, p does not vary in the cross-section of locations, so that:

$$\frac{d \log \frac{\Theta_r}{1 - \Theta_r}}{d \log w_r} = \sigma_r - 1$$

and capital cost shares a lower wherever wages are higher, since $\sigma_r < 1$.

Relationship to Model in Section 4. The model presented in Section 4 nests the neoclassical model presented here as a well-defined limit when the function F_r is CES (corresponding to the case where $\epsilon = 0$ and $\zeta \to 1$). In particular, note that when we set $\epsilon = 0$ and $\zeta < 1$ in equation OA.12 of our theory, we obtain

$$\frac{d \log w_r}{d \log p} = -\frac{\zeta \Theta_r}{1 - \zeta \Theta_r},$$

additionally sending $\zeta \to 1$ recovers expression OA.17 above.

C. DATA CONSTRUCTION

C.1 Defining Commuting Zones

We assign counties to 1990 USDA ERS commuting zones as constructed by Tolbert and Sizer (1996). However, there are 11 counties that change or are added over our time period that we manually assign. We merge these counties with adjacent counties or their precursor counties. In particular, we combine Federal Information Processing Standards (FIPS) Codes 12025 with 12086, 08014 with 08013, 51780 with 51083, 30113 with 56029, 02231 with 02282, 02105 with 02282, 02230 with 02282, 02195 with 02280, 02275 with 02280, 02275 with 02280, and 02198 with 02201.[44]

In general, these are minor adjustments, with only the first three being associated with substantial population counts (the first is a subdivision of Miami-Dade County, the second involves the creation of a new county in the Denver-Boulder Metro Area, the third involves a minor subdivision of Halifax County, Virginia). The last seven adjustments all involve a complete reordering of extremely remote Alaskan commuting zones primarily in the Wrangell area.

We do not use Alaskan or Hawaiian commuting zones in our counterfactual analysis or model calibration (but include them in the national-level aggregate statistics for completeness). We are left with 722 commuting zones out of the 741 original USDA ERS commuting zones.

C.2 Price Index Data

Figure 5 relies on the BEA asset prices "Table 1.5.4. Price Indexes for Gross Domestic Product" from the FRED database available at `https://fred.stlouisfed.org/release/tables?rid=53&eid=14833`. All prices are relative to the BLS Consumer Price Index for Urban Consumers (CPI-U), available as FRED series CPI-AUCSL. We compute annual averages of the price indices for equipment capital and intellectual property and their sub-components and report them in the two panels of Figure 5. For the model, we take the ICT price index as the simple average for "Information Processing" and "Software" investment prices for 190

[44]Combining 30113 with 56029 is the only cross-state merge, attributing remote parts of Yellowstone National Park to Park County, WY. This is also popularly known as the "Zone of Death," where theoretically one could commit any crime up to and including murder without charge.

to 2015.

C.3 LBD

In processing the LBD data, we aggregate the administrative, establishment-level Longitudinal Business Database (LBD) from the US Census Bureau from 1980 to 2015. The underlying LBD reports establishment categories in different classification systems, starting with the Standard Industrial Classification (SIC) and then transitioning to the North American Classification System (NAICS) in 1997. The NAICS system has been further updated in subsequent years. We use Fort and Klimek (2016) to update historical SIC records into consistent NAICS records.

We trim outlier data, removing establishments without employment or payroll data, as well as omitting firms with mean worker pay greater than $1,000,000 per year.

We additionally exclude a small number of agricultural establishments. Coverage of NAICS 61 is sparse, as the majority of national employment is in the public sector, which is not covered by the LBD.

For 2013, we merge the LBD with the the Annual Capital Expenditures Survey (ACES) and the Information and Communication Technology Survey (ICTS) to produce spending on ICT at the firm level. We use this to produce firm-level ICT investments per employee, as in Table 3.

C.4 Census

To construct our ”Census” data set, we combine the 1970, 1980, 1990, and 2000 Decennial Censuses and the 2010 and 2015 American Community Survey files from Ruggles et al. (2015).

We drop all observations that are not in the labor force, have zero income, are employed in the government or agriculture, or are missing an industry identifier. We split workers into those with at least a college degree (”college”) and those without (”non-college”) and those in cognitive non-routine occupations (CNR) and all others (non-CNR) following Rossi-Hansberg et al. (2019).

We aggregate the data to 722 commuting zones (see Tolbert and Sizer (1996)) covering the entirety of the continental United States. To do this, we use the crosswalks by Autor and Dorn (2013) to map Census Public Use Microdata areas (PUMAs) native to the Census files to commuting zones. In 1980, the crosswalk uses the county groups in the Census data since no PUMA codes are available.

We aggregate all our data to 1-digit NAICS sectors which are designed to capture the principal functional differences between groups of industries. To do so, we create a crosswalk from the Census industry identifiers to NAICS codes using the 2000 cross-section of the data that includes both codes.

We define the average wage within a location-sector pair as the ratio of its total payroll to its total employment, using Census-provided sampling weights.

To construct a household rental price index, we regress the logarithm of household-level gross rents on the dwelling age, number of rooms, number of bedrooms, number of units in the building, and commuting-zone-year fixed effects, weighting by household sampling weights. The resulting commuting zone fixed effects serve as the rental price index for each year. We display the resulting rent price indices for 1980 and 2015 in Figure OA.5.

C.5 QCEW

For some of our aggregate wage, employment, and establishment statistics (such as Figures 2 and OA.6) , we use the publicly-available BLS Quarterly Census of Employment and Wages (QCEW). The data cover most of the US workforce and use unemployment insurance records as the source. We drop observations located in the synthetic counties designated as "Overseas Locations," "Multicounty," "Out-of-State," or "Unknown Or Undefined" and counties with a privacy disclosure flag.

Prior to 1990, the QCEW uses the SIC industry classification standard. To convert this to the modern NAICS industry standard we again use the Fort and Klimek (2016) crosswalks to the NAICS 2012 classification for the SIC 1977 codes for data from 1980-1986 and the SIC 1987 codes for 1987-1990. We make two small adjustments: we classify "SIC 1520" as a non-Business Services industry and "SIC 9999" (non-classifiable establishments) as a non-Business Services industry.

C.6 CPS

The Current Population Survey (CPS) conducted by the US Census Bureau and the BLS is used to get data on employee characteristics by firm size. We obtain a cleaned version of this dataset from IPUMS (Ruggles et al. (2015)). Since 1992 the CPS has consistently asked the size of an employer. There are six employer sizes, "<10 employees", "10-24", "25-99", "100-499", "500-999", and "1000+."[45] We drop workers who worked more than 168 hours in a week and part-time workers who work less than 30 hours in a "usual" week. We classify workers with a bachelor's degree through a doctorate degree as "college educated." All other workers, including those with an associate's degree (both academic and vocational based) are classified as those without a college degree.

For sector of employment, we use an adapted crosswalk of Fort and Klimek (2016) to map from 1990 SIC codes (which itself deviates from many Census products) to 2007 1-digit NAICS sector codes.

C.7 CBP

As a robustness exercise, we document the increase in the wage-density gradient in the US Census Bureau's County Business Patterns (CBP) database in Figure OA.11.

We perform minimal processing of the data, first aggregating counties to commuting zones following Tolbert and Sizer (1996) and then adjusting wages by CPI-U. Wages are computed as total payroll divided by the number of reported employees.

We additionally use the 2015 CBP to generate the spatial distribution of establishments by size in Figure 7. We aggregate the establishment count bins for locations with "1-4", "5-9", "10-19", and "20-49" employees to create panel (A) and "250-499", "500-999", and "1000+" employees for panel (B).

C.8 BEA Fixed Asset and Value Added Data

For Figure 6, we use data from the BEA on fixed cost capital stocks (in 2012 dollars) by industry and capital type. We compute the stock of proprietary software using codes ENS2 and ENS3, pre-packaged software with code ENS1, and hardware with codes EP1A-EP31. These data have been converted from SIC codes to consistent BEA-specific NAICS codes that we aggregate into our 1-digit NAICS sectors.

We additionally use data on value added and employee compensation from the BEA industry accounts. As data on employee compensation are only available after 1987, for 1980, we use data on employee compensation from the QCEW, for which we map SIC codes to NAICS codes using Fort and Klimek (2016).

[45] Data on employer size first started in 1988; however, the first few iteration changed the size categories of employers. The question reached its current form in 1992, so we use that as the first year.